8

The Costs and Benefits of Meaning

Morton Ann Gernsbacher, Rachel R. W. Robertson, and Necia K. Werner

In this world of text, there are many benefits to be reaped. But typically these benefits come at some cost. In this chapter, we provide empirical data demonstrating the benefits reaped and the costs incurred in the process of understanding the meaning of a sentence. We use as our empirical tool ambiguous words, which have enjoyed a rich past of empirical inquiry (Foss, 1970; Foss & Jenkins, 1973; Hogaboam & Perfetti, 1975; Lackner & Garrett, 1972; Perfetti & Goodman, 1970; Swinney & Hakes, 1976). One of the most interesting questions in the study of lexical ambiguity concerns what happens to alternative homonym meanings after one meaning is selected. For example, in the sentence *She lit the match*, what happens to the contest meaning of *match* when the firestick meaning is selected? Simpson and Kang (1994) recognized this as one question that is sadly underrepresented in the literature.

In this chapter, the emphasis is on six experiments designed to explore the facets of this very question. The theory we focus on in these experiments is that the selection of one meaning (such as the firestick meaning in *She lit the match*) leads to an immediate and active suppression of other meanings (Faust & Gernsbacher, 1996; Gernsbacher, 1990; Gernsbacher & Robertson, 1995). This theory stems from what Gernsbacher called the *structure building framework*. Over the last two decades, Gernsbacher (1990, 1991a, 1995, 1997c) posited that the goal of language comprehension is to build a coherent structure. In this structure-building process, the reader first lays a foundation. Related information gets mapped onto the structure, and unrelated information is shifted into a new substructure.

Two general cognitive mechanisms, suppression and enhancement, play a role in building this structure (Gernsbacher & Faust, 1991a, 1991b). The mechanism of suppression *dampens* the activation of information no longer necessary or relevant, and the mechanism of enhancement *boosts* the activation of infor-

This research was supported by Grants R01 NS 29926 from the National Institutes of Health and DASW0194-K-004 from the Army Research Institute. The first four experiments were presented in November 1994 at the 35th annual meeting of the Psychonomic Society in St. Louis, Missouri. Experiment 5 was presented in April 1999 at the 71st annual meeting of the Midwestern Psychological Association in Chicago. We are grateful to David Gorfein and James Erickson for their helpful comments on the manuscript and to David Balota for raising the question that inspired Experiment 6.

mation that is relevant. These suppression and enhancement signals could be triggered by the configural or, in St. John's (1992) term, the gestalt-level representation of a phrase, sentence, passage, or other meaningful unit. In relation to the costs and benefits of meaning, the suppression of irrelevant information could potentially cause a cost to the reader (i.e., slow down comprehension) if the suppressed information needs to become accessible. Conversely, the enhancement of relevant information could be beneficial (i.e., speed up comprehension) if the reader must comprehend information that has already been enhanced.

On the basis of the idea of suppression and enhancement, we assumed that participants would find the following pair of ambiguous sentences relatively easy to comprehend:

He lit the match. (same-meaning prime)
He blew out the match. (target)

The idea here is that the firestick meaning of *match* becomes enhanced when the prime sentence is read, and so participants would be relatively fast to comprehend the target sentence (thus reaping a benefit). Conversely, we also assumed that the following pair of sentences would be more difficult to comprehend:

He lit the match. (different-meaning prime)
He won the match. (target)

In this case, the firestick meaning of *match* becomes enhanced when the prime sentence is read. As the contest meaning of *match* has been suppressed, participants should be slower to comprehend the target sentence (thus incurring a cost).

In this chapter, six experiments examined the costs incurred and the benefits reaped when reading pairs of sentences of this type. In the first four experiments, costs and benefits were quantified from a neutral baseline (e.g., *He saw the match*) and a no-meaning baseline (e.g., *He prosecuted the match*), as a function of participants' comprehension skill and the distance between the prime and target sentences. The final two experiments explored the role of these costs and benefits as measured by the effect of dominant, subordinate, and no-meaning primes on dominant and subordinate targets. We attempt to explain the results of all six experiments in terms of the mechanisms of suppression and enhancement.

Experiment 1

Sixty University of Wisconsin–Madison undergraduates participated in the experiment for extra credit in an introductory psychology course. In all the experiments reported here, all participants were native English speakers. We constructed 24 target sentences, each of the form pronoun, verb, noun phrase, with the final noun phrase ending in a homonym (e.g., *She blew out the match*). For each of the 24 target sentences, we constructed a same-meaning prime sentence (*She lit the match*), a different-meaning prime sentence (*She won the match*), and a neutral-meaning prime sentence (*She saw the match*). The three prime sentences and their associated target sentence differed only by their verb. We also constructed 336 filler sentences, similar in length and style to the experimental sentences

Therefore, our list of 384 total sentences comprised 48 experimental sentences (arranged as 24 pairs of experimental sentences) and 336 filler sentences (arranged as 168 pairs of filler sentences). Both sentences in each of the 24 pairs of experimental sentences made sense, and both sentences ended in the same homonym. Of the 168 pairs of filler sentences, 24 pairs ended in the same homonym, and neither sentence made sense; 24 pairs ended in the same homonym and only the first sentence made sense; 24 pairs ended in the same homonym and only the second sentence made sense; 24 pairs ended in two different homonyms and both sentences made sense; 24 pairs ended in two different homonyms and neither sentence made sense; 24 pairs ended in two different homonyms and only the first sentence made sense; and 24 pairs ended in two different homonyms and only the second sentence made sense.

We created three material sets by randomly assigning one of the three versions of each experimental prime sentence to one of the three material sets. An equal number of same-meaning, different-meaning, and neutral prime sentences occurred within each material set, and each prime sentence occurred in each of its three versions across the three material sets. The target and filler sentences remained the same across the three material sets. The complete set of materials for all experiments reported here is available on-line at http://psych.wisc.edu/lang/material.html.

The sentences were presented using a computer. Participants were asked to read each sentence and decide rapidly and accurately whether it made sense by pressing either the "yes" or "no" button on the response pads. The testing facilities and procedures were identical for all experiments reported here.

Table 8.1 displays participants' average response times to target sentences after reading same-, neutral-, and different-meaning primes. In all the experiments reported here, participant data were included only if the prime sentence and its subsequent target sentence were responded to correctly. In this experiment, the benefits were calculated by subtracting participants' response times (in ms) to targets preceded by same-meaning primes from the response times to targets preceded by neutral-meaning primes. The costs were calculated by subtracting participants' response times to targets preceded by neutral-meaning primes from the response times to targets preceded by different-meaning primes. We found that participants reaped a statistically reliable 88-ms benefit (MSE = 231,629.71), $F(1, 59) = 10.72$, $p < .001$, and incurred a statistically reliable 86-ms cost (MSE = 221,996.98), $F(1, 59) = 10.27$, $p < .002$ (i.e., participants were faster to say that a target made sense if it was preceded by a same-meaning prime and were slower to make this judgment if the target was preceded by a different-meaning prime).

The costs and benefits of meaning were also quantified using participants' percentage error rates on the target sentences, as shown in Table 8.1. As with the response times, we subtracted participants' error rates to the target sentences following same-meaning primes from their response times to target sen-

Table 8.1. Participants' Data From Experiment 1

Target sentence	Response time to target (ms)	% Error rate to target	*SD* (ms)
She blew out the match			
After SAME meaning primes:	1,107 (29)	11 (1%)	330
(She lit the match)			
After NEUTRAL meaning primes:	1,195 (35)	17 (2%)	388
(She saw the match)			
After DIFFERENT meaning primes:	1,281 (41)	25 (2%)	408
(She won the match)			

Note: Numbers in parentheses represent the standard errors of the mean.

tences following neutral-meaning primes and quantified a statistically reliable 6% benefit (MSE = 7.50), $F(1, 59) = 8.08, p < .005$. By subtracting error rates on the target sentences following neutral-meaning primes from error rates on the target sentences following different-meaning primes, we calculated a statistically reliable 8% cost (MSE = 10.80), $F(1, 59) = 11.63, p < .001$. These costs and benefits can be seen in Figure 8.1, which represents participants' average response times in milliseconds and percentage error rates to target sentences.

Based on the mechanisms of suppression and enhancement, one could assume that after participants comprehend the sentence *She lit the match*, the firestick meaning of match is enhanced, and the contest meaning is suppressed. This would explain the benefit that participants received when reading a target sentence preceded by a same-meaning prime: The appropriate meaning of the homonym was already enhanced and easily accessible. Similarly, after participants comprehend the sentence *She won the match*, the contest meaning of match is enhanced, and the firestick meaning is suppressed. This also explains the costs that were incurred when participants read a target sentence preceded by a different-meaning prime: The meaning that they needed was suppressed and not readily available.

But what happens when participants comprehend the neutral-prime sentence, *She saw the match*? One possibility is that neither the firestick meaning nor the contest meaning is enhanced or suppressed. In this case, participants would receive neither a benefit nor a cost from reading a target sentence preceded by a neutral prime, which would explain the trend we see in Figure 8.1.

An alternative explanation is that for each neutral sentence prime, approximately half of our participants enhanced one meaning and the other half enhanced the other meaning. The net result would be that half reap a benefit and half incur a cost. The average response times and error rates to the target sentences following the neutral primes would be halfway between the response times and error rates in the same- versus different-meaning conditions, which is what we observed.

However, if, when reading a target sentence after reading a neutral prime, half the participants were reaping a benefit and half were incurring a cost, then the neutral prime condition should be associated with more variance. As illus-

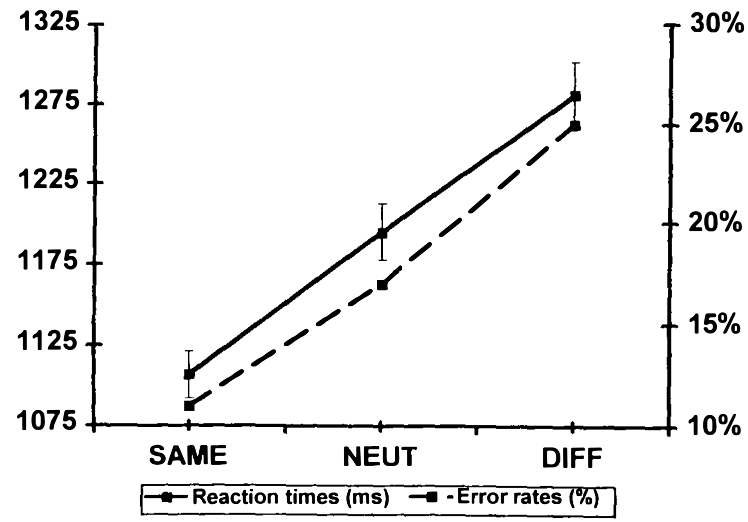

Figure 8.1. Average response times (in ms) and average error rates to target sentences in Experiment 1. The solid line represents average response times (in ms), with the scale on the left, and the dashed line represents average error rates, with the scale on the right. SAME = same-meaning prime; NEUT = neutral-meaning prime; DIFF = different-meaning prime.

trated in Table 8.1, the average of each participant's standard deviation of his or her response time was not larger in the neutral than in the same- or different-meaning prime conditions. The same was true when we investigated each item's standard deviation. This supports the claim that neither meaning is enhanced or suppressed when participants read neutral-meaning primes.

Experiment 2

In our second experiment, we used a corroborating baseline that also demonstrated the costs and benefits of meaning. In this experiment, we again preceded each target sentence by one of three primes, but nonsense primes were used as the baseline instead of neutral-meaning primes. Because neither meaning of the homonym can be inferred from a nonsense prime, neither meaning should be enhanced or suppressed. We should therefore see results identical to those in Experiment 1, in which we predicted that participants were neither enhancing nor suppressing either meaning of the sentence-final homonym in the neutral prime sentences.

The participants were 75 University of Wisconsin–Madison undergraduates who participated for extra credit and members of the Madison community who participated for monetary compensation. We used the same 24 target sentences, 24 different-meaning prime sentences, and 24 same-meaning prime sentences from Experiment 1. To replace the neutral prime sentences, we constructed 24 no-meaning prime sentences (e.g., *She prosecuted the match*). Because this meant that there were 24 new sentences for which the correct response to the task question was "no," we included 24 additional filler sentences for which the correct response was "yes" to balance the total number of sentences that did and did not make sense. In a paper-and-pencil norming task, at least 90% of the participants confirmed that each no-meaning prime sentence did indeed not make sense, and when asked to write what meaning the sentence-final word could imply, no one individual meaning was selected for each no-meaning sentence by more than 60% of the participants.

Our list of 256 sentences comprised 48 experimental sentences (24 pairs) and 208 filler sentences (104 pairs). Of the 24 experimental pairs, 16 ended in the same homonym and both sentences made sense and 8 ended in the same homonym and only the second sentence made sense. Of the 108 pairs of filler sentences, 16 pairs ended in the same homonym and neither sentence made sense; 16 pairs ended in the same homonym and only the first sentence made sense; 8 pairs ended in the same homonym, and only the second sentence made sense; 16 pairs ended in two different homonyms and both sentences made sense; 16 pairs ended in two different homonyms and neither sentence made sense; 16 pairs ended in two different homonyms and only the first sentence made sense; and 16 pairs ended in two different homonyms and only the second sentence made sense. As described in Experiment 1, three material sets were created from these target and filler pairs.

Figure 8.2 displays the participants' average response times and error rates in Experiment 2. Responses to the target sentences in the no-meaning condition were used as a baseline. In this experiment, as in all the experiments that follow, the benefits were calculated by subtracting participants' response times (ms) and error rates (%) to targets preceded by same-meaning primes from the response times and error rates to targets preceded by no-meaning primes. The costs were calculated by subtracting participants' response times and error rates to targets preceded by no-meaning primes from the response times and error rates to targets preceded by different-meaning primes. In this experiment, participants reaped a reliable 66 ms – 7% benefit, [(MSE = 163,772.01), $F(1, 74)$ = 8.26, $p < .005$ – (MSE = 11.76), $F(1, 74)$ = 8.53, $p < .004$], and incurred a reliable 127 ms – 8% cost [(MSE = 602,866.80), $F(1, 74)$ = 30.41, $p < .0001$ – (MSE = 16.01), $F(1, 74)$ = 11.61, $p < .001$].

These results reiterate what was found in Experiment 1: Participants find it more difficult to judge if a target sentence makes sense if it is preceded by a different-meaning prime, and thus incur a cost. Participants also find it easier to make that decision if the target is preceded by a same-meaning prime, and thus reap a benefit. The results also support our claim that in the baseline condition, participants are neither enhancing nor suppressing either meaning of the homonym.

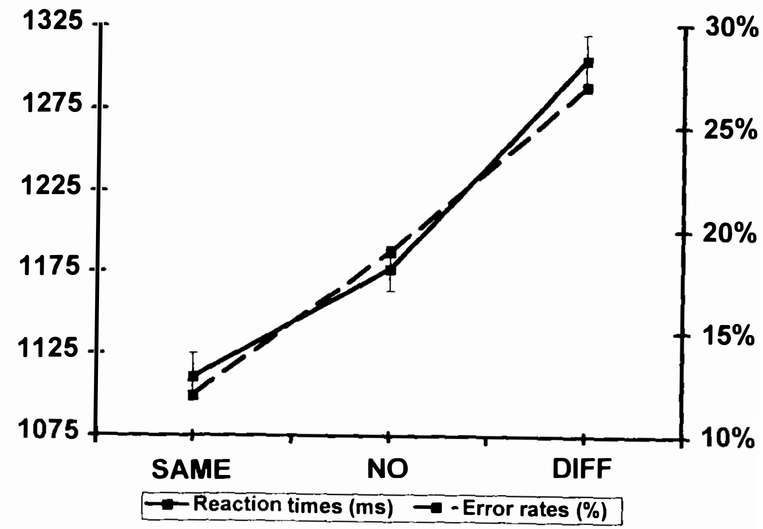

Figure 8.2. Average response times (in ms) and average error rates to target sentences in Experiment 2. The solid line represents average response times (in ms), with the scale on the left, and the dashed line represents average error rates, with the scale on the right. SAME = same-meaning prime; NO = no-meaning prime; DIFF = different-meaning prime.

Experiment 3

Previous literature has suggested that less-skilled comprehenders have a less-efficient suppression mechanism than skilled comprehenders and so have greater difficulty in suppressing irrelevant information (Faust, Balota, Duchek, Gernsbacher, & Smith, 1997; Gernsbacher, 1993, 1997b; Gernsbacher & Faust, 1991a, 1995; Gernsbacher, Varner, & Faust, 1990). More specifically, it has been suggested that less-efficient comprehenders have greater difficulty suppressing the incorrect meaning of a homophone (Gernsbacher & Faust, 1991b; Gernsbacher & Robertson, 1995). Experiment 3 explored the differences that such difficulties make on the costs incurred and benefits gained by more- versus less-skilled readers while reading target sentences preceded by same-, different-, or no-meaning prime sentences.

The participants were 72 University of Wisconsin–Madison undergraduates who participated for extra credit and members of the Madison community who participated for monetary compensation. Half were classified as more-skilled readers and half were classified as less-skilled readers on the basis of their performance on the reading component of the Multi-Media Comprehension

Battery (Gernsbacher & Varner, 1988), which was administered before the sentence verification portion of the task. The 36 more-skilled readers scored higher than 75% on the comprehension test, and the 36 less-skilled readers scored lower than 70% on the comprehension test. The mean performance of the 36 more-skilled readers was 81% correct (*SD* = 5%), and the mean performance of the 36 less-skilled readers was 64% correct (*SD* = 8%). The materials and design were identical to those in Experiment 2.

The results for Experiment 3 appear in Figure 8.3. The more-skilled comprehenders, whose data are represented by the bottom two lines, reaped a reliable 70 ms – 6% benefit [(*MSE* = 87,025.90), $F(1, 35) = 5.07$, $p < .03$ – (*MSE* = 5.01), $F(1, 35) = 4.67$, $p < .04$], and they incurred a reliable 118 ms [(*MSE* = 250,896.04), $F(1, 35) = 14.63$, $p < .0003$; although their 5% cost was not statistically reliable]. The less-skilled comprehenders, whose data are represented by the top two lines, also reaped a reliable 96 ms – 9% benefit [(*MSE* = 167,663.83), $F(1, 35) = 9.22$, $p < .004$ – (*MSE* = 8.00), $F(1, 35) = 5.15$, $p < .03$]; however, they incurred very few costs. The 10 ms – 1% cost was not statistically reliable [(*MSE* = 1,865.54), $F < 1$ – (*MSE* = .347), $F < 1$]. These data support the hypothesis that the costs incurred by reading a different-meaning prime are due to suppression; less-skilled comprehenders, who are characterized by less-efficient suppression mechanisms, are less able to get rid of the inappropriate meaning of a homonym than comprehenders with better suppression mechanisms.

Thus, less-skilled comprehenders reap the same benefits as the more-skilled comprehenders but do not incur the same costs. These data speak against the claim raised in this volume by Perfetti and Hart (chapter 5) that previous demonstrations of less-skilled comprehenders' worse suppression is due to poor lexical access (or lexical representation). If that were the case, then the less-skilled comprehenders in Experiment 3 should have shown deficits in what we call enhancement (i.e., they should have reaped weaker benefits for having read a same-meaning prime). However, they reaped benefits equivalent to those reaped by more-skilled comprehenders, suggesting that at least by this assay, their challenges with what we call suppression are not attributable to weaker lexical representations.

Experiment 4

In a fourth experiment, we asked: How lasting are the costs and benefits that we observe in the present experiments? We explored this question by presenting half the prime sentences immediately before their respective target sentences, as we had done before, and presenting half the prime sentences five sentences immediately before their respective target sentences. The answer to this question could help us adjudicate between episodic retrieval explanation of the costs and benefits that we have (Lowe, 1998; Neill & Valdes, 1992; Neill, Valdes, Terry, & Gorfein, 1992) and a suppression and enhancement explanation. According to an episodic retrieval explanation, there is no suppression or enhancement of prime information; rather a memory trace between the prime and target is formed. If the memory trace conflicts, there is slowed response time; if it is compatible, there is speeded response time. Thus, episodic retrieval

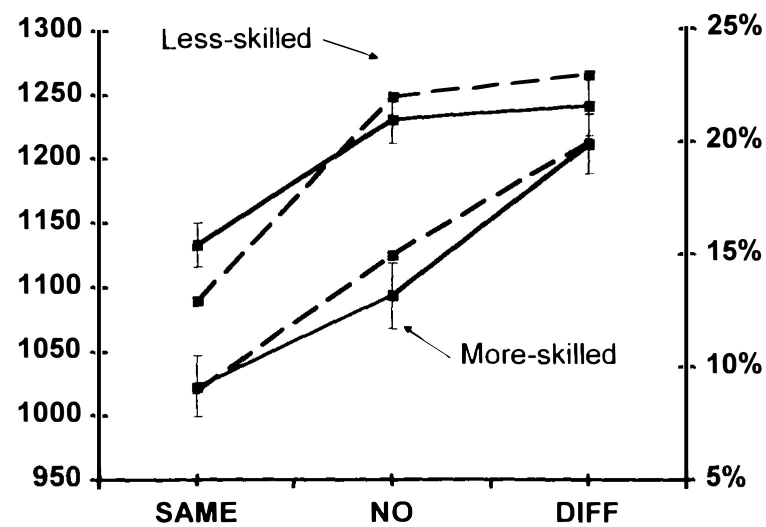

Figure 8.3. Average response times (in ms) and average error rates to target sentences by more- versus less-skilled comprehenders in Experiment 3. The solid lines represent average response times (in ms), with the scale on the left, and the dashed lines represent average error rates, with the scale on the right. SAME = same-meaning prime; NO = no-meaning prime; DIFF = different-meaning prime.

is a single-mechanism account rather than a dual-mechanism account (as in suppression and enhancement).

Although episodic retrieval explanations are usually pit against only inhibition accounts (e.g., Conway, 1999; Hasher, Zacks, Stoltzfus, Kane, & Connelly, 1996; Houghton & Tipper, 1994; May, Kane, & Hasher, 1995; Tipper, 1985; Tipper & Cranston, 1985; Tipper & Milliken, 1996), in this chapter we have advocated two mechanisms: suppression and enhancement. Our interest in examining the time course of the costs and benefits of meaning was to see if other variables would dissociate our putative mechanisms of suppression. If so, a single mechanism explanation of these effects seems less tenable.

The participants were 180 University of Wisconsin–Madison undergraduates who participated for extra credit and members of the Madison community who participated for monetary compensation. The materials were the same as those used in Experiment 2. We altered the design of the experiment by adding an additional manipulation. Half of the experimental prime sentences were presented five sentences before their respective target sentence, as illustrated in Table 8.2. The other half of the experimental prime sentences were presented immediately preceding their respective targets.

Table 8.2. Example of Experimental Materials in Experiment 4

Immediately before	Five before
She fed the cap.	She fed the cap.
She dressed the letter.	She dressed the letter.
She disliked the perch.	*She lit / saw / won the match*
She arrested the perch.	She disliked the perch.
She shuddered the hood.	She arrested the perch.
She stacked the deck.	She shuddered the hood.
She lit / saw / won the match.	She stacked the deck.
She blew out the match.	*She blew out the match.*

We used the master list from Experiment 1 and randomly assigned 12 of the experimental pairs (four same-meaning primes, four different-meaning primes, and four neutral primes) to be separated pairs. From this altered master list, we created a total of six material sets. The first three material sets were created by randomly assigning one of the three versions of each experimental prime sentence to one of the three material sets (as in Experiments 1 and 2). The pairs that were separated remained constant for the first three material sets. The next three material sets were the same as the first three except the separated sentence pairs from the first three material sets became immediate pairs in the last three material sets, and vice versa. There were 12 fillers that ended in the same homonym and that were separated by four intervening sentences.

The results, displayed in Figure 8.4, demonstrated that when the prime sentences immediately preceded the target sentences, participants reaped a reliable 88 ms – 8% [(*MSE* = 699,293.95), $F(1, 179) = 7.98$, $p < .005$ – (*MSE* = 7.51), $F(1, 179) = 14.63$, $p < .0002$] benefit and incurred a reliable 77 ms – 4% cost [(*MSE* = 526785.38), $F(1, 179) = 6.01$, $p < .01$ – (*MSE* = 2.669), $F(1, 179) = 5.20$, $p < .02$]. When the prime sentences preceded the target sentence by five sentences, participants also reaped a reliable 53 ms – 7% benefit [(*MSE* = 252,832.25), $F(1, 179) = 3.615$, $p < .04$ – (*MSE* = 7.511), $F(1, 179) = 14.63$, $p < .0002$]; however, they did not incur a cost when measured by either their response time or their error rate (the 5 ms – 2% cost was not statistically significant) [(*MSE* = 1935.19), $F < 1$ – (*MSE* = .544), $F(1, 179) = 1.06$, $p < .30$]. This experiment demonstrates another situation in which these costs and benefits are dissociated: The costs are relatively short-lived, but the benefits last longer.

Indeed, the benefits, but not the costs, persist over more intervening sentences. In a further experiment, the primes either preceded the targets by seven sentences or were presented immediately before the targets. As we observed before, when the targets were presented immediately after the primes, participants reaped a reliable 53 ms – 5% benefit [(*MSE* = 241,073.06), $F(1, 179) = 3.88$, $p < .04$ – (*MSE* = 4.23), $F(1, 179) = 10.00$, $p < .002$] and a reliable 123 ms – 5% cost [(*MSE* = 1,377,121.18), $F(1, 179) = 22.16$, $p < .0001$ – (*MSE* = 2.67), $F(1, 179) = 6.32$, $p < .01$]. When the targets were presented seven sentences after the primes, participants continued to reap a reliable 83 ms – 8% benefit [(*MSE* = 621,518.749), $F(1, 179) = 10.00$, $p < .002$ – (*MSE* = 9.025), $F(1, 179) = 21.37$, $p <$

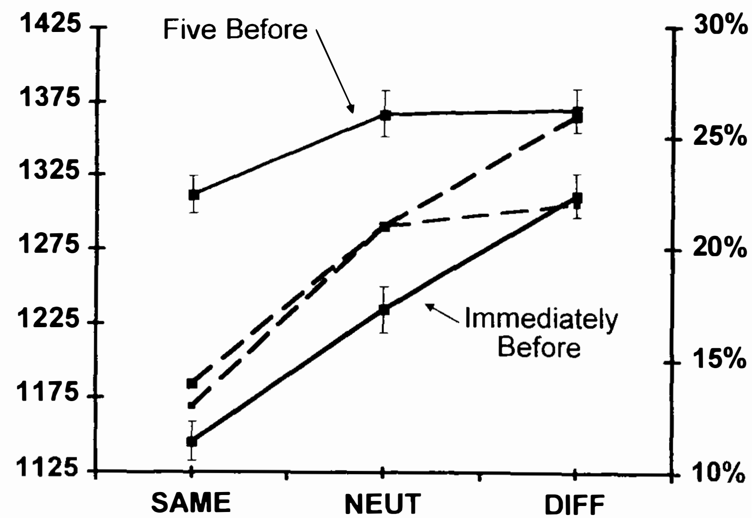

Figure 8.4. Average response times (in ms) and average error rates to target sentences in Experiment 4, in which the target followed the prime sentence immediately or five sentences later. The solid lines represent average response times (in ms), with the scale on the left, and the dashed lines represent average error rates, with the scale on the right. SAME = same-meaning prime; NEUT = neutral-meaning prime; DIFF = different-meaning prime.

.0001] but failed to incur a cost (the 39 ms – 3% cost was unreliable) [(*MSE* = 135,767.81), $F(1, 179) = 2.18$, $p < .14$ – (*MSE* = .711), $F(1, 179) = 1.68$, $p < .20$]. This replication experiment confirms our conclusion that the benefits we observed are relatively long-lived, whereas the costs are not.

Experiment 5

Although participants may recognize the multiple meanings of a homonym based on the context in which it is presented, these meanings are not necessarily used with equal frequency in everyday language. One meaning of a homonym may be considered the "dominant" meaning (or the meaning that is most often thought of when a homonym is seen in isolation), and the other, less-common meanings, may be thought of as "subordinate." Recent homonym literature has placed a heavy emphasis on the importance of weighing these disparate meanings and recognizing the profound effect they can have on how quickly participants comprehend an ambiguous sentence (Gorfein, Berger, & Bubka, 2000; Simpson, 1994; Simpson & Kang, 1994; Simpson & Krueger, 1991). Thus, the final two experiments reported here focus on the dissocations of costs and benefits in relation to prime and target sentences that end in dominant and subordinate homonyms (whereas our previous experiments used homonyms that were relatively balanced in their two meanings).

One hundred and twenty University of Wisconsin–Madison undergraduates participated for extra credit in an introductory psychology class. The dominant and subordinate meanings of various homonyms were verified in a paper-and-pencil norming task. Only homonym meanings that were defined by at least 60% of participants were considered to be dominant, and only meanings defined by 40% or fewer were considered subordinate. If more than one subordinate meaning for a homonym was generated, we chose the meaning with the highest percentage less than or equal to 40% as the subordinate.

For each of 36 homonyms that met our criteria in the word norming task, we then constructed six two-sentence sets: a dominant-meaning prime with a dominant-meaning target (Same-Dom), a subordinate prime with a subordinate target (Same-Sub), a subordinate prime with a dominant target (Diff-Dom), a dominant prime with a subordinate target (Diff-Sub), a nonsensical prime with a dominant target (Non-Dom), and a nonsensical prime with a subordinate target (Non-Sub). These pairs were pseudorandomly organized across six conditions so that participants saw only one pair of sentences for any one homonym.

We also constructed 60 pairs of filler sentences also ending in homonyms. Fillers comprised three different types of pairs: two nonsensical sentences (a no/no pair), a nonsensical prime followed by a target that made sense (a no/yes pair), or a prime that made sense followed by a nonsensical target (a yes/no pair). All participants saw all 24 no/no, 12 no/yes, and 24 yes/no filler pairs.

Before the final material sets were constructed, the sentences were rated in a second paper-and-pencil norming task. The newly created experimental and filler sentences were pseudorandomly listed, and participants were asked to decide whether each sentence made sense. Participants were also asked to write what meaning the sentence-final word could imply (they were asked to do so even for the nonsense sentences). In the norming task, at least 95% of our participants confirmed that the sentences with the subordinate and dominant meanings of the homonym made sense and that the nonsensical sentences did not. When asked to write what meaning the sentence could imply, no one individual meaning was selected for each nonsense sentence by more than 60% of the participants. Also, according to our criteria, at least 95% of participants had to agree on one particular meaning before we included a dominant or subordinate sentence in the final materials.

The six conditions were constructed from the experimental and filler sentences that met the norming criteria. So, in any one condition, participants saw 6 of each of the Same-Dom, Same-Sub, Diff-Dom, Diff-Sub, Non-Dom, and Non-Sub experimental pairs, with the 60 filler pairs randomly distributed among the experimental pairs.

Figure 8.5 represents participants' reaction times and error rates to subordinate-meaning target sentences, depending on whether they were preceded by a same-, different-, or no-meaning prime sentence. When subordinate targets were preceded by a same-meaning (subordinate) prime, participants reaped a reliable 176 ms – 5% benefit [(*MSE* = 1,863,843.75), $F(1, 119) = 30.39$, $p < .0001$

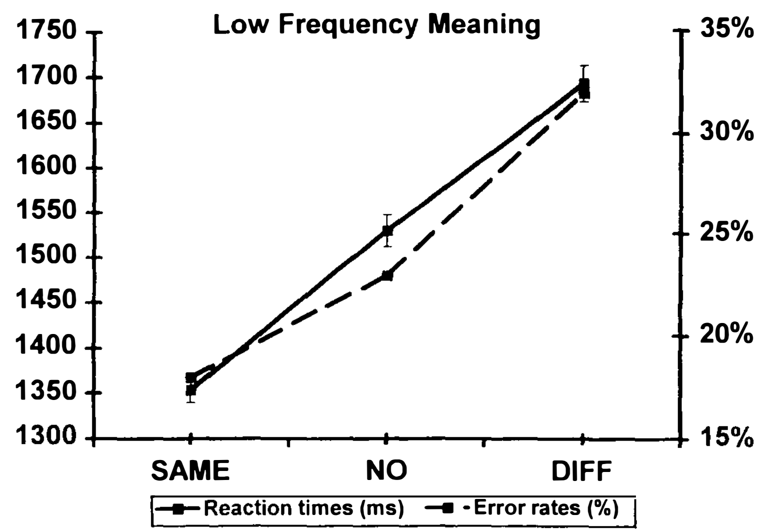

Figure 8.5. Average response times (in ms) and average error rates to target sentences, following a subordinate-meaning prime, in Experiment 5. The solid lines represent average response times (in ms), with the scale on the left, and the dashed lines represent average error rates, with the scale on the right. SAME = same-meaning prime; NO = no-meaning prime; DIFF = different-meaning prime.

– (*MSE* = 0.176), $F(1, 119) = 9.11$, $p < .003$]. Conversely, when subordinate targets were preceded by a different-meaning (dominant) prime, participants incurred a reliable 166 ms – 9% cost [(*MSE* = 1,657,014.02), $F(1, 119) = 27.02$, $p < .0001$ – (*MSE* = 0.474), $F(1, 119) = 24.54$, $p < .0001$].

Figure 8.6 represents participants' reaction times and error rates to dominant-meaning target sentences, depending on whether they were preceded by a same-, different-, or no-meaning prime sentence. When dominant targets were preceded by a different-meaning (subordinate) prime, participants continued to incur a reliable 170 ms – 15% cost [(*MSE* = 1,738,252.61), $F(1, 119) = 28.34$, $p < .001$ – (*MSE* = 1.45), $F(1, 119) = 75.15$, $p < .0001$]. Contrary to what we saw with the subordinate targets, participants reaped no benefit at all when the dominant targets were preceded by a same-meaning (dominant) prime (a statistically unreliable 18 ms – 2% benefit), [(*MSE* = 18,462.60), $F < 1$ – (*MSE* = 0.023), $F(1, 119) = 1.17$, $p < .28$].

Experiment 6

After the results for Experiment 5 were presented at the Midwestern Psychological Association meetings in 1999, David Balota suggested that perhaps the

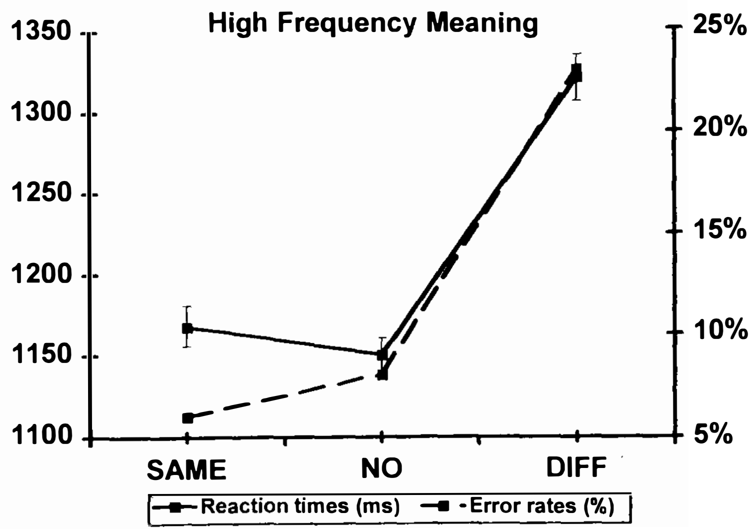

Figure 8.6. Average response times (in ms) and average error rates to target sentences, following a dominant-meaning prime, in Experiment 5. The solid lines represent average response times (in ms), with the scale on the left, and the dashed lines represent average error rates, with the scale on the right. SAME = same-meaning prime; NO = no-meaning prime; DIFF = different-meaning prime.

failure to find a benefit from enhancement with the high-frequency meanings may not have been because of the frequency per se, but because the target sentences in which they occurred were more tightly constrained. In other words, perhaps the dominant-meaning target sentences provided more context than did the subordinate-meaning target sentences. This seemed doubtful, as in the original materials we deliberately avoided using explicit sentences such as *She withdrew money from the bank*, and instead used more context-neutral sentences such as *She closed the bank*. In addition, subordinate- or dominant-meaning sentences were not included in the materials unless 95% of participants could give the intended meaning in the norming task conducted prior to the reading task in Experiment 5. This disagreement turned into a high-stakes ($5) bet from which the following rating task was born.

Balota suggested that we gather all of our prime, target, and filler sentences (including the nonsense filler sentences) and ask participants simply to rate on a scale of 1 to 5 how well the sentence-final homonym fit the meaning of the sentence. Then, by averaging the responses for each sentence, we could see if the ratings for the dominant-meaning sentences (e.g., *She applauded the boxers*) differed from the ratings for the subordinate-meaning sentences (e.g., *She folded the boxers*).

Forty University of Wisconsin–Madison undergraduates completed this rating task for extra credit in an introductory psychology course. Much to our chagrin, participants' average overall rating on the 5-point scale for the dominant-meaning prime sentences was 4.68 (*SE* = 0.30), whereas subordinate-meaning prime sentences were rated on average at 4.21 (*SE* = 0.56). Dominant-meaning target sentences were rated at 4.61 (*SE* = 0.30), and subordinate-meaning target sentences were rated at 4.13 (*SE* = 0.56). What looks here to be numerically small differences are actually supported by *F* values close to 20. Indeed, the original sentences for the prior experiment differed; the dominant-meaning homonyms in both the prime and target sentences were better fits in their respective contexts. The $5 was humbly bestowed on the deserving winner, and one final response time experiment was conducted: Would we still find a lack of benefit from enhancement with the high-frequency meanings, even if all the sentences were rated as contextually equivalent on the 5-point scale?

One hundred and twenty University of Wisconsin–Madison undergraduates participated for extra credit in an introductory psychology class to help us find out. The goal of this experiment was to replicate Experiment 5 with sentences that had relatively equivalent means on the previously described 5-point rating scale. To achieve this, we "tightened" the offending low-frequency sentences so that they were more contextually constrained and "loosened" the more contextually obvious high-frequency sentences. Thus, subordinate-meaning sentences such as *She consulted with the cabinet* were replaced with *She was a member of the cabinet*, which changed the rating from 3.80 to 4.88. Interestingly, even some seemingly minor alterations, such as modifying *She appointed the cabinet* to *She was appointed to the cabinet* changed the rating from 4.25 to a surprising 4.80. Similarly, rewriting the contextually constrained dominant-meaning sentence *She inflated the ball* to the less-obvious *She repaired the ball* changed the rating from 4.88 to 4.04.

We rewrote a total of 53% of the experimental sentences from Experiment 5 and replaced approximately equal numbers of high- and low-frequency prime and target sentences (28% dominant primes, 22% dominant targets, 21% subordinate primes, and 24% subordinate targets). The old sentences (minus the sentences that were removed) and the newly created replacement sentences were pseudorandomly listed, and participants were again asked to rate, on a scale of 1 to 5, how well the meaning of the sentence-final word fit the overall meaning of the sentence. The new means for the high- and low-frequency primes and targets proved to be relatively equivalent: dominant primes, 4.54 (0.37); dominant targets, 4.46 (0.37); subordinate primes, 4.47 (0.42); subordinate targets 4.53 (0.30; the numbers in parentheses are the standard errors of the means).

Figure 8.7 represents participants' reaction times and error rates to subordinate-meaning target sentences, depending on whether they were preceded by a same-, no-, or different-meaning prime sentence. When subordinate-meaning target sentences were preceded by a same-meaning (subordinate) prime, participants reaped a reliable 60-ms benefit [(MSE = 316,245.60), $F(1, 119) = 6.526$, $p < .01$, although the benefit in error rate failed to reach reliability]. Conversely, when subordinate targets were preceded by a different-meaning (dominant) prime, participants incurred a reliable 153 ms – 12% cost [(MSE = 1,085,146.02), $F(1, 119) = 22.392$, $p < .0001$ – (MSE = 22.82), $F(1, 119) = 36.99$, $p < .0001$].

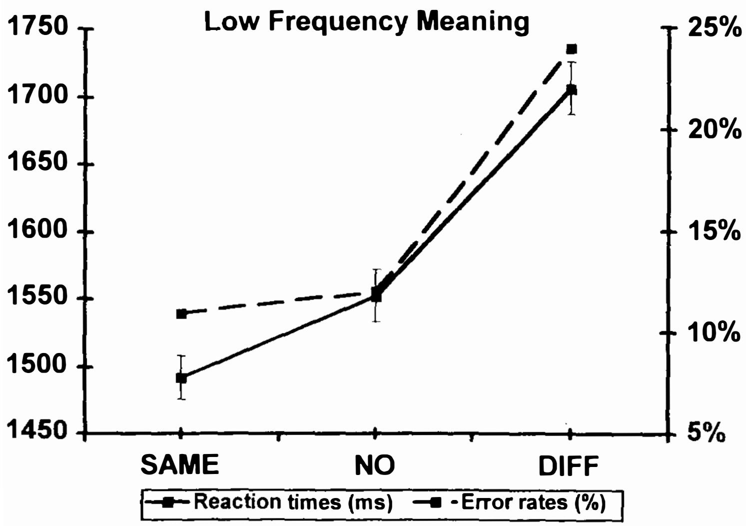

Figure 8.7. Average response times (in ms) and average error rates to target sentences, following a subordinate-meaning prime, in Experiment 6. The solid lines represent average response times (in ms), with the scale on the left, and the dashed lines represent average error rates, with the scale on the right. SAME = same-meaning prime; NO = no-meaning prime; DIFF = different-meaning prime.

Figure 8.8 represents participants' reaction times and error rates to dominant-meaning target sentences, depending on whether they were preceded by a same-, different-, or no-meaning prime sentence. When dominant targets were preceded by a different-meaning (subordinate) prime, participants continued to incur a reliable 233 ms – 12% cost [(*MSE* = 2,942,849.07), $F(1, 119) = 60.73$, $p < .0001$ – (*MSE* = 28.704), $F(1, 119) = 46.53$, $p < .0001$]. And again, contrary to what we observed with the subordinate targets, participants reaped no benefit at all when the dominant targets were preceded by a same-meaning (dominant) prime (a statistically unreliable –1 ms – 0% benefit) [(*MSE* = 24,220.50), $F(1, 119) = .5$, $p < .48$ – (*MSE* = 0.017), $F(1, 119) = .027$, $p < .87$].

Thus, Experiment 6 replicated Experiment 5, despite the better-matched sentence contexts. We note, however, the difference in normative (subjective) data we obtained when participants are asked to choose which meaning a sentence-final homonym conveys as opposed to rate the "fit" of a sentence context. But from our response time data, we conclude that when a different-meaning prime precedes a target sentence, the sentence-final homonym is suppressed; therefore, we see reliable costs in both cases. Conversely, when a same-meaning prime precedes a target sentence, the sentence-final homonym is already highly activated; thus, the benefits are negligible.

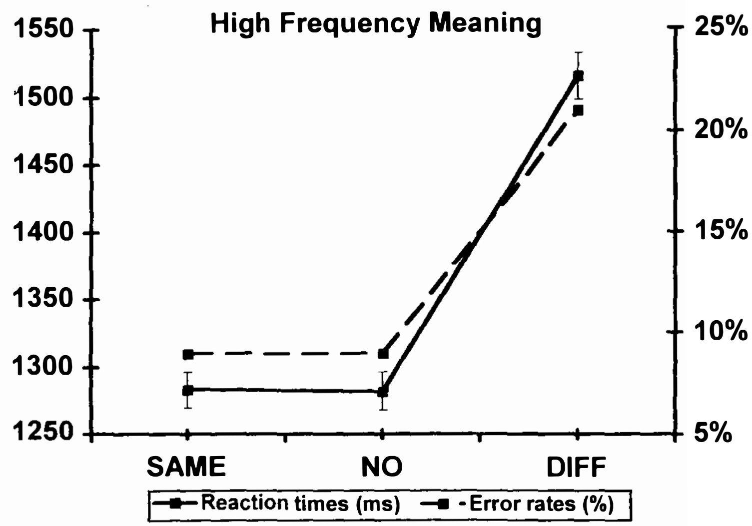

Figure 8.8. Average response times (in ms) and average error rates to target sentences, following a dominant-meaning prime, in Experiment 6. The solid lines represent average response times (in ms), with the scale on the left, and the dashed lines represent average error rates, with the scale on the right. SAME = same-meaning prime; NO = no-meaning prime; DIFF = different-meaning prime.

In sum, the present experiments explored the costs incurred and benefits reaped during the comprehension of ambiguous target sentences when they are preceded by ambiguous prime sentences and demonstrated the role of suppression and enhancement in the dissociation of these costs and benefits.

We argue that the dissociations we have seen here are not simply due to a ceiling effect. If we look at the slowest third of the participants in Experiment 1 (see Figure 8.9), whose response times were an average 100 ms slower in the baseline condition than any of the means shown so far, these participants still incurred a reliable 149 ms – 9% cost [(*MSE* = 221,853.58), $F(1, 19) = 6.006$, $p < .02$ – (*MSE* = 5.625), $F(1, 19) = 3.985$, $p < .05$]. Similarly, when we look at the slowest third of the participants in Experiment 2 (see Figure 8.10), we see that although their response times ranged into the 1,500-ms range, they still incurred a reliable 134 ms cost [(*MSE* = –8.6%/222, 994), $F(1, 24) = 8.269$, $p < .006$]. These extreme group data support the idea that the dissociations between costs and benefits that we observed for comprehension skill and for the distance between the primes and targets are most likely not due to a ceiling effect.

These dissociations also suggest that the costs and benefits are not due to episodic retrieval. According to an episodic retrieval explanation, when participants read each prime sentence ending in a homonym, they neither suppress nor enhance any meaning. When they later read a target sentence that is simi-

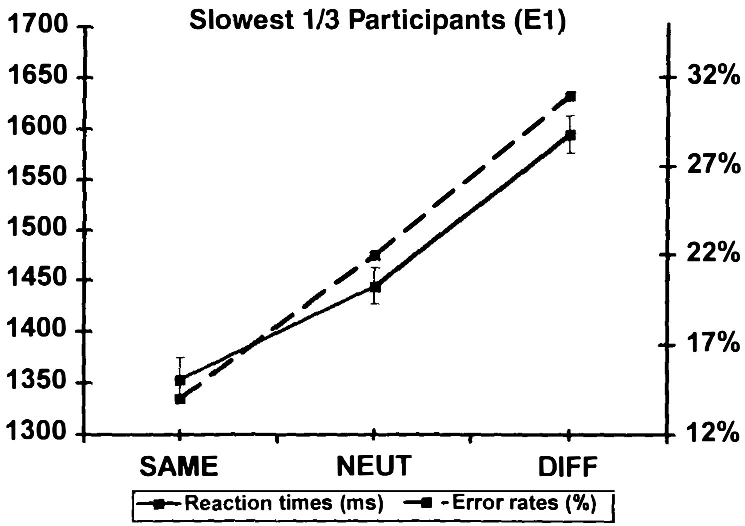

Figure 8.9. Average response times (in ms) and average error rates to target sentences of the slowest third of the participants (n = 20) from Experiment 1 (E1). The solid lines represent average response times (in ms), with the scale on the left, and the dashed lines represent average error rates, with the scale on the right. SAME = same-meaning prime; NEUT = neutral-meaning prime; DIFF = different-meaning prime.

lar to the prime sentence, they are reminded of the prime sentence. To the degree that the target sentence matched participants' memory of the prime sentence, they were facilitated or impeded in responding to the target sentence. However, an episodic retrieval explanation would have to account for why our participants' memory of different prime sentences fades more quickly than their memory of same prime sentences and our less-skilled comprehenders had poorer immediate memory of only the different prime sentences. Given that participants did not know when they read each sentence whether a later sentence would imply the same meaning or a different meaning, we conclude the following from these data.

The benefits that we have observed in these experiments are due to enhancement, and the costs are due to suppression. When participants read a sentence that contains a homonym, they suppress the meaning that is not implied by the sentence. When they later read a sentence that implies the previously suppressed meaning, they incur a cost. We also conclude from these data that the costs incurred by suppressing an irrelevant meaning are relatively short-lived, whereas the benefits reaped by enhancing a relevant meaning last longer. Furthermore, the less-skilled comprehenders, presumably because they are characterized by less-efficient suppression, incur little cost but reap the

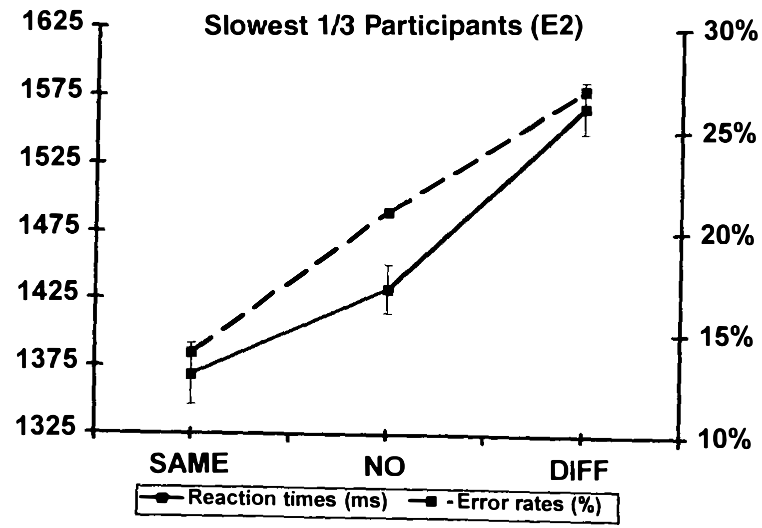

Figure 8.10. Average response times (in ms) and average error rates to target sentences of the slowest third of the participants (n = 25) from Experiment 2 (E2). The solid lines represent average response times (in ms), with the scale on the left, and the dashed lines represent average error rates, with the scale on the right. SAME = same-meaning prime; NO = no-meaning prime; DIFF = different-meaning prime.

same benefits as the more-skilled comprehenders. Finally, the dissociations that we have observed suggest that these costs and benefits should not be attributed to episodic retrieval.