Journal of Experimental Psychology:
Learning, Memory, and Cognition
2005, Vol. 31, No. 6, 1360–1371

0278-7393/05/$12.00 DOI: 10.1037/0278-7393.31.6.1360

The Representation of Homophones: Evidence From the Distractor-Frequency Effect

Michele Miozzo
Columbia University

Alfonso Caramazza
Harvard University

Current models of word production offer different accounts of the representation of homophones in the lexicon. The investigation of how the homophone status of a word affects lexical access can be used to test theories of lexical processing. In this study, homophones appeared as word distractors superimposed on pictures that participants named orally. The authors varied distractor frequency, a variable that has been shown to modulate the interference that distractors produce on picture naming. The results of 3 experiments converged in showing that words interfered in proportion to their individual frequency in the language, even if they have high-frequency homophone mates. This effect of specific-word frequency is compatible with models that assume (a) distinct lexical representations for the individual homophones and (b) that access to such representations is modulated by frequency. The authors discuss the extent to which current models of word production satisfy these constraints.

Keywords: word production, homophones, frequency effect, lexical processing, speaking

Many studies of speaking have shown that frequently occurring words are produced faster and more accurately than are less frequently occurring words (see Garrett, 2001, for a review). The received view is that part of this frequency effect involves lexical access, which occurs more smoothly the higher the frequency of the word (e.g., Jescheniak & Levelt, 1994; Wingfield, 1968). Consistent with this view, word-frequency effects are believed to offer a window on lexical processing. However, providing a precise characterization of word frequency has proven to be a challenging task for language scientists. Inflected words such as *tables* or *walked* are an illustrative example: Does their production reflect the frequency of their individual forms or, rather, the frequency of their stems *table* and *walk*? Homophones, words such as *hair/hare*, which have identical sounds but different meanings, are another case in point. One hypothesis is that the production of each homophone correlates with the frequency of its concept in the language, so that *hair* would be produced faster and more accurately than would *hare*, a word that is generally used far less frequently. Alternatively, homophone production could depend on the frequency with which word phonology is produced: In the case of words such as *hair* and *hare*, which have identical phonology, they would have identical frequencies and, everything else being equal, identical production latencies. Here we attempt to clarify the role of homophone frequency in speech production, an issue important for understanding the representation of homophones in the lexicon but also with broader implications for theories of lexical organization.

In the effort to elucidate the role of homophone frequency in word production, researchers have examined two types of frequency counts, namely specific-word frequency (i.e., the occurrences of *hare*) and cumulative homophone frequency (i.e., the occurrences of *hare* + *hair*) (Bonin, & Fayol, 2002; Caramazza, Costa, Miozzo, & Bi, 2001; Dell, 1990; Jescheniak & Levelt, 1994; Jescheniak, Meyer, & Levelt, 2003; Miozzo, Jacobs, & Singer, 2004). The objective has been to determine which of these frequency measures is the better predictor of homophone production. Researchers have used different techniques to elicit verbal responses, including picture naming and word translation, with different groups of participants (i.e., neurologically intact speakers and brain-damaged individuals with speaking impairments). The rationale of these experiments is taken in a straightforward manner from the perspective of current models of word production, which commonly assume a “localized” lexical representation for each of the words in a speaker’s lexicon (Caramazza, 1997; Dell, 1986; Levelt, Roelofs, & Meyer, 1999). This differs from connectionist models in which word knowledge is represented in a distributed fashion. Evidence of an effect of specific-word frequency would be consistent with the hypothesis that distinct lexical representations are accessed when producing homophones such as *hare* and *hair* and that the locus of the frequency effect is at the point where these representations are accessed. We refer to this hypothesis of homophone representation as the *independent representation* (IR) hypothesis (Caramazza et al., 2001). By contrast, the finding that cumulative homophone frequency is a crucial variable in word production would point to the conclusion that a common lexical representation is accessed in the production of *hare* and *hair* and that the frequency effect occurs at the point at which this common representation is accessed (Jescheniak & Levelt, 1994). Because

---

Michele Miozzo, Department of Psychology, Columbia University; Alfonso Caramazza, Department of Psychology, Harvard University.

This research was supported by a Keck Foundation Grant and by National Institutes of Health Grants DC006242-2 (to Michele Miozzo) and DC04542 (to Alfonso Caramazza). We thank Kristen Geiger and Melissa Jacobs for their help with the preparation and administration of the experiments.

Correspondence concerning this article should be addressed to Michele Miozzo, who is now at Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3ED, United Kingdom. E-mail: mm854@cam.ac.uk

homophones have the same phonology, whereas their syntactic properties can differ—as with homophones from different grammatical classes (e.g., *red* [noun]/*red* [adjective]) – the representation shared by homophones would encode phonology alone. We refer to this proposal as the *shared representation* (SR) hypothesis. The IR and SR hypotheses are compatible with different models of lexical processing, as we will see in greater detail below. Here we simply wish to emphasize that, to the extent that the frequency effect helps discriminate between the IR and SR hypotheses, it would also further constrain proposals about the functional architecture of the lexicon. We should also note that although the IR and the SR hypotheses have been the favored accounts within the perspective of current theories of word production, alternative accounts can be offered, as we will see later in the article when we discuss a connectionist proposal. Because of the prominence that IR and SR hypotheses have acquired in the debate on the homophone versus specific-word frequency effect, the predictions and results of our experiments will be considered from the perspective of those accounts alone, deferring the discussion of alternative proposals to the end of the article.

Unfortunately, the empirical findings have been mixed. Homophone production has been shown to vary as a function of specific-word frequency in picture naming tasks in different languages, as well as in anomia, a language impairment caused by brain damage (Bonin, & Fayol, 2002; Caramazza et al., 2001; Miozzo et al., 2004). However, an error elicitation paradigm revealed an effect of cumulative homophone frequency (Dell, 1990). And in word translation tasks, both types of frequency effects were found in experiments conducted by different research teams (Caramazza et al., 2001; Jescheniak & Levelt, 1994; Jescheniak et al., 2003).

It could be argued that the conflict in the results stems from the use of a variety of paradigms. But the task of linking the individual results to the characteristics of specific experimental paradigms is unlikely to succeed. These paradigms vary on so many dimensions—type of stimuli, response modality, participants' language, to mention a few—that it is difficult to pinpoint the factors in each paradigm that led to a particular finding (for a detailed discussion, see Caramazza, Bi, Costa, & Miozzo, 2004; Caramazza et al., 2001; Jescheniak et al., 2003). An alternative and probably more fruitful approach is to seek convergence of evidence, that is, to determine the type of frequency effect that is observed more frequently in a range of tasks involving homophone production. This is the approach we adopted in the present study. Specifically, we investigated the roles of specific-word and homophone frequencies in the picture–word interference paradigm, a task that psycholinguists have used to elucidate a number of aspects of word production (see Levelt et al., 1999, for a recent review). Below we introduce this task and describe how it is suited for the investigation of homophone production.

The Distractor-Frequency Effect

In the picture–word interference paradigm, pictures are shown along with words and participants are instructed to name the pictures and ignore the words. In a variant of this task, colored words are presented and participants name the word colors. Two recent studies (Burt, 2002; Miozzo & Caramazza, 2003) have reported slower responses for low-frequency (LF) distractors than for high-frequency (HF) distractors[1] in picture- and color-naming tasks. Robust effects of distractor frequency were obtained with HF and LF distractors that were semantically and phonologically unrelated to the picture name and that were matched for a number of variables known to affect the recognition of written words, including neighborhood size, bigram frequency, word length, grapheme-to-phoneme consistency, and concreteness (Miozzo & Caramazza, 2003).

The various accounts of the word interference effect hinge on the idea that it takes longer to select the picture name because distractors strongly activate their corresponding lexical nodes in the output lexicon (see, e.g., Caramazza & Costa, 2000; Glaser, 1992; Roelofs, 2003; Schriefers, Meyer, & Levelt, 1990). One may extend this claim to the distractor-frequency effect and thus consider it an "output" effect arising at the level of picture name selection and for this reason suited to our purpose of studying homophone frequency in production. But defining the locus of the word frequency effect is not straightforward. Since word frequency affects the recognition of written words, an "input" account that views the distractor-frequency effect as arising at the level of word recognition stands as a plausible alternative. An "input" account can be articulated in various ways. Within the input account we considered in an earlier article (Miozzo & Caramazza, 2003) the picture–word interference paradigm was viewed as a sort of dual-task, wherein picture naming is performed simultaneously with word reading. Performing two tasks simultaneously can be particularly demanding if, given the limited amount of cognitive resources available, fewer resources can be allocated to the individual tasks. To the extent that word recognition subtracts resources from the principal task of picture naming, this task becomes more difficult in the picture–word interference paradigm. Word interference could then reflect the amount of resources required to recognize the word distractors, and frequency could be one of the factors modulating the amount needed for word recognition. One might then reasonably assume that the recognition of LF words requires more resources than does the recognition of HF words, a difference that could explain the greater interference observed with LF distractors.

If a resource account were correct, we should find that distractor interference is reduced in conditions that facilitate distractor recognition and is increased when distractor recognition becomes more difficult. These predictions have not received strong empirical support, however. In two experiments (Burt, 2002; Miozzo & Caramazza, 2003), distractor recognition was facilitated by repeatedly presenting the distractors in the context of a lexical decision task that preceded the picture–word interference paradigm. Repetition facilitated distractor recognition, as evidenced by a reduction of lexical decision latencies. However, this manipulation led to an attenuation of interference only in Burt's (2002) experiment, whereas in Miozzo and Caramazza's (2003) experiment, no effects

---

[1] Two earlier studies (Fox, Shor, & Steinman, 1971; Klein, 1964) reported larger interference for HF words. As discussed in Burt (2002) and Miozzo and Caramazza (2003), these findings are methodologically limited—for example, they were based on a small number of items and used very rare words. In contrast, the finding of larger interference for LF words has reliably been replicated in two systematic investigations of the distractor-frequency effect (Burt, 2002; Miozzo & Caramazza, 2003) but also in other studies (La Heij, Mark, Sander, & Willeboordse, 1998; Monsell, Taylor, & Murphy, 2001).

of distractor repetition were observed. The reverse prediction—increased interference when distractor recognition is more demanding—was tested by Miozzo and Caramazza (2003) by showing distractors with alternated cases (e.g., *cHaIr*, *HaNd*), a manipulation that has been shown to impair word recognition (see, e.g., Besner & McCann, 1987). Contrary to the prediction of the input account, however, distractor interference did not increase when distractors were shown with alternated cases.

Instead, Miozzo and Caramazza (2003) found that procedures that affected phonological processing also affected the distractor-frequency effect. For example, the distractor-frequency effect was significantly reduced with sound-related picture–word pairs like *ball–wall* or *ball–shawl*. The interaction between the effects of distractor frequency and phonological relatedness suggests that the two effects involve common mechanisms. Consistent with the widely held view that sound-related distractors affect the retrieval of picture name phonology (see, e.g., Meyer & Schriefers, 1991; Starreveld & La Heij, 1996), the common mechanisms are likely to be those implicated in the production of word phonology. Evidence of this kind favors the interpretation of the distractor-frequency effect as an output effect, in particular one that involves the phonological retrieval of spoken words.

If our interpretation of the distractor-frequency effect is correct, contrasting predictions follow from the IR and the SR hypotheses about the interference that homophones would produce when shown as distractors in picture naming. The IR hypothesis, which assumes that distinct lexical representations exist for homophones, anticipates that homophones would interfere in proportion to their individual frequencies. On the other hand, an effect of cumulative homophone frequency is expected by the SR hypothesis, which proposes shared lexical representations for homophones. These contrasting predictions were tested in the experiments reported here.

Experiment 1: LF Homophones

In Experiment 1, we tested words like *ewe* whose specific-word frequency is low but whose homophone *you* is used far more commonly. Consequently, the cumulative homophone frequency of *ewe*, which is given by the sum of the frequencies of *ewe* + *you*, is high. We refer to words like *ewe* as LF homophones and words like *you* as HF homophones. We compared the interference produced by LF homophones with two groups of control words: LF controls (words with the same specific-word frequency as the LF homophones) and HF controls (words whose specific-word frequency was comparable to the cumulative homophone frequency of the LF homophones). To illustrate, we compared the LF homophone *ewe*, the LF control *bow*, and the HF control *she*.

If specific-word frequency is the critical variable, responses to *ewe* should be similar to those of the LF control *bow* but slower than those to the HF control *she* because low-frequency distractors have been shown to lead to relatively slow responses in the picture–word interference task. This result pattern would be consistent with the IR hypothesis. If, however, cumulative homophone frequency is the crucial variable here, *ewe* should interfere as much as the HF control *she* but less than the LF control *bow*. This result pattern would be in line with the SR hypothesis.

Method

Participants. The participants for all the studies reported in this article were students at Columbia University who received partial course credit in an introductory psychology course or payment for their participation. No participant took part in more than one of the experiments. All participants considered English to be their native language. Experiment 1 had 12 participants.

Material. We selected three sets of 27 distractors based on their specific-word frequency. For specific-word frequencies, we used the lemma frequency counts reported in Francis and Kučera (1982). Lemma frequency corresponds to the sum of the occurrences of the inflected forms of a word (e.g., *chair* + *chairs* or *walk* + *walks* + *walked* + *walking* + *to walk*). The only exceptions were words with irregular inflections (e.g., *write/wrote/written*), for which we used the sum of the counts of their regular inflected forms (e.g., *write* + *writes* + *writing* + *to write*; unpublished data from our laboratories indicate that this measure is more precise than lemma frequencies for words with irregular inflections). Cumulative homophone frequency corresponds to the sum of the specific-word frequencies of all the words with the same phonology (e.g., [a] *brake*+[to] *brake*+[a] *break*+[to] *break*). Log transformed frequency counts were used in the analyses reported in the present article. The set of LF homophones comprised 27 words that had a homophone of considerably higher specific-word frequency, $t(26) = 11.9$, $p < .0001$; see Table 1 for frequency means. We selected only HF–LF homophone pairs that had different spellings—that is, heterographs (e.g., *ewe/you*, *knot/not*, *reel/real*)—so as to remove any ambiguity in word recognition. To increase the chance of observing a frequency effect, we chose LF–HF homophone pairs with the largest frequency discrepancy. The pairs that met this criterion were composed, for the most part, of words of different grammatical classes—for example, noun–pronoun ([a] *ewe/you*) or noun–verb ([a] *rite*/[to] *write*). Each LF homophone was matched to a LF of the same grammatical class and with the same specific-word frequency range (<25 counts per million). In English, it is almost impossible to select a list of words without homophones. The LF controls had homophones of relatively low (specific-word) frequency, as demonstrated by the fact that the LF controls had very similar specific-word and cumulative homophone frequencies (*M*s = 10 and 12, respectively). LF homophones and LF controls were also matched for grammatical category. A third set of words (HF controls; $N = 27$) had specific-word frequencies comparable to the cumulative homophone frequencies of the LF homophones ($t < 1$). These words were of the same grammatical categories as the HF homophones. For example, the HF control of *ewe* was *she*, a pronoun like *you*.

In a number of cases, the grammatical class of English written words is ambiguous— consider for example *flour*, which can be either a noun or a verb. We used frequency to decide a word's grammatical class: we assigned a word the grammatical form that had the highest frequency count. Thus, *flour* was treated as a noun, because it is more commonly used as a noun than as a verb.

The three sets of distractors were controlled for variables that modulate word recognition, including bigram frequency, number of neighbors, length (number of letters), and grapheme-to-phoneme correspondence (see Table 1 for summary). One-way analyses of variance (ANOVAs) showed that the three distractor sets were matched on these variables (*p*s > .1). The only exception was bigram frequency, $F(2, 78) = 4.17$, $MSE = 3486874.9$, $p < .05$, which, as revealed by post hoc analyses (Newman–Keuls), was significantly greater for LF controls than for HF controls ($p < .05$; *M*s = 1,360 and 2,076, respectively).

The LF homophones, LF controls, and HF controls were superimposed on a set of 27 pictures. All picture–word pairs were neither semantically nor phonologically related (see list in Appendix A). We also showed 27 filler pictures, each of which was presented with three unrelated word distractors. Fillers were in all respects identical to the other items and were used as warm-up stimuli for the first three trials of each experimental block. Fillers were not included in any of the analyses.

Table 1
Written Word Distractors Shown in Experiments 1–3

	LF distractor		HF distractor	
	Homophone	Control	Homophone	Control
Variable	*ewe*	*bow*	*you*	*she*
	Experiment 1 and 2			
Number	27	27	27	27
Specific-word frequency[a]	7	10	1,723	1,202
Cumulative-homophone frequency[a]	1,796	12	1,796	1,266
Bigram frequency[b]	1,663	1,360	2,076	1,958
Neighborhood size[c]	7.50	7.90	6.00	7.80
Length (no. of letters)	4.20	4.20	4.00	4.00
Grapheme–phoneme consistency[d]	0.90	0.92	0.91	0.87
	Experiment 3			
	hare	*dune*	*hair*	*hall*
Number	24	24	24	24
Specific-word frequency[a]	6	5	192	200
Cumulative-homophone frequency[a]	205	6	205	204
Bigram frequency[b]	3,774	3,077	3,700	3,389
Neighborhood size[c]	7.00	5.40	7.10	8.40
Length (no. of letters)	4.80	4.80	4.60	4.60
Grapheme–phoneme consistency[d]	0.93	0.94	0.87	0.87
Concreteness[e]	4.90	4.70	4.70	4.70

Note. LF = low frequency; HF = high frequency.
[a] Frequency counts were from N.W. Francis and H. Kučera (1982). [b] We report the mean of the words' bigram frequencies; counts were from the English Lexicon Database (Washington University). [c] Coltheart's *N* count, a crude measure of a word's neighborhood density was obtained from the English Lexicon Database (Washington University). [d] This is a parameter with a 0–1 range (1 = maximum predictability) that is based on the probability with which each grapheme or graphemic cluster of a word maps onto a particular phoneme (R.S. Berndt, J.A. Reggia, & C.C. Mitchum, 1987). [e] This is a parameter with a 0–7 range (0 = *highly concrete,* 7 = *highly abstract*).

On each presentation, pictures appeared with a different distractor, and each distractor was shown only once in the course of the experiment. Distractors were typed in capital letters (Geneva, 20-point, bold font) and were positioned in the area around fixation that allowed optimal distractor recognition. For a given picture, words always appeared in the same location. Picture–word pairs of the various experimental conditions were evenly distributed across four blocks. In a block, a given picture appeared once. Trials were randomized with the constraints that items of the various experimental conditions were evenly distributed across the block and were not to appear in consecutive trials. Two randomizations were used.

Procedure. Participants were tested individually in a dimly lit testing room while seated at a distance of about 80 cm from the computer screen. The experiment started with a naming task to familiarize participants with the pictures and their names. Pictures were shown on a computer screen for unlimited exposure and with a string of 5 *X*s superimposed on the picture. The *X*s appeared in the same position as the words and served as a control for visual interference in picture identification. If participants produced a name other than the one selected by the experimenter, they were corrected. Next, participants were instructed to name the pictures as fast as they could without making mistakes. There were two training blocks before the experiment proper. In each training block, pictures appeared once, with word distractors not shown later in the experiment proper.

Trials were structured as follows: Participants initiated the trial by pressing the space bar; the fixation point (a cross) appeared for 700 ms and was immediately replaced by a picture–word pair. Stimuli remained on the screen until participants initiated their response, up to 600 ms. Stimulus presentation was controlled by the PsychLab program (Gum & Bub, 1988). Response latencies were measured from the onset of the stimulus by means of a voice key (Lafayette Instruments). Accuracy was monitored by the experimenter.

Analyses. Responses scored as errors included responses in which (a) participants produced an unexpected name, (b) verbal disfluencies were noticed (stuttering, utterance repairs, production of nonverbal sounds that triggered the voice key), and (c) recording failures occurred. Errors were discarded from analyses. Responses that exceeded a participant's mean by three standard deviations were treated as outliers and were also excluded from analyses. Planned comparisons were conducted to determine whether the three sets of distractors (LF homophones, LF controls, and HF controls) had similar effects on response latencies and accuracy. All analyses made use of one-way repeated measure ANOVAs. For response latencies and errors we report F_1 analyses (based on participants' means) and F_2 analyses (based on items' means). Error rates are typically low in the picture–word interference paradigm, and we report the results of error analyses only if significant ($p < .05$).

Results and Discussion

The results of Experiment 1 were clear-cut and overall in line with the predictions of the IR hypothesis. HF controls led to responses that were not only faster than the responses to the LF controls (775 vs. 807 ms), $F_1(1, 11) = 10.6$, $MSE = 577.2$, $p < .01$; $F_2(1, 26) = 8.7$, $MSE = 1700.4$, $p < .01$, but also faster than the responses to the LF homophones (775 vs. 806 ms), $F_1(1, 11) = 6.0$, $MSE = 941.7$, $p < .05$; $F_2(1, 26) = 16.1$, $MSE = 754.8$, $p < .001$. Moreover, response latencies were almost identical for LF homophones and LF controls (*M*s: 806 vs. 807 ms; *F*s < 1). In essence, the data of Experiment 1 indicate that homophones interfere like other words matched for specific-word frequency. This pattern of results is in line with the IR hypothesis but is seemingly incompatible with the SR hypothesis. We should further note that the size of the distractor-frequency effect in Experiment 1 (about 30 ms) was comparable to that found by Miozzo and Caramazza (2003) in several experiments that investigated the effect of distractor frequency with the same task.

The LF controls had lower bigram frequencies than did the HF controls. It is unlikely that the LF controls interfered more than the HF controls because of their lower bigram frequencies. If such were the case, we should not have found a difference between the HF homophones and the HF controls, as these distractors were closely matched for bigram frequency. The explanation that the results of Experiment 1 reflect variations of specific-word frequency provides a more comprehensive account and is the one that we consider in the following discussion.

A drawback of the design of Experiment 1 is that it did not include HF homophones, and therefore it did not provide direct evidence about the interference produced by these words. The reason for adopting this design was our concern that the inclusion of HF homophones (e.g., *you*) would affect the way in which the LF homophones (e.g., *ewe*) would be processed. The interpretation of the findings of Experiment 1, however, rests on the assumption that the HF homophones (e.g., *you*) would interfere as much as the HF controls (e.g., *she*). But this assumption might be false. What if the HF homophones selected for Experiment 1 interfered (for whatever reason) more than did their HF controls? If this were the case, the results of Experiment 1 would not be incompatible with the SR hypothesis. These doubts concerning the HF homophones are assessed in Experiment 2 presented below, in which we examined how HF homophones interfered when used as distractors for picture naming.

In Experiment 1 we selected the homophone pairs with the largest frequency discrepancy so as to maximize the chance of obtaining a distractor-frequency effect. The homophone pairs with these characteristics were for the most part of different grammatical classes, such as the noun–pronoun pair *ewe–you* or the noun–verb pair *rite–write*. In order to match the homophones to control words as closely as possible, we chose LF homophones and LF controls that were of the same grammatical class; the same held between HF homophones and their HF controls. Thus, the LF noun *ewe* was paired with the LF noun *bow*, whereas the HF pronoun *you* was paired with the HF pronoun *she*. Consequently, the majority of LF homophones and HF controls differed not only with respect to frequency but also with respect to grammatical class. In particular, whereas LF distractors were mostly nouns (22/27, 81%), fewer HF distractors were nouns (12/27, 44%). We cannot rule out that the LF homophones interfered more than the HF distractors because nouns were represented relatively more frequently among the LF homophones. This account gains some plausibility if we consider that the task used in Experiment 1 involved noun production. It is not unreasonable to think that nouns would produce greater competition than would words of other grammatical classes in a task involving noun production. Contrary to these suspicions, Pechmann and Zerbst (2002) did not observe longer naming latencies when the distractors were nouns than when they were functors. Nevertheless, to eliminate any doubts about the possible confounding of grammatical class, in Experiment 3 we showed LF–HF homophone pairs of the same grammatical class (nouns).

Experiment 2: HF Homophones

Experiment 2 was carried out to determine whether the HF homophones would interfere to the same extent as did the HF controls selected for Experiment 1.

Method

The method for Experiment 2 is as described in the *Method* section of Experiment 1, except 16 participants took part in Experiment 2. The materials and procedure were identical to those of Experiment 1, with one exception: LF homophones were replaced by their HF homophones (the characteristics of these words are described in the *Method* section of Experiment 1). For example, the picture *bottle* appeared in Experiment 1 with the LF homophone distractor *ewe* and in Experiment 2 with the HF homophone *you*. As in Experiment 1, distractors were controlled for bigram frequency, number of neighbors, length, and grapheme-to-phoneme correspondence. The three distractor sets were well matched for all of these variables ($p > .15$) except for bigram frequency, $F(2, 78) = 3.9$, $MSE = 3{,}981.3$, $p < .05$. As shown by Newman–Keuls post hoc analyses, this difference emerged because LF controls had significantly ($p < .05$) higher bigram frequencies than did HF homophones and HF controls (see means in Table 1). This difference is potentially problematic if we find that the LF distractors interfere more than the HF distractors: This result could arise because of the difference in bigram frequency, specific-word frequency, or both. A similar confound was present in Experiment 1, but as we saw bigram frequency could not account for the results of Experiment 1. In light of these results, it seems unlikely that bigram frequency could explain the greater interference of LF distractors in Experiment 2. (The results of Experiment 3 also rule out an effect of bigram frequency.)

Results

Means for errors and response latencies are shown in Table 2. Naming latencies were overall slower in Experiment 2 than in Experiment 1, as demonstrated by the responses to LF controls and HF controls, which in Experiment 2 were 36 and 31 ms longer, respectively. An overall tendency for the participants to respond

Table 2
Mean Response Latencies and Percentage Error Rates (ER) Obtained in Experiments 1–3

	LF distractor		HF distractor	
	Homophone	Control	Homophone	Control
Experiment	*hare*	*dune*	*hair*	*hall*
1				
M	806	807		775
ER	0.90	0.30		1.50
2				
M		843	800	806
ER		3.20	3.00	1.80
3				
M	778	773	757	752
ER	1.90	2.20	1.20	1.00

Note. LF = low frequency; HF = high frequency.

more slowly in Experiment 2 can also explain why responses were almost identical between LF (806 ms; Experiment 1) and HF homophones (800 ms; Experiment 2). Analyses showed that response latencies were statistically identical between HF homophones and HF controls (800 vs. 806 ms; *F*s < 1). Both groups of HF distractors interfered less than did LF controls (HF homophones vs. LF controls: $F_1(1, 15) = 33.6$, $MSE = 441.8$, $p < .0001$, $F_2(1, 26) = 11.5$, $MSE = 2324.2$, $p = .002$; HF controls vs. LF controls: $F_1(1, 15) = 37.2$, $MSE = 282.9$, $p < .0001$, $F2(1, 26) = 8.3$, $MSE = 2102.2$, $p = .01$. In sum, the results of Experiment 2 revealed that the HF homophones interfered like other nonhomophonic words matched for frequency (and grammatical class). Experiment 2 has implications for the interpretation of the critical finding of Experiment 1, which showed that LF homophones produced larger interference that HF controls. On the basis of Experiment 2, we can rule out that this crucial finding reflected anomalies with the HF homophones.

Experiment 3: Homophonic Nouns

In Experiment 3, we compared the interference produced by homophonic nouns such as *hare* and *hair*, which differ significantly in terms of specific-word frequency. This procedure departs from those adopted in previous reaction time studies, in which the behavior of the HF homophones was inferred on the basis of the responses to HF controls. Experiment 3 allowed us to determine the interference produced by HF homophones and so to directly compare how pairs of HF–LF homophones affected picture naming. Predictions are straightforward from the perspective of the IR and the SR hypotheses: Whereas the former predicts larger interference for LF homophone nouns, the latter predicts comparable levels of interference between LF and HF homophone nouns.

In Experiment 3, we also tested control (nonhomophonic) noun distractors that were matched to HF and LF homophones for specific-word frequency. To illustrate, the distractors shown in Experiment 3 included *hare* (LF homophone), *dune* (LF control), *hair* (HF homophone), and *hall* (HF control). Controls were introduced to assure that the distractor-frequency effect could be obtained within the frequency range of the homophones chosen for Experiment 3. Note that because LF and HF homophones were matched to LF and HF controls, Experiment 3 offered another opportunity to test the predictions of the IR and the SR hypotheses that we tested in Experiment 1. Namely, whereas the IR hypothesis predicts that the LF homophones would interfere as much as the LF controls and more than the HF controls, the SR hypothesis predicts that the LF homophones would interfere less than would the LF controls and as much as would the HF controls. Thus, according to the IR hypothesis, HF versus LF homophones and HF versus LF controls should lead to distractor-frequency effects that are quantitatively similar and independent of a word's homophone status; however, the SR hypothesis predicts that HF versus LF homophones and HF versus LF controls should lead to quantitatively different effects depending on the homophone status of the distractor words.

Experiment 3 also served as a control for Experiment 1. In Experiment 1, we observed that LF homophones interfered more than HF controls, a finding that we cannot univocally attribute to frequency because the LF distractors and HF controls were, for the most part, of different grammatical classes. The results of Experiment 3, in which we used only nouns, clarify the role of grammatical class in Experiment 1. We should note that the inclusion of noun distractors makes it easier to control for concreteness, a variable known to affect distractor interference (Lupker, 1979). Distractors in Experiment 3 were also matched for concreteness.

Participants were shown only one word in each of the HF–LF homophone pairs selected for Experiment 3 (either *hair* or *hare*). In this way, we removed any possible influence that the presentation of a word could have on the response to its homophone.

Method

Participants. Thirty-four participants took part in Experiment 3. An equal number of participants was randomly assigned to Group A and Group B (which are explained below).

Materials. A set of 96 nouns was used as distractors. The set consisted of 24 pairs of heterographic homophones and 48 matched control words. One member of each homophone pair had a higher specific-word frequency, $t(23) = 10.6$, $p < .0001$ (see Table 1 for means). Each homophone was paired with a control word (a noun) that was closely matched for specific-word frequency ($t < 1$) and had the same length (number of letters). Only 30/96 (31%) of the distractors had homographic homophones listed in Francis and Kučera (1982; e.g., the pair [a] *flower*/[to] *flower*). Nouns that had homographic homophones listed had a frequency that was at least twice the frequency of their homographic homophones (nouns: $M = 167$; homographic homophones: $M = 7.7$), $t(29) = 12.1$, $p < .0001$. These data give us confidence that the distractors selected for Experiment 3 were most probably recognized as nouns. We selected control nouns that either had no homophones or had homophones in the low-frequency range, as indicated by the similarities of their specific-word and cumulative homophone frequencies (see means in Table 1). Homophones and control nouns were also controlled for bigram frequency, neighborhood size, length (number of letters), grapheme-to-phoneme regularity, and concreteness (see means in Table 1). Two-way (homophone vs. controls; high- vs. low-frequency) ANOVAs were carried out to determine whether each of these features varied uniformly within each distractor set. We consistently found a lack of interactions (*p*s $> .20$), a finding that suggests that the distractor sets were well matched for the features controlled in Experiment 3.

HF–LF homophones and controls were divided into two groups (Group A and Group B). Each group included either the LF homophone or its HF homophone. For example, *hare* was in Group A, whereas *hair* was in Group B. The frequency-matched controls of the individual homophones were included in the same group. Thus, *dune*, the LF control for *hare*, was in Group A, whereas *hall*, the HF control for *hair*, was in Group B. Distractors of different frequencies and type (homophones vs. controls) were equally represented within Groups A and B.

The homophones and their controls were mounted on 24 pictures, and care was taken to avoid semantically or phonologically related pairs (see list in Appendix B). We also selected 28 pictures, each of which was paired with 4 semantically and phonologically unrelated distractors. These pictures served as fillers. Participants saw only the distractors of either Group A or Group B. Participants presented with Group A had the distractors of Group B substituted by filler words, and Group B had the distractors of Group A substituted by fillers. In this way, each picture was presented four times in the course of the experiment. Participants named a total of 208 pictures in Experiment 3.

Procedure and analyses. The procedure and analyses were identical to those described in the *Method* section of Experiment 1.

Results

As is evident from an inspection of Table 2, responses were faster with pictures paired with HF homophones than those paired with LF homophones. A parallel discrepancy appeared between HF and LF controls. The variables distractor (homophones vs. controls) and distractor frequency (high vs. low) were analyzed by means of two-way ANOVAs. The effect of distractor frequency was significant, $F_1(1, 33) = 9.1$, $MSE = 1425.6$, $p < .01$, $F_2(1, 23) = 17.5$, $MSE = 599.6$, $p < .001$, a finding that reflects longer naming latencies associated with LF than with HF distractors, the hallmark of the distractor-frequency effect. ANOVAs revealed neither a main effect of the variable distractor (Fs <1) nor signs of a Distractor × Distractor Frequency interaction (Fs with ps $> .25$). The latter results indicate that homophones behaved like nonhomophonic words of the same frequency and interfered as a function of their specific-word frequency.

The question of whether the LF words included in an experiment are represented in the participants' vocabularies is always a legitimate concern in studies that use LF words. The most probable consequence of including unfamiliar distractors in our Experiment 3 would be a reduction of the frequency-effect size. Consistent with previous findings, which have shown that (orthographically plausible) nonwords interfered less than did familiar words (e.g., Klein, 1964), unfamiliar low-frequency words are expected to generate even less interference than high-frequency words, thereby rendering the possibility of finding a distractor-frequency effect less likely. The fact that a sizable effect of distractor frequency occurred indicates that, by and large, low-frequency distractors were familiar.

Nonetheless, there is a legitimate concern that two of the low-frequency homophone distractors, *hart* and *tulle*, may have been unfamiliar to the participants. To assess the contribution of their familiarity more directly, we reran the ANOVAs described above, this time excluding those two homophone distractors as well as their paired distractors. Other concerns of a different nature can be raised for the distractors *aunt* and *leader*. *Aunt* is not consistently homophonic with *ant* across American English dialects (Wells, 1990). The morphologically complex word *leader* is not ideal for testing models like the one proposed by Levelt et al. (1999), which maintain a lemma–lexeme distinction and hold that lexemes encode morphemes (the notions of lemma and lexeme will be discussed in detail later in the article). In such a model, a word like *leader* is assumed to be stored in a decomposed form (*lead* + *er*). Because *leader* and *liter* have distinct lexemes within such models, these words could produce different levels of interference. *Aunt* and *leader*, along with their paired distractors, were also excluded from the post hoc analysis, which was then based on 20 of the 24 distractor quadruplets shown in Experiment 3. Even with this smaller set, means were comparable between LF distractors (homophones = 785 ms, controls = 780 ms) and slower than the means of HF distractors (homophones = 761 ms, controls = 762 ms). The effect of distractor frequency remained significant, $F_1(1, 33) = 9.55$, $MSE = 1466.9$, $p < .01$, $F_2(1, 19) = 13.0$, $MSE = 692.8$, $p = .001$, and there was no evidence of an effect of distractor type or a Distractor Frequency × Distractor Type interaction (Fs < 1).

Further confirmation that distractors interfered as a function of their specific-word frequency and not their cumulative frequency comes from the results of a regression analysis in which the two types of frequencies were entered as predictors of the naming latencies recorded in Experiment 3. Although specific-word frequency was a significant predictor, $t\,(45) = -2.51$, $p = .01$, $r^2 = .09$, cumulative frequency failed to affect the responses in the picture–word interference task ($t < 1$; note that this pattern of results held even when the potentially problematic items we described above were excluded from the analysis (specific-word frequency with $t\,(37) = 2.3$, $p = .002$; cumulative frequency with $t < 1$, $r^2 = .11$).

Overall, the results of Experiment 3 replicated and further extended the findings of Experiment 1. It is important to note that because Experiment 3 escapes the criticisms raised for Experiment 1, the data of Experiment 3 allow us to reach firmer conclusions about how homophones interfere, conclusions that can be compared with the predictions derived from the IR and SR hypotheses. Although the homophone findings of Experiment 3 are completely in line with the predictions of the IR hypothesis, they are inconsistent with the predictions of the SR hypothesis. In summary, the data of Experiment 3 lend further support to the IR hypothesis.

General Discussion

The experiments reported in the present article were aimed at determining how homophones interfere when used as distractors in picture naming. The results are clear-cut: A homophonic word interferes as a function of its frequency in the language—namely, its specific-word frequency. Our results have direct implications for understanding what variables affect homophone processing in speaking. Previous data provided an ambiguous picture of whether homophone processing varies as a function of specific-word or cumulative homophone frequency. Even more important than showing once again the effects of specific-word frequency is that we obtained this evidence with a new paradigm.

Our results are relevant for the issue of homophone processing in speaking only if it can be shown that the effect of distractor frequency involves mechanisms of word production, that is, if it is an output effect rather than an input effect originating at the level of word recognition. The claim that this is an output effect has received empirical support from previous studies such as that of Miozzo and Caramazza (2003), which extensively tested the likelihood that the distractor-frequency effect reflected asymmetries in the recognition of HF and LF words. They found no support for that possibility. Instead, their results aligned well with the hypothesis that the distractor-frequency effect involves mechanisms for word production. It is important to note that results seemingly inconsistent with the hypothesis of a frequency effect originating at the level of word recognition also emerge from the present study. The first piece of evidence pointing in that direction comes from a regression analysis carried out on the responses of Experiment 3. We assessed the contribution of specific-word frequency and bigram frequency, a variable that reflects the orthographic structure of the distractors and which has been shown to affect word recognition (Broadbent & Gregory, 1968; Rice & Robinson, 1975; Westbury & Buchanan, 2002). Specific-word frequency contributed significantly to picture naming times, $t(45) = -3.02$, $p = .003$, $r^2 = .09$, whereas the contribution of bigram frequency was negligible in this respect ($t < .01$). (The different response latencies between the participants in Experiments 1 and 2 represented a confounding element that prevented us from extending the regression analysis to those experiments).

Our finding that LF homophone distractors (*hare*) and LF control distractors (*smog*) produced comparable levels of interference provides a second, more indirect piece of evidence. Homophones have also been tested in tasks that presumably demand the recognition of the words' orthography and semantics but do not require access to the phonological form for production. These tasks include lexical decision (e.g., Grainger, Nguyen Van Kang, & Segui, 2001; Pexman, Lupker, & Jared, 2001; Rubinstein, Lewis, & Rubinstein, 1971), semantic categorization (Jared & Seidenberg, 1991; Van Orden, 1987), and eye movement measures in sentence reading (e.g., Folk, 1999). The prevalent result with these tasks is longer responses to LF homophones relative to frequency-matched controls. In essence, different result patterns emerged in those tasks and in the picture–word interference paradigm reported here. In light of this cross-task difference, it seems unlikely that mechanisms implicated in visual word recognition were (primarily) responsible for our findings with LF homophones, for otherwise we would have expected parallel effects in the picture–word interference paradigm and in tasks tapping visual word recognition. Taken together, not only do our results further support the output account of the distractor-frequency effect, they also indicate that an account of this type can be extended to heterographic homophones, the distractors used in our experiments. In the following paragraphs, we explore the implications of our results with homophone distractors for theories of word production.

Although current models of word production differ along a number of dimensions, the vast majority of these models share two assumptions that are of particular relevance for homophone processing. First, most models hold a localized view of lexical representations, thereby proposing distinct lexical nodes or separate lexical entries for each of the words in a speaker's lexicon. Second, the majority of the models assume distinct lexical representations for homophonic words. At least in principle, these models can incorporate the IR hypothesis, which postulates distinct lexical nodes for homophone pairs like *hare* and *hair* and which is supported by the data presented here. A point of divergence, as we discuss below, concerns whether this distinction is absolute, so that homophonic words do not share any lexical representations, or whether this distinction is only partial, so that homophonic words share at least some part of their lexical representations.

The IR hypothesis can be incorporated in lexical models that assume that a single lexical layer mediates the semantic and phonological contents of words stored in the lexicon (Caramazza, 1997; Harley, 1999). This type of model is schematically represented in Figure 1A, and we refer to it as the *one-lexical-layer model*. Here a syntactically specified lexical node serves as a binder of its phonological features. In the course of lexical access, activation spreads from the semantic system to lexical nodes that are possible candidates for production and then to the phonological features linked to these nodes. Within a model of this type, homophones can take only distinct lexical nodes. And if frequency were to affect access to lexical nodes, homophone production should vary as a function of specific-word frequency, a prediction consonant with the finding we obtained here with the picture–word interference paradigm.

A different hypothesis about the architecture of the lexical system assumes that there are two lexical layers between the semantic and phonological contents of words, such that there are two types of lexical nodes associated with each word (Dell, 1986; Garrett, 1992; Levelt et al., 1999). One lexical node, the lemma, encodes the syntactic characteristics of the word; the other lexical node, the lexeme, specifies the morphological form of the word (see Figure 1B). We refer to this architecture as the *two-lexical-layer model*. Lexical access unfolds in this model in three steps: Activation spreads first from the semantic system to lemmas, then from lemmas to lexemes, and then from lexemes to their phonological contents. Proponents of two-lexical-layer models have typically assumed that whereas homophones have distinct lemmas, they have common lexemes (Cutting & Ferreira, 1999; Dell, 1990; Jescheniak & Levelt, 1994). This configuration captures the fact that though homophone pairs can have different syntactic features (e.g., *ewe* is a noun, *you* a pronoun; *him* is masculine, *hymn* has no gender), their phonology is identical. In other words, within this type of model the IR hypothesis is embodied at the lemma level. If one assumed that frequency affected lemma access, homophone production would be expected to depend crucially on specific-word frequency.

Dell (1990) proposed a two-lexical-layer model that locates frequency at the lemma level. Like other models, HF words are assumed to reach higher activation levels than are LF words, a difference that explains the advantage speakers have in producing HF words. Interactivity is a defining feature of Dell's model, where activation propagates from lemmas downward to lexemes and then to phonological features and bounces back to lemmas from the layers beneath. As a result of interactivity, the lexemes of homophones receive activation from multiple lemmas in Dell's model. For example, activation converges to the lexeme /heir/ from the lemma of *hair* and *hare*. This leads to the possibility that the LF homophone *hare* would receive a sizable amount of activation from the lemma of its HF mate *hair*, with the result that *hare* will be substantially more activated than a LF nonhomophonic word such as *smog*. Differences in the activation levels between *hare* and *smog* were found in computer simulations of Dell's model. Thus, even if frequency were encoded at the lemma level, Dell's model would still anticipate that LF homophones behave differently than LF control words, contrary to the results of our experiments.

The model proposed by Jescheniak and Levelt (1994) is a two-lexical-layer model in which activation propagates in a feed-forward fashion. Their model locates frequency at the lexeme levels, which leads to the incorrect prediction that cumulative frequency should be the principal determinant of speed of lexical access of homophones. This model also predicts that LF distractor homophones should behave like HF distractor homophones in the picture–word interference task used in the experiments reported here. An alternative variant of two-lexical-layer and feed-forward models that would be consistent with the findings of specific-word frequency effects is one that would extend the IR assumption to the lexeme level, and then assume that lexeme access is also sensitive to frequency (see Figure 1C). In this way, distinct lemmas and lexemes would correspond to the homophones *hare* and *hair*, and access to these nodes would reflect the frequency of the individual homophones. This type of solution does not face the problem of confining the frequency effects to a single level of lexical nodes. Here the challenge is to explain how distinct representations would emerge for words that sound alike at a level of processing where only phonology is encoded. In addition, there is recent evidence from the production of pronominal clitics (pronouns that attach to verbs as in the Italian phrase *portalo* [bring it$_{\text{masculine}}$]) that the locus of the frequency effect is at the level where a word's

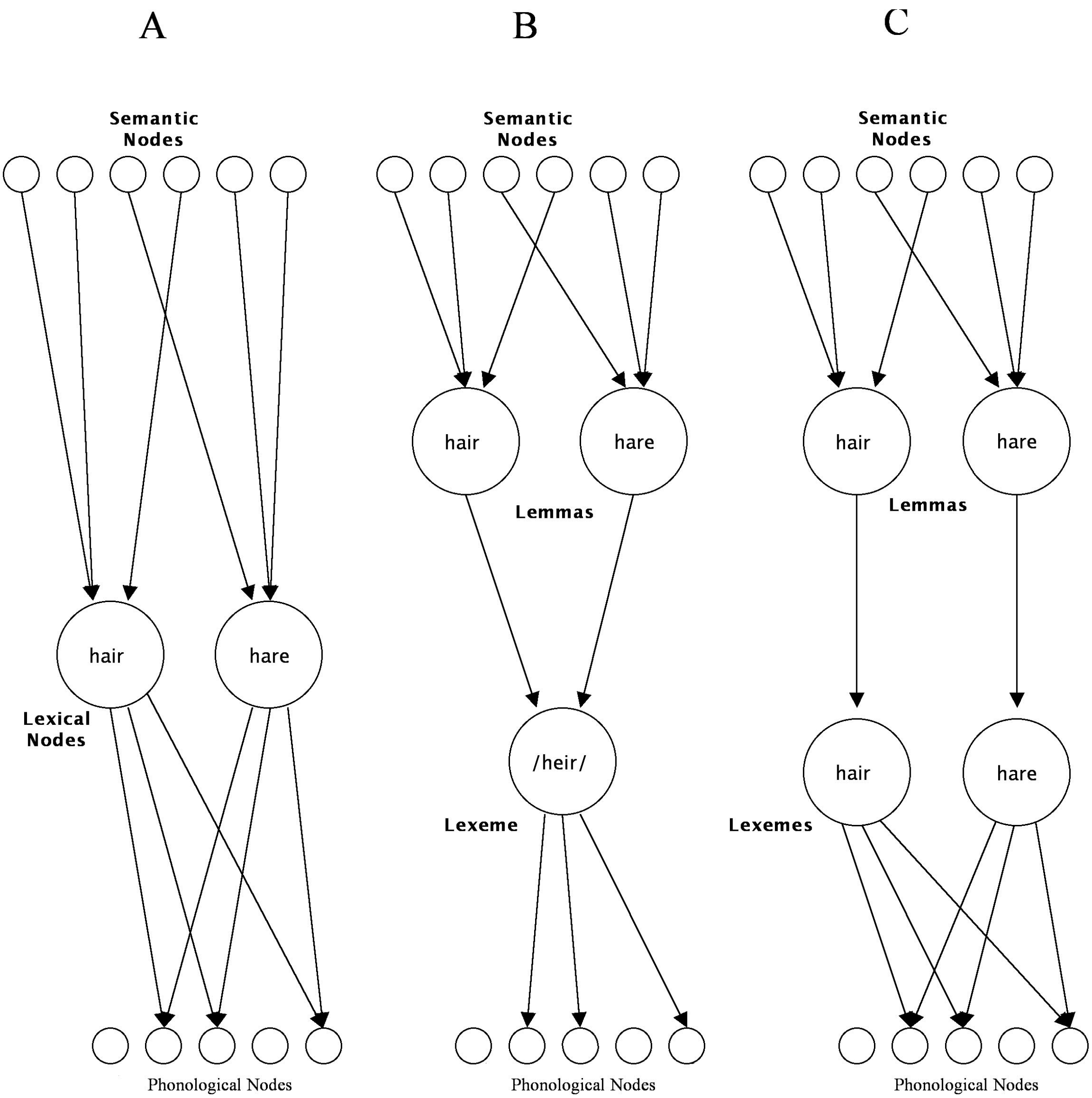

Figure 1. Schematic illustration of three hypotheses about homophone representations in the lexicon accessed for word production.

syntactic properties are specified, thereby ruling out an account that locates the frequency effect at the level of lexeme representations (Finocchiaro & Caramazza, in press). These researchers found a frequency effect in pronoun production even though the noun's phonological form is not accessed in the course of producing the pronoun.

Thus far, our discussion has focused on the predictions derived from localist lexical models. We would now like to consider a connectionist approach that assumes distributed representations and in which frequency effects emerge naturally from the strength of the whole system of connections. Joanisse and Seidenberg's (1999) model includes word production (among other lexical tasks) and could provide a starting point for examining homophone production from a connectionist perspective. Their model assumes distinct layers of representations for the semantic and phonological features of words and bidirectional links between these layers. Words are represented in a distributed fashion in both layers. As in other connectionist models, syntactic features such as grammatical class or verb tense are not explicitly encoded. These properties emerge as a by-product of the interaction between semantics and phonology, instead. In a model like that of Joanisse and Seidenberg (1999), different semantic units would correspond to the homophones *hair* and *hare*, but their phonological units would overlap completely. In connectionist models, where activation circulates in both directions between layers, words tend to be coactivated in proportion to the features they share. As a consequence of this general property of connectionist models, one would expect the semantic units of the word *hair* to be activated during the production of its homophone *hare* and to contribute to the activation of the phonology of *hare*. Similar levels of coactivation would not appear with control words, whose phonology is not completely shared. A prediction in models with massive interaction would seem to be that homophones should behave differently from other nonhomophonic words. Our finding of virtually identical response with homophones and control words would thus seem to be at variance with these types of connectionist models, which assume global and massive forms of interaction. Inhibitory mechanisms or "clean-up" units could be devices that might be introduced to prevent a reverberation of activation between homophonic words. Whether these solutions would enable connectionist models to obtain the specific-word effects reported here and in other experiments is presently unknown.

In conclusion, our data from the picture–word interference paradigm have confirmed the specific-word frequency effect that has emerged in the majority of the paradigms used so far to investigate homophone frequency (Bonin, & Fayol, 2002; Caramazza et al., 2001; Miozzo et al., 2004). Of course, further evidence is required to more precisely establish the type of frequency that is critical for homophone production, so as to clarify the conflicting results in the literature. But as a clearer picture of the homophone-frequency effect emerges, language researchers can count on this phenomenon to test theories of lexical processing in word production.

References

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention & performance XII: The psychology of reading* (pp. 201–219). London: Erlbaum.

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology, 14,* 289–313.

Broadbent, D. E., & Gregory, M. (1968). Visual perception of words differing in letter digram frequency. *Journal of Verbal Learning and Verbal Behavior, 7,* 569–571.

Gum, T., and Bub, D. (1988). PsychLab [Computer software]. Montreal, Quebec, Canada: Montreal Neurological Institute.

Burt, J. S. (2002). Why do noncolor words interfere with color naming. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1019–1038.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14,* 177–208.

Caramazza, A., Bi, Y., Costa, A., & Miozzo, M. (2004). What determines the speed of lexical access: Homophone or specific-word frequency? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 278–282.

Caramazza, A., & Costa, A. (2000). The semantic interference effect in the picture–word interference paradigm: Does the response set matter? *Cognition, 75,* 51–64.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The representation of homophones: Evidence from the frequency effect in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1430–1450.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 318–344.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5,* 313–349.

Finocchiaro, C., & Caramazza, A. (in press). The locus of the frequency effect in word production: Evidence from the production of clitics in Italian. *Language and Cognitive Processes.*

Folk, J. R. (1999). Phonological codes are used to access the lexicon during silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 892–906.

Fox, L. A., Shor, R. E., & Steinman, R. J. (1971). Semantic gradients and interference in naming color, spatial direction, and numerosity. *Journal of Experimental Psychology, 91,* 59–65.

Francis, N. W., & Kučera, H. (1982). *Frequency analysis of English usage.* Boston: Houghton Mifflin.

Garrett, M. (2001). Now you see it, now you don't: Frequency effects in language production. In E. Dupoux (Ed.), *Language, brain, and cognitive development* (pp. 227–240). Cambridge, MA: MIT Press.

Garrett, M. F. (1992). Disorders of lexical selection. *Cognition, 42,* 143–180.

Glaser, W. R. (1992). Picture naming. *Cognition, 42,* 61–106.

Grainger, J., Nguyen Van Kang, M., & Segui, J. (2001). Cross-modal repetition priming of heterographic homophones. *Memory & Cognition, 29,* 53–61.

Gum, T., & Bub, D. (1988). PsychLab [Computer software]. Montreal, Quebec, Canada: Montreal Neurological Institute.

Harley, T. A. (1999). Will one stage and no feedback suffice in lexicalization? *Behavioral and Brain Sciences, 22,* 45.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120,* 358–394.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824–843.

Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific-word frequency is not all that counts in speech production. Some comments on Caramazza et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 432–438.

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences USA, 96,* 7592–7597.

Klein, G. S. (1964). Semantic power measured through the interference of words with color naming. *American Journal of Psychology, 77,* 576–588.

La Heij, W., Mark, P., Sander, J., & Willeboordse, E. (1998). The gender-congruency effect in picture–word tasks. *Psychological Research, 61,* 209–219.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–75.

Lupker, S. J. (1979). The semantic nature of responses competition in the picture–word interference task. *Memory & Cognition, 7,* 485–495.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture–word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 1146–1160.

Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture–word interference paradigm. *Journal of Experimental Psychology: General, 132,* 228–258.

Miozzo, M., Jacobs, M. L., & Singer, N. J. W. (2004). The representation of homophones: Evidence from anomia. *Cognitive Neuropsychology, 21,* 840–866..

Monsell, S., Taylor, T. J., & Murphy, K. (2001). Naming the color of a word: Is it responses or task sets that compete? *Memory & Cognition, 29,* 137–151.

Pechmann, T., & Zerbst, D. (2002). The activation of word class information during speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 233–243.

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 139–156.

Rice, G. A., & Robinson, D. O. (1975). The role of bigram frequency in the perception of words and nonwords. *Memory & Cognition, 3,* 513–518.

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

Rubinstein, H., Lewis, S. S., & Rubinstein, M. A. (1971). Evidence for phonetic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10,* 645–657.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language, 29,* 86–102.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning Memory, and Cognition, 22,* 896–918.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15,* 181–198.

Wells, J. C. (1990). *Longman pronunciation dictionary* (2nd ed.). Harlow, UK: Longman.

Westbury, C., & Buchanan, L. (2002). The probability of the least likely non-length-controlled bigram affects lexical decision reaction times. *Brain & Language, 81,* 66–78.

Wingfield, A. (1968). Effects of frequency on identification and naming objects. *American Journal of Psychology, 81,* 226–234.

Appendix A

Picture–Word Pairs Shown in Experiments 1 and 2

	Word Distractor			
Picture	LF homophone	HF homophone	LF control	HF control
Apple	brake	break	chord	offer
Arrow	nun	none	mat	once
Bed	pore	poor	gill	easy
Belt	sew	so	tap	out
Bottle	ewe	you	bow	she
Bread	teem	team	pave	king
Candle	bawl	ball	gnaw	trip
Comb	sundae	Sunday	maggot	Monday
Cross	chute	shoot	chive	marry
Deer	knot	not	robe	at
Door	flour	flower	cheer	wheel
Drum	knight	night	lyric	woman
Flag	hymn	him	veil	her
Hammer	seam	seem	rope	tell
Hanger	sine	sign	moss	rule
Lamp	hare	hair	smog	tree
Lemon	grate	great	avail	small
Lion	thyme	time	knoll	year
Mouse	dye	die	bun	act
Mushroom	rite	write	cane	begin
Nose	hoarse	horse	ardent	paper
Pipe	inn	in	rib	on
Plate	heir	air	wool	tax
Rope	pane	pain	claw	mass
Tent	oar	or	hog	but
Tomato	reel	real	chip	free
Truck	ladder	latter	ribbon	former

Note. Low-frequency (LF) and high-frequency (HF) homophones were shown in Experiment 1 and Experiment 2, respectively.

Appendix B

Picture–Word Pairs Shown in Experiment 3

	Noun distractor			
Picture	LF homophone	HF homophone	LF control	HF control
Foot	heir	air	feud	aid
Tie	flour	flower	wheat	bottle
Spoon	mussel	muscle	nectar	breath
Book	steppe	step	stucco	club
Horse	beech	beach	vault	coast
Cow	ark	arc	elm	egg
Frog	liter	leader	bonus	father
Fork	hare	hair	dune	hall
Pig	gait	gate	bout	lake
Ring	prophet	profit	thunder	league
Anchor	pane	pain	bean	loss
House	ant	aunt	ink	meal
Broom	cymbal	symbol	musket	moment
Bottle	sundae	Sunday	morgue	Monday
Drum	tee	tea	hue	pie
Camel	pear	pair	fern	poet
Finger	berth	birth	mulch	price
Bed	tulle	tool	sable	roof
Truck	manor	manner	baron	summer
Bell	hart	heart	lard	voice
Sock	mettle	metal	puddle	wagon
Tomato	knight	night	trench	water
Envelope	rein	rain	loaf	wine
Lemon	thyme	time	chive	year

Note. LF = low frequency; HF = high frequency.

Received August 11, 2004
Revision received June 10, 2005
Accepted July 1, 2005 ■