Journal of Experimental Psychology:
Learning, Memory, and Cognition
2001, Vol. 27, No. 6, 1430–1450

0278-7393/01/$5.00 DOI: 10.1037//0278-7393.27.6.1430

The Specific-Word Frequency Effect: Implications for the Representation of Homophones in Speech Production

Alfonso Caramazza and Albert Costa
Harvard University

Michele Miozzo
Columbia University

Yanchao Bi
Harvard University

In a series of experiments, the authors investigated whether naming latencies for homophones (e.g., /nΛn/) are a function of specific-word frequency (i.e., the frequency of *nun*) or a function of cumulative-homophone frequency (i.e., the sum of the frequencies of *nun* and *none*). Specific-word but not cumulative-homophone frequency affected picture-naming latencies. This result was obtained in 2 languages (English and Chinese). An analogous finding was obtained in a translation task, where bilingual speakers produced the English names of visually presented Spanish words. Control experiments ruled out that these results are an artifact of orthographic or articulatory factors, or of visual recognition. The results argue against the hypothesis that homophones share a common word-form representation, and support instead a model in which homophones have fully independent representations.

Homophones are words that have the same pronunciation but differ in meaning, spelling, or grammatical class. How are homophones represented and accessed in speech production? Two hypotheses have been proposed. One view holds that homophones share a common lexical–phonological representation, but because they have different meanings and often also different grammatical properties (e.g., *sun/son*; *the watch/to watch*; *him/hymn*), they have different semantic and lexical–grammatical representations (Cutting & Ferreira, 1999; Dell, 1990; Jescheniak & Levelt, 1994; Levelt, Roelofs, & Meyer, 1999).[1] We refer to models of this type as *shared representation* (SR) models. There are four levels of representation in these models: semantic/conceptual nodes, lemma nodes, lexeme nodes, and phonological nodes. Lemmas specify the word's grammatical properties, whereas lexemes specify their phonological contents. Figure 1A schematically represents this hypothesis. The alternative hypothesis attributes no special status to homophones. Each word, homophonic and nonhomophonic, is represented independently (Caramazza, 1997; Harley, 1999). We refer to models of this type as *independent representation* (IR) models. One such proposal is schematically represented in Figure 1B. Here there are only three levels of representation in lexical access: semantic/conceptual nodes, lexical nodes, and phonological nodes.

The results of two recent studies have been interpreted as providing support for the SR hypothesis (Dell, 1990; Jescheniak & Levelt, 1994). In both studies, the authors investigated the effects of homophone frequency on naming performance. Given the assumption that homophones share a common representation, the effective frequency of the homophone word form would be the sum of the frequencies of the homophonic words. For example, the frequency of the word form /nΛn/ would be the sum of the frequencies of the homophonic words *nun* and *none*. We refer to this property of homophones as *cumulative-homophone frequency*; the term *specific-word frequency* will be used to refer to the frequency of individual words (*nun* or *none*). Jescheniak and Levelt (1994) reasoned that if word frequency were to affect the retrieval of word forms in production, a clear prediction would follow from the SR hypothesis: The retrieval of homophonic words should be determined by cumulative-homophone frequency

Alfonso Caramazza, Albert Costa, and Yanchao Bi, Department of Psychology, Harvard University; Michele Miozzo, Department of Psychology, Columbia University.

The work reported here was supported in part by National Institutes of Health Grants NS22201 and DC0452 awarded to Alfonso Caramazza. Albert Costa was supported by a Fulbright Fellowship from the Spanish Government; Michele Miozzo was supported by a start-up grant from Columbia University and a Keck Foundation grant. We thank Dr. Hua Shu for her assistance in the experiments conducted in Beijing, China. We thank Dr. Nuria Sebastian for her assistance in the experiments conducted in Barcelona, Spain. We thank Delia Kong, Jie Zhuang, and Elena Tenconi for their help in participant testing.

Correspondence concerning this article should be addressed to Alfonso Caramazza, Cognitive Neuropsychology Laboratory, Department of Psychology, William James Hall, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138. Electronic mail may be sent to caram@wjh.harvard.edu.

[1] Note that whether homophonic words are homographic (*the watch/to watch*) or heterographic (*him/hymn*) has not been considered to be a relevant factor in theories of lexical access in speech production. Nevertheless, below we will consider the possible role of orthographic form in theories of phonological lexical access.

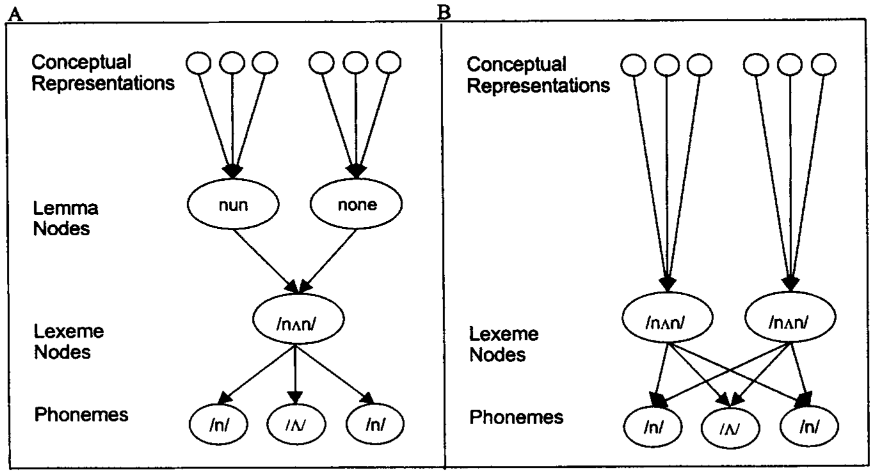

Figure 1. Schematic representation of the shared (Panel A) and independent (Panel B) representation hypotheses.

and not by specific-word frequency.[2] Thus, for example, although *nun* and *none* have very different frequencies, retrieval performance for the two words should be the same.[3]

In the context of a larger study addressing the locus of the frequency effect in lexical access, Jescheniak and Levelt (1994) tested the hypothesis that cumulative-homophone and not specific-word frequency affects word-production times. They tested bilingual Dutch–English speakers. Participants were given English words and were required to produce the Dutch equivalent as fast as possible. Three types of Dutch words were included in the study: (a) low-frequency homophonic words that have a high-frequency homophone mate, (b) nonhomophonic words matched to the homophones on specific-word frequency (specific-word frequency controls), and (c) nonhomophonic words matched to the homophones on homophone frequency (cumulative-homophone frequency controls).[4] An English example will illustrate the two types of control words used in the study. If the target were the word *nun*, one control word, *owl*, would have the same frequency as the word *nun*—the specific-word frequency control—and the other control word, *tooth*, would have the same frequency as the sum of frequencies of the homophones *nun* and *none*—the cumulative-homophone frequency control. On the assumption that the frequency effect is located at the stage of word form (lexeme) retrieval—the level at which homophones presumably share a representation—Jescheniak and Levelt's version of the SR hypothesis makes a straightforward prediction: Naming latencies for *nun* should be similar to naming latencies for the cumulative-homophone frequency control (*tooth*) rather than for the specific-word frequency control (*owl*). This outcome is expected, even though the frequency of the word *nun* is the same as that of *owl* (both low-frequency words) and it is lower than that of *tooth* (a high-frequency word). The results showed that mean naming latencies for the homophones (*nun*) were about 100 ms faster than that for the specific-word controls (*owl*), and roughly equal to that for the homophone frequency controls (*tooth*). Jescheniak and Levelt (1994) interpreted this result as evidence that the locus of the frequency effect in lexical access is at the lexeme level and that homophones share a common lexeme (as in Figure 1A).

Dell (1990) also reported results that are consistent with the SR hypothesis. Dell assessed the effects of frequency on the probability of making sound errors in producing function/content word homophones (e.g., *him/hymn*, *would/wood*).[5] Dell used an error-inducing paradigm, in which participants were required to produce simple phrases (e.g., "not/knot the mop") as quickly as possible. There were two main findings: (a) the overall error-incidence rate was the same for content and function words, and (b) frequency affected the rate of sound errors, with higher frequency words resulting in fewer errors. More important for present purposes, a post-hoc analysis revealed that the log frequency of the function words (e.g., *not*) predicted the rate of sound errors for their

---

[2] There is a controversy in the literature concerning whether word frequency or age of acquisition (AoA) better accounts for variation in picture-naming latencies (Barry, Morrison, & Ellis, 1997; Ellis & Morrison, 1998; Hodgson & Ellis, 1998; Morrison, Chappell, & Ellis, 1997; Morrison & Ellis, 1995; Morrison, Ellis, & Quinlan, 1992). Although our discussion of the putative homophone frequency effect is couched in terms of word frequency, we do not intend to prejudge the issue of which of the two factors best explains lexical access performance. Because the relevant research on homophone representation has been based on the dimension *word frequency*, we will follow this practice. Nonetheless, below we also consider the factor of AoA on naming latencies.

[3] This prediction only follows on the assumption that specific-word frequency does not contribute to performance. However, the SR hypothesis does not require such a strong assumption. It could be that both homophone frequency and specific-word frequency contribute to performance. Thus, a more general prediction is that retrieval performance is a function of both specific-word and cumulative-homophone frequencies and therefore we would not expect performance on homophonic words of different specific-word frequencies to be exactly the same. However, we would still expect low-frequency homophones to benefit from the frequency of other high-frequency mates.

[4] It should be noted that because Dutch has a transparent orthography, all the homophones used in Jescheniak and Levelt's (1994) study were also homographs.

[5] In this study, all the homophones were heterographs (e.g., *not/knot*).

homophonic content words (e.g., *knot*) better than the frequency of the content words themselves.

Dell (1990) interpreted these results as support for the SR hypothesis. However, unlike Jescheniak and Levelt (1994), he ascribed the locus of the frequency effect to the lemma level. This is possible in Dell's model because of the interactivity of activation between lemma and lexeme levels. The shared word-form node of a homophone (e.g., /nΛn/) sends activation back to its lemma cohort (*nun, none*), which in turn sends activation down to its shared lexeme node (/nΛn/), and so on for a number of iterations. In this way, a nontarget lemma node affects the activation level of the lexeme node that it shares with the target lemma. Given the further assumption that frequency modulates the level of activation transmitted by a node, a high-frequency lemma will send relatively more activation to its lexeme than will a low-frequency lemma. As a consequence, the lexeme of a low-frequency homophone (*nun*) will reach a higher activation level if its lemma cohort (*nun, none*) has high-frequency members. The results of simulation experiments confirmed the feasibility of this conclusion, thereby showing that the existence of a homophone effect in lexical retrieval does not, on its own, uniquely determine the locus of the frequency effect within the lexical access system. That is, the frequency effect could be located at the retrieval of either the lexeme nodes or lemma nodes, depending on other processing assumptions implemented in the model.

The results obtained by Jescheniak and Levelt (1994) and Dell (1990) are not incompatible with the IR hypothesis (Caramazza & Miozzo, 1998). A cumulative-homophone effect is expected in an architecture such as that depicted in Figure 1B if we were to assume interactivity between the lexical node layer and the segmental layer. Dell's simulation experiments with this type of architecture also confirmed the feasibility of this expectation. The cumulative-homophone effect could arise in an IR interactive model because the phonological segments of the target word would receive activation from two different lexical nodes: the target lexical node (*nun*), and its high-frequency homophone(s) (*none*). That is, the activation of the target lexical node (*nun*) would activate its phonological segments (/n/, /Λ/, /n/), which in turn would activate all the lexical nodes with which they are connected (e.g., *nun* and *none*). These lexical nodes will, in turn, send some activation back down to their phonological segments. On the assumption that the levels of activation of the phonemes depend, among other things, on the frequency of the lexical nodes with which they are connected, phonemes should be retrieved more easily when the low-frequency word has a high-frequency homophone.[6]

Although the cumulative-homophone frequency effect does not, on its own, distinguish between the SR and IR hypotheses of homophone representation, it still plays an important role in distinguishing between specific assumptions and models of lexical access. One should consider the case of Jescheniak and Levelt's (1994) model. This model makes three major assumptions: (a) activation only spreads forward and discretely (i.e., noncascaded processing); (b) homophones share a common lexeme representation; and (c) frequency affects the activation/selection of lexemes but not lemmas. Given these assumptions, the model predicts a cumulative-homophone frequency effect. The failure to observe a cumulative-homophone frequency effect would undermine the model; that is, the results would indicate that at least one of the model's assumptions would have to be modified. For example, we could retain the SR assumption, and locate the frequency effect at the lemma level. With this modification, we would not expect a cumulative-homophone frequency effect, because the speed of lexical access would be determined by specific-word frequencies and not by cumulative-homophone frequencies. Alternatively, we could give up the SR assumption, and keep the frequency effect at the lexeme level, where homophonic words would be represented by distinct lexemes for each word (e.g., separate lexemes for *nun* and *none*). This model also predicts the absence of a cumulative-homophone frequency effect.

Because of its interactive nature, predictions about the locus of frequency effects are more complex in the case of Dell's (1990) model. However, as already noted, the interactivity assumption predicts a cumulative-homophone frequency effect for the two lexical access systems shown in Figure 1. Failure to obtain a homophone frequency effect would challenge the assumption of feedback activation in lexical access. It should be noted, however, that the effects of interactivity can only be revealed by simulation studies, where the consequences of selecting different parameter values for the model's variables (e.g., connectivity strength) can be varied systematically (Dell & O'Seaghdha, 1991). We would then be able to ascertain whether there are parameter settings that allow us to obtain the absence of a homophone frequency effect, as well as other relevant facts about lexical access, such as the lexical bias effect observed in the slips-of-the-tongue data (Dell & Reich, 1981).

The importance of the cumulative-homophone frequency effect in distinguishing among models of lexical access has recently been highlighted by Levelt et al. (1999; see also Roelofs, Meyer, & Levelt, 1998). These authors have argued that this effect is incompatible with lexical models that do not distinguish between lemma and lexeme levels of representation (Caramazza, 1997; Caramazza & Miozzo, 1997; Harley, 1999; Starreveld & La Heij, 1995, 1996). We have seen that this claim is too strong, because one-lexical-layer models (Figure 1B) that assume interactivity between the lexical and segmental layers can account for the cumulative-homophone effect (Caramazza & Miozzo, 1998; Dell, 1990). However, the demonstration of a reliable cumulative-homophone frequency effect would undermine the one-lexical-layer model proposed by Caramazza (1997), which does not assume interactivity between the lexical and the segmental layers. Given the potentially crucial role played by the cumulative-homophone frequency effect in distinguishing among assumptions and models of lexical access, it is surprising that neither the naming latencies

[6] Dell's (1990) simulations show that the cumulative-homophone effect can be explained by a model that (a) embodies the IR hypothesis and (b) does not postulate a lemma–lexeme distinction. Nevertheless, Dell argued against a model of this sort, on the grounds that it would predict what he called an effect of "outcome word-frequency." That is, the model would predict that sound errors would more likely result in a word response for high-frequency outcomes (e.g., *late rack* → *rate lack*) than for low-frequency outcomes (e.g., *lake rag* → *lake lag*). Because this effect has not been observed in the speech-error data, Dell opted for a model with a lemma–lexeme distinction, and located the frequency effect at the lemma level. Whatever the merits of this argument, the point here is simply that the existence of a homophone frequency effect is insufficient to determine the locus of the effect.

effects reported by Jescheniak and Levelt (1994) nor the error effects reported by Dell (1990) have been replicated. Therefore it is crucial to establish the reliability of the phenomenon. Here we report several experiments that investigate the effect of homophone frequency on naming time.

In the following experiments we investigate the extent to which there is a cumulative-homophone frequency effect in picture naming. We follow Jescheniak and Levelt (1994) in assuming that the frequency effect in picture naming arises (primarily) during lexical access. In their Experiment 2, Jescheniak and Levelt asked participants to perform a picture recognition task. They used the same pictures that had previously been found to show a frequency effect in a picture-naming task (their Experiment 1). The authors argued that if the frequency effect observed in the picture-naming task were due to differences in the time needed to recognize high- and low-frequency pictures, one would have expected to find a frequency effect in the picture recognition task. The results did not support this prediction. Instead, no frequency effect was obtained in the object recognition task, leading the authors to conclude that the frequency effect arises primarily at the level of lexical selection (see Wingfield, 1967, 1968, for converging evidence; however, also see Kroll & Potter, 1984, for results suggesting a frequency effect in object recognition).

In Experiments 1 and 2, we examined whether picture-naming latencies are affected by the specific-word or the cumulative-homophone frequency of a homophonic word. We used a picture-naming task, because it provides a simple and direct way to test the cumulative-homophone frequency effect in speech production. Given the assumption that the frequency effect in picture naming reflects processes at the level of lexical access, we can use the task to assess whether specific-word or cumulative-homophone frequency determines naming latencies. Speakers of different languages were tested: In Experiment 1, we tested English speakers; in Experiment 2, we tested Chinese speakers. In Experiment 3 we attempted a direct replication of Jescheniak and Levelt's (1994) bilingual translation study. In the latter experiment, we tested English–Spanish bilinguals.

Our focus on the effect of word frequency in naming homophones reflects the emphasis that has been placed on this property of words to examine the structure of lexical representations (e.g., Cutting & Ferreira, 1999; Dell, 1990; Jescheniak & Levelt, 1994). However, it must be noted that there is some controversy regarding the role of word frequency in lexical processing. A number of researchers have argued that the frequency effect in picture naming is actually an effect of the *age of acquisition* (AoA) of words (e.g., Morrison & Ellis, 1995; Morrison, Ellis, & Quinlan, 1992; but see Lewis, Gerhand, & Ellis, 2001). Research aimed at resolving this issue is inconclusive. First, word frequency and AoA are highly correlated. Second, it appears that word frequency and AoA may each independently account for part of the variance in picture-naming tasks (Barry et al., 1997; Ellis & Morrison, 1998). Consequently, although our focus will be primarily on the effects of word frequency in naming homophones, we will also report the effects of AoA on naming performance.

Experiment 1A: Picture Naming in English

Using the criteria of Jescheniak and Levelt (1994), we selected three sets of pictures: (a) HomName pictures (e.g., *nun*), whose names in English have higher frequency homophone mates (*none*), (b) controls matched on specific-word frequency (e.g., *owl*), and (c) controls matched on cumulative-homophone frequency (e.g., *tooth*). The issue addressed here is whether picture-naming latencies are determined by a picture's specific-word frequency or by its cumulative homophone frequency. Operationally, this translates into the following question: Are HomName pictures named as fast as specific-word frequency controls or are they named as fast as cumulative-homophone frequency controls? The SR hypothesis predicts that naming latencies for the HomName condition should be comparable with those for the cumulative-homophone frequency controls, and faster than the naming latencies for the specific-word frequency controls. The IR hypothesis predicts the reverse pattern of results: HomName picture latencies should be comparable with those for the specific-word frequency controls, and slower than the naming latencies for the cumulative-homophone frequency controls.

Method

Participants. Thirty native English speakers who were students at Harvard University participated in Experiment 1A. Participants in this and in the other experiments reported here were paid for their participation, unless indicated otherwise.

Materials. The set of HomName pictures consisted of 26 pictures, each having one or more homophone mates. Each HomName picture had at least one homophone mate of higher frequency. Each HomName picture was paired with a picture matched for specific-word frequency ($F < 1$), and with a picture matched for cumulative-homophone frequency ($F < 1$).[7] Mean frequencies (from Francis & Kučera, 1982) are shown in Table 1. (The means reported here are for the 25 pictures retained for analysis. One item, and its associated controls, was not analyzed because it was mistakenly included in the homophone set. See Appendix A for the complete list of stimuli.) The mean frequencies of the names of HomName and specific-word frequency pictures were lower than the mean frequency of the names of the cumulative-homophone frequency pictures (both $ps < .01$). The three picture sets were also matched for number of syllables and for familiarity ratings, which were obtained by having 10 native English speakers score the familiarity of the picture names on a scale from 1 (*very unfamiliar*) to 5 (*very familiar*) (see Table 1; $Fs < 1$). In all the ratings collected in this study, when the target word was a homonym (e.g., *watch*), a synonym of the intended meaning was printed next to the target word (e.g., *watch–clock*). Another set of pictures ($N = 52$) was used as fillers and served as warm-up stimuli at the beginning of each block. Participants were shown a total of 130 pictures (78 experimental pictures and 52 fillers). Responses to filler pictures were not included in the analyses.

Procedure. The experiment began with a training block in which participants were asked to name the entire set of 130 pictures. If they produced a name that differed from the one expected by the experimenters, they were immediately asked to use the designated name. In the experiment proper, each picture was shown three times. The experiment was divided into three blocks of 130 pictures, and in each block a given picture appeared once. Both in the training block and in the experiment proper,

[7] In a language such as English, the operational distinction between homophones and nonhomophones can only be a relative one, because any noun can be used as a verb, and vice versa. Therefore, the words used in the control conditions could also be homophones. However, the frequency of any homophonic word with a control word was of such low frequency as to be effectively insignificant. This point is clearly illustrated in Table 1, where specific-word frequency control pictures had almost identical specific-word and cumulative-homophone frequencies.

Table 1
Mean Frequency (Specific and Cumulative), Familiarity, and Length for the Pictures of Experiment 1A

Picture set	Example	Specific frequency	Cumulative frequency	Familiarity	Length
Homophone name pictures	nun	28.2 (0–195)	142.0 (15–906)	3.2	4.3
Specific-word frequency matched pictures	owl	28.0 (1–160)	28.2 (1–160)	3.4	4.6
Homophone frequency matched pictures	tooth	116.8 (16–393)	117.9 (16–393)	3.9	4.4

Note. Frequency ranges are shown in parentheses.

participants were asked to name the pictures as fast as possible without making errors. Order of presentation was randomized with the constraint that pictures of a given experimental condition would not appear in more than three consecutive trials. Block order was randomized for each participant. Each trial had the following structure: a fixation point (a cross) was shown in the center of the screen for 700 ms, and was then replaced by the picture. The picture remained on the screen for 600 ms. Participants initiated the next trial by pressing the space bar. Response latencies were measured from the onset of the stimulus to the beginning of the naming response by means of a voice key (Lafayette Instrument Company, Lafayette, IN). Stimulus presentation was controlled by the PsychLab program (Bub & Gum, University of Victoria, British Columbia, Canada). Response accuracy was manually recorded by the experimenter.

Analyses. Responses scored as errors included (a) names that were not the ones designated as target responses, (b) verbal dysfluencies (stuttering, utterance repairs, production of nonverbal sounds which triggered the voice key), and (c) voice-key failures. Erroneous responses and responses longer than 3 s or shorter than 300 ms were excluded from the analyses. Outliers, responses exceeding a participant's mean by three standard deviations, were also eliminated. These exclusionary criteria were also applied in the other picture-naming experiments reported here. In the analyses of response latencies, two variables were examined: picture set (HomName pictures vs. specific-word frequency pictures vs. cumulative-homophone frequency pictures) and presentation (first vs. second vs. third). These variables were treated as within-subject variables with one exception: In $F2$ analyses, picture set was considered a between-subject variable. The same analyses were repeated with participants' error rate as a dependent measure. The results of the latter analyses are reported only if significant. As already noted, one of the HomName items was eliminated from the analyses because of a selection error. The control pictures associated with this item were also excluded.

Results and Discussion

The data of one participant were excluded because of an exceedingly high error rate (14%). Errors and outliers accounted for 3.6% of the data. Mean errors and mean response latencies for each picture set are shown in Table 2. Analyses of variance (ANOVAs) on the naming latencies revealed a significant effect of picture set, $F1(2, 56) = 68$, $MSE = 878.4$, $p < .0001$; $F2(2, 72) = 4.1$, $MSE = 12{,}670$, $p < .02$, and presentation, $F1(2, 56) = 3.4$, $MSE = 4{,}341$, $p < .05$; $F2(2, 144) = 18$, $MSE = 675$, $p < .001$. We did not find signs of interaction between these variables ($ps > .29$), suggesting that the size of the effect of frequency remained constant across picture repetitions. Jescheniak and Levelt (1994) and Levelt, Praamstra, Meyer, Helenius, and Salmelin (1998) previously reported the lack of interaction between repetition and frequency in picture naming (but see Griffin & Bock, 1998).

To clarify the nature of the main effect of picture set, we directly compared the responses obtained with the three groups of stimuli. The 50-ms difference between HomName and cumulative-homophone frequency pictures was highly significant, $F1(1, 28) = 115$, $MSE = 949$, $p < .0001$; $F2(1, 48) = 8.5$, $MSE = 11{,}191$, $p < .005$. The 38-ms difference between specific-word frequency and cumulative-homophone frequency pictures was also significant, $F1(1, 28) = 59$, $MSE = 1{,}077$, $p < .0001$; $F2(1, 48) = 4.9$, $MSE = 11{,}614$, $p < .03$. HomName pictures were named slightly slower than specific-word frequency pictures (12 ms), a difference that reached significance only in the $F1$ analysis, $F1(1, 28) = 9.7$, $MSE = 609$, $p < .004$; $F2 < 1$.

We also assessed the contribution of specific-word and cumulative-homophone frequencies on naming latencies by means of regression analysis. The SR hypothesis of lexical access predicts that cumulative-homophone frequency is a better predictor of naming latencies; the IR hypothesis predicts the opposite outcome: viz., that specific-word frequency is a better predictor of naming latencies. Log-frequency counts were used in these analyses. Only

Table 2
Mean Naming Latencies (and Error Rates) for the Pictures of Experiment 1A

Picture set	Example	Presentation 1	Presentation 2	Presentation 3	Average
Homophone name pictures	nun	755 (4.8)	782 (3.0)	754 (3.3)	764 (3.7)
Specific-word frequency matched pictures	owl	748 (4.1)	768 (3.3)	741 (4.6)	752 (4.0)
Homophone frequency matched pictures	tooth	709 (2.9)	724 (2.8)	708 (3.2)	714 (3.2)

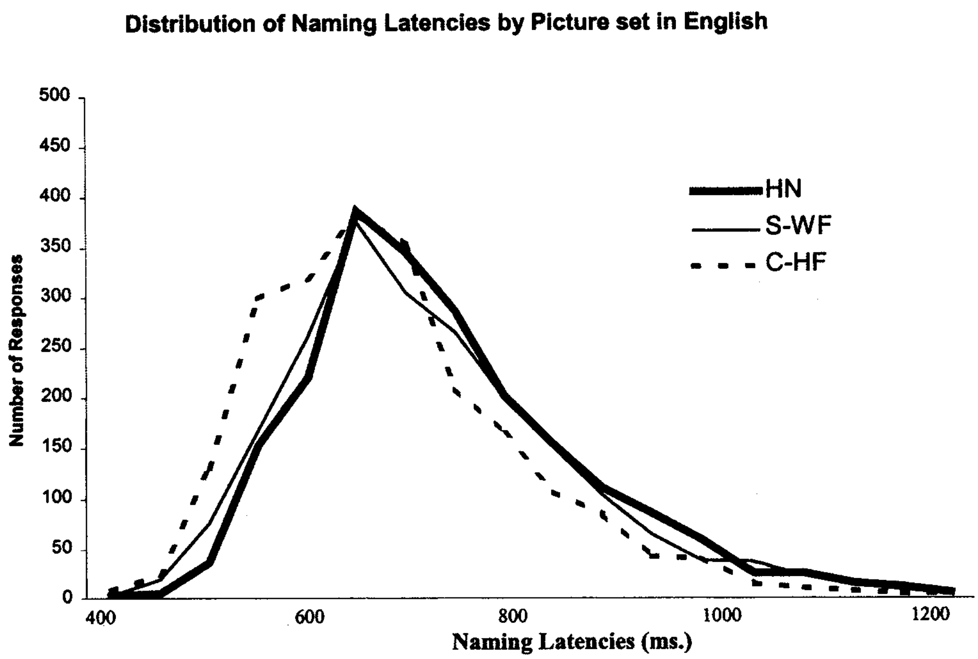

Figure 2. Distribution of naming latencies in Experiment 1A. The size of the interval is 50 ms. The lower interval begins at 400 ms, the highest interval ends at 1,300 ms. HN = homophone name; S-WF = specific-word frequency; C-HF = cumulative-homophone frequency.

specific-word frequency was a significant predictor (specific-word frequency: $R^2 = .11; p < .003$; cumulative-homophone frequency: $R^2 = .03; p < .12$). When the two variables were analyzed in a stepwise regression, we found that the inclusion of the cumulative-homophone frequency variable does not add any significant explanatory power to the regression model (Model 1, specific-word frequency: $R^2 = .116$; Model 2, specific-word frequency and cumulative-homophone frequency: $R^2 = .124$). These results show that naming latencies are only affected by specific-word frequency.

In the introduction, we noted that the controversy concerning whether word frequency or AoA better accounts for naming latencies remains unresolved. In order to assess if naming latencies are predicted by the age at which the different words are acquired, we obtained AoA ratings for the words used in our experiments. Fourteen native English speakers rated the items of Experiment 1 on a 7-point scale (1 = *acquired between 0–2 years old*, 2 = *acquired between 2–4 years old*, etc.). AoA ratings were also obtained for the pictures' higher frequency homophones (e.g., *none*, *dear*, *son*). AoA was found to be significantly correlated with word frequency ($r = .50, p < .001$). When AoA and word frequency were included in a regression analysis, specific-word AoA explained a larger share of the variance ($R^2 = .27; p < .001$) than specific-word frequency ($R^2 = .11; p < .003$). Furthermore, when the two variables were analyzed in a stepwise regression, the inclusion of specific-word frequency did not add appreciably to the variance explained by specific-word AoA alone (Model 1, AoA: $R^2 = .27$; Model 2, AoA and specific-word frequency: $R^2 = .27$). We also considered a measure of AoA that could be considered comparable with cumulative-homophone frequency: the AoA of the homophone word that was first learned (MinHomophone AoA). For example, the MinHomophone AoA of the pair *bear/bare* was equivalent to the AoA of *bear*, the word of the pair that was first learned. In a regression analysis, MinHomophone AoA accounted for a smaller share of the variance than specific-word AoA ($R^2 = .23$). Furthermore, when both variables were entered in a stepwise regression analysis, MinHomophone AoA does not add appreciably to the variance explained by specific-word AoA (Model 1, specific-word AoA: $R^2 = .27$; Model 2, specific-word AoA and MinHomophone AoA: $R^2 = .28$). We return to the implications of the AoA result in the General Discussion.

The findings of Experiment 1A are clear: Both HomName and specific-word frequency pictures were named slower than cumulative-homophone frequency pictures (see Figure 2). Our results did not show any benefit from having a higher frequency homophonic mate in picture naming. We failed to replicate the homophone frequency effect observed with other word production tasks by Dell (1990) and Jescheniak and Levelt (1994).

There are two possible confounding factors in our experiment that may explain the absence of a homophone frequency effect. One is that HomName pictures might have been especially difficult to recognize. If such were the case, any benefit that would have accrued to HomName pictures from having a higher frequency homophone mate would be masked by the longer time required for recognizing these stimuli. This possibility was tested in Experiment 1B, in which Italian speakers named the pictures shown in Experiment 1A. When translated into Italian, the picture sets do not vary systematically in terms of homophone status. If the results obtained with HomName pictures in English stemmed from a relative difficulty in recognizing those pictures, we would expect Italian speakers to also name them significantly more slowly than their specific-word frequency controls. In particular, we would expect the difference in naming latencies between the HomName and specific-word frequency controls to be larger for the Italian than for the English speakers. This is because, in English, the HomName words should have the benefit of their homophone mates, whereas this benefit does not exist for their translations in Italian.

The other potentially confounding factor is that the HomName words are named more slowly than the other words because of their articulatory structure. Perhaps the articulatory programs of HomName words are compiled and executed more slowly than those of the control words, or the HomName words triggered the voice key later than the control words. These confounding factors might have made the existence of a homophonic effect invisible. To assess this possibility, we used the delayed-naming task (e.g., Balota & Chumbley, 1985; but see Goldinger, Azuma, Abramson, & Jain, 1997). In this task, words are named after a short interval, which gives speakers enough time to recognize the word and retrieve its name. The naming latencies observed in this task are assumed to reflect articulatory processing. A group of English speakers saw the written names of the pictures of Experiment 1A and named these words after a 1-s interval. In previous studies, no effects of frequency (e.g., Jescheniak & Levelt, 1994, Experiments 3 and 7; Monsell, Doyle, & Haggard, 1989; Savage, Bradley, & Forster, 1990) or AoA (Ellis & Morrison, 1998) were found in delayed-naming (1–4 s) conditions. If the lack of the homophone effect observed in Experiment 1A was due to articulatory factors, response latencies should be longer for HomName pictures than for specific-word frequency and for cumulative-homophone frequency pictures in the delayed-naming task. This prediction was tested in Experiment 1C.

Experiment 1B: Picture Naming in Italian

Method

Participants. Seventeen native Italian speakers who were students at the University of Padua, Padua, Italy, participated in Experiment 1B.

Materials and procedure. We excluded five of the pictures of the English version of the experiment, because they depicted concepts that are unfamiliar to Italian speakers (e.g., *skunk*), or because in Italian they have more than one commonly used name (e.g., *skull* can be named either *cranio* or *teschio*). Once we eliminated these pictures, 10 pictures remained unpaired and were therefore also excluded. Thus, we were left with 21 of the 26 triplets of HomName, specific-word frequency, and cumulative-homophone frequency pictures used in Experiment 1A. The words retained for Experiment 1B are shown in Appendix A. The Italian names of HomName, specific-word frequency, and cumulative-homophone frequency pictures were comparable in terms of frequency (M = 87, 50, and 91, respectively; $F < 1$; norms from Istituto Italiano di Linguistica) and length (number of letters; $F < 1$). Italian speakers were also shown the fillers used in Experiment 1A ($N = 52$). Procedure and analyses were identical to the ones of Experiment 1A (see *Method* section).

Results and Discussion

Errors and outliers accounted for 0.7% of the responses. Mean naming latencies and error rates for the HomName, specific-word frequency, and cumulative-homophone frequency pictures are shown in Table 3. Error rate decreased with repetition, $F1(2, 32) = 5.9$, $MSE = 0.4$, $p < .01$. The results of the ANOVAs indicated that naming latencies varied across picture sets, $F1(2, 32) = 14$, $MSE = 2{,}519$, $p < .0001$; $F2(2, 60) = 2.9$, $MSE = 15{,}980$, $p = .05$, and became faster with repetition, $F1(2, 32) = 7.0$, $MSE = 3{,}452$, $p < .01$; $F2(2, 60) = 18$, $MSE = 1{,}598$, $p < .01$. The interaction between picture set and repetition was not significant, $F1(4, 64) = 1.0$, $MSE = 634$, *ns* ($F2 < 1$). Additional analyses revealed that naming latencies were faster for cumulative-homophone frequency than for HomName pictures, $F1(1, 16) = 14$, $MSE = 2{,}292$, $p = .001$; $F2(1, 40) = 2.6$, 14,597; $p = .11$, and for specific-word frequency pictures, $F1(1, 16) = 23$, $MSE = 2{,}953$, $p < .001$; $F2(1, 40) = 5.5$, $MSE = 16{,}687$, $p = .02$. HomName pictures were named as fast as pictures matched for specific-word frequency (both $Fs < 1$). This result does not support the hypothesis that HomName pictures were particularly difficult to recognize. If anything, naming latencies for the specific-word frequency pictures were slower than for the HomName pictures. Thus, we have no evidence that a true homophone frequency effect (for the English HomName pictures) is masked by the fact that the HomName pictures are more difficult to recognize than the specific-word frequency pictures.

Experiment 1C: Delayed Naming in English

Method

Participants. Ten native English speakers who were students at Harvard University participated in Experiment 1C.

Materials and procedure. Three sets of written words ($N = 376$) were shown in Experiment 1C: (a) the names of HomName, specific-word frequency, and cumulative-homophone frequency pictures used in Experiment 1A ($N = 78$); (b) the experimental words used in Experiment 3A ($N = 107$); and (c) fillers (stimuli that were used as fillers in Experiments

Table 3
Mean Naming Latencies (and Error Rates) for the Pictures of Experiment 1B

Picture set	Italian name	Presentation 1	Presentation 2	Presentation 3	Average
Homophone name pictures	*suora* (nun)	787 (0.7)	768 (0.6)	747 (0.6)	767 (0.6)
Specific-word frequency matched pictures	*gufo* (owl)	812 (1.1)	784 (0.6)	755 (0.5)	784 (0.7)
Homophone frequency matched pictures	*dente* (tooth)	746 (0.7)	736 (0.8)	712 (0.4)	731 (0.6)

Note. English translations of Italian names are shown in parentheses.

1A and 3A; $N = 182$). Words were printed in upper case, with Geneva 20-point bold font. In each trial, a fixation point (a cross) was shown for 700 ms, and was immediately replaced by the written word, which remained in view for 600 ms. Following a blank interval, a cue (an asterisk) appeared and participants named the word. The interval lasted 1 s for the words of Sets A and B, and 700 or 1,300 ms for the words of Set C. Intervals varied to prevent participants from anticipating the appearance of the cue. Participants were instructed to prepare their response and, when the cue appeared, to name the word as fast as possible. Participants proceeded to the next trial by pressing the space bar. The results of the stimuli of Set B will be presented in Experiment 3C. Fillers were not included in the analyses.

Results and Discussion

In the analyses that follow, we only considered the data from the 75 words analyzed in Experiment 1, excluding therefore the 3 words that were also excluded in that experiment (one from each condition). Errors, responses that were too fast (< 200 ms) or too slow (>1,800 ms) and those that exceeded participants' means by three standard deviations were excluded from the analyses (2.6% of the data). The same exclusionary criteria were applied to the other delayed-naming experiments we report below. As can be seen in Table 4, production latencies were not statistically different for the names of the HomName, specific-word frequency, and cumulative-homophone frequency pictures ($ps < .2$). Our results are in line with those of other delayed-naming tasks, which also demonstrated an absence of word frequency or AoA effects (e.g., Ellis & Morrison, 1998; Jescheniak & Levelt, 1994; Monsell et al., 1989; Savage et al., 1990). The results of the delayed-naming experiment have direct implications for the picture-naming task of Experiment 1A. The results do not support the hypothesis that we failed to obtain a homophone effect in Experiment 1A because the names of the HomName pictures were articulated slower and/or they triggered the microphone later than the other items.

Summary of Experiments 1A–1C

Experiment 1A revealed that the naming latencies of HomName pictures (e.g., *nun*), which have a higher frequency mate (*none*), are a function of specific-word frequency and not of cumulative-homophone frequency. The results of Experiments 1B and 1C exclude the possibility that the results of Experiment 1A arose because we accidentally selected HomName pictures that were especially hard to recognize, or whose names took longer to articulate than their specific-word frequency controls. Our results contrast with those of Jescheniak and Levelt (1994), who reported that cumulative-homophone frequency predicts production latencies for HomName words in a translation task. Therefore, it is prudent to attempt to replicate our picture-naming experiment with a new set of stimuli. Unfortunately, we could not find a sufficient number of pictures whose names, in English, met the constraints for designing a properly controlled experiment. We decided instead to carry out a replication in Chinese.

Experiment 2A: Picture Naming in Chinese

In Experiment 2A, we examined the effect of homophone frequency on picture naming in Chinese (Mandarin). The experiment was modeled after Experiment 1A and thus included three sets of stimuli: HomName, specific-word frequency, and cumulative-homophone frequency pictures. The experimental question addressed here is whether specific-word frequency or cumulative-homophone frequency best predicts naming latencies. We addressed this question by creating three sets of words that differed in the degree to which they were comparable on the dimension of specific-word frequency versus cumulative-homophone frequency. That is, we constructed word sets that met two criteria: (a) the specific-word frequencies of HomName and specific-word frequency controls were similar, but lower than that of the cumulative-homophone frequency controls (see means in Table 5), and (b) the cumulative-homophone frequencies of HomName and cumulative-homophone frequency pictures were high compared with that of specific-word frequency pictures.

Method

Participants. Twenty-eight native Mandarin speakers who were students at Beijing Normal University, Beijing, China, participated in Experiment 2A.

Materials and procedure. Thirty-two pictures were selected for each picture set (HomName, specific-word frequency, and cumulative-homophone frequency). Only words with identical segments and tone were considered homophones. All pictures had monomorphemic names (see list in Appendix B). The specific-word frequencies of the names of HomName and specific-word frequency pictures were similar ($M = 46$ vs. 61, respectively; $F < 1$; norms from Xiandai Hanyu Pinlv Cidian, 1986) but less frequent than the names of cumulative-homophone frequency pictures ($M = 737$; both $ps < .06$). The cumulative-homophone frequencies of the names of HomName and cumulative-homophone frequency pictures did not differ (M = 1327 vs. 1897, respectively; $p > .10$) but were more frequent than the names of specific-word frequency pictures ($M = 118$; both $ps < .01$). We also included 21 fillers, which were not examined in any of the analyses. Procedure and analyses were identical to the ones described in Experiment 1A. Recording of naming latencies was controlled by the dual-screen version of DMASTR (Forster & Forster, 1990).

Results and Discussion

Following the same criteria as in Experiment 1A, 2.6% of the data were excluded from the analyses. Table 6 shows the

Table 4
Mean Naming Latencies and Error Rates for the Words of Experiment 1C

Written word	Example	Naming latency	Error rate
Homophone words	nun	390	3.2
Specific-word frequency matched words	owl	387	2.0
Homophone frequency matched words	tooth	377	2.5

Table 5
Mean Frequency (Specific and Cumulative) for the Pictures of Experiment 2A

Picture set	Chinese name	Specific frequency	Cumulative frequency
Homophone name pictures	桃 (peach)	46 (8–145)	1,327 (218–4,329)
Specific-word frequency matched pictures	刷 (brush)	61 (3–145)	118 (12–262)
Homophone frequency matched pictures	床 (bed)	737 (223–2,202)	1,897 (252–19,372)

Note. English translations of Chinese names are shown in parentheses. Frequency ranges are shown in parentheses.

distribution of mean response latencies and errors, as a function of picture set (HomName vs. specific-word frequency vs. cumulative-homophone frequency pictures) and presentation (first vs. second vs. third). Both the main effects of picture set, $F1(2, 54) = 100$, $MSE = 831$, $p < .0001$; $F2(2, 93) = 5.6$, $MSE = 18{,}142$, $p < .005$, and presentation, $F1(2, 54) = 64$, $MSE = 1{,}917$, $p < .0001$; $F2(2, 186) = 183$, $MSE = 847$, $p < .0001$, were significant. There was no evidence of interaction between these variables ($Fs < 1$). To determine the extent to which responses varied across picture sets, we carried out additional analyses, in which we directly contrasted the naming latencies observed in the various sets. Cumulative-homophone frequency pictures were named significantly faster than HomName pictures, $F1(1, 27) = 156$, $MSE = 1{,}067$, $p < .0001$; $F2(1, 62) = 11$, $MSE = 18{,}214$, $p < .002$. HomName pictures were named slightly slower than specific-word frequency pictures, a difference that reached significance in the analysis by subject, $F1(1, 27) = 64$, $MSE = 710$, $p < .001$, but not in the analysis by item, $F2(1, 62) = 2.6$, $MSE = 20{,}845$, $p = .10$. The 32-ms difference between cumulative-homophone frequency and specific-word frequency pictures was significant in the subject analysis, $F1(1, 27) = 54$, $MSE = 717$, $p < .0001$, and marginally significant in the item analysis, $F2(1, 62) = 3.1$, $MSE = 15{,}369$, $p < .08$ (see Figure 3).

As in Experiment 1A, naming latencies were entered into a regression analysis in which specific-word frequency and cumulative-homophone frequency were treated as independent variables. Cumulative-homophone frequency explained virtually none of the variance ($R^2 = .001$). In contrast, specific-word frequency accounted for a significant proportion of the variance ($R^2 = .15$, $p < .0001$). Furthermore, the variance accounted for by the two variables together is not significantly larger than that explained by specific-word frequency alone (Model 1, specific-word frequency: $R^2 = .15$; Model 2, specific-word frequency and cumulative-homophone frequency: $R^2 = .174$).

The results show that HomName pictures were named slower than the cumulative-homophone frequency controls. This result parallels that found in English (Experiment 1A). However, HomName pictures were also named slower than the specific-word frequency controls. This may reflect the fact that the recognition of HomName pictures or the articulation of their names was particularly difficult. That is, it may be that there is a homophone frequency effect, but the effect is masked by the difficulties in recognizing HomName pictures. Therefore, as argued earlier for the English variant of this experiment, if we eliminate the homophone status of the HomName pictures, one should expect to observe even larger differences between the HomName and the specific-word frequency pictures. This hypothesis is examined in the next two experiments. In Experiment 2B, the pictures used in the Chinese experiment were shown to English speakers. If the hypothesis that HomName pictures are especially difficult to recognize were correct, we should observe the difference between HomName and specific-word frequency matched controls to be larger than that observed in Experiment 2A. This is because, in Experiment 2A, HomName words were expected to benefit from having high-frequency homophone mates. In Experiment 2C, the Chinese characters corresponding to the picture names were shown in a delayed-naming task. This experiment was designed to evaluate the possibility that the effects observed in Experiment 2A reflect the ease with which their names can be articulated.

Table 6
Mean Naming Latencies (and Error Rates) for the Pictures of Experiment 2A

		Presentation			
Picture set	Chinese name	1	2	3	Average
Homophone name pictures	桃 (peach)	828 (3.5)	775 (2.7)	746 (1.7)	783 (2.6)
Specific-word frequency matched pictures	刷 (brush)	796 (3.7)	739 (3.8)	712 (2.1)	749 (3.2)
Homophone frequency matched pictures	床 (bed)	756 (2.0)	713 (2.8)	683 (1.6)	717 (2.1)

Note. English translations of Chinese names are shown in parentheses.

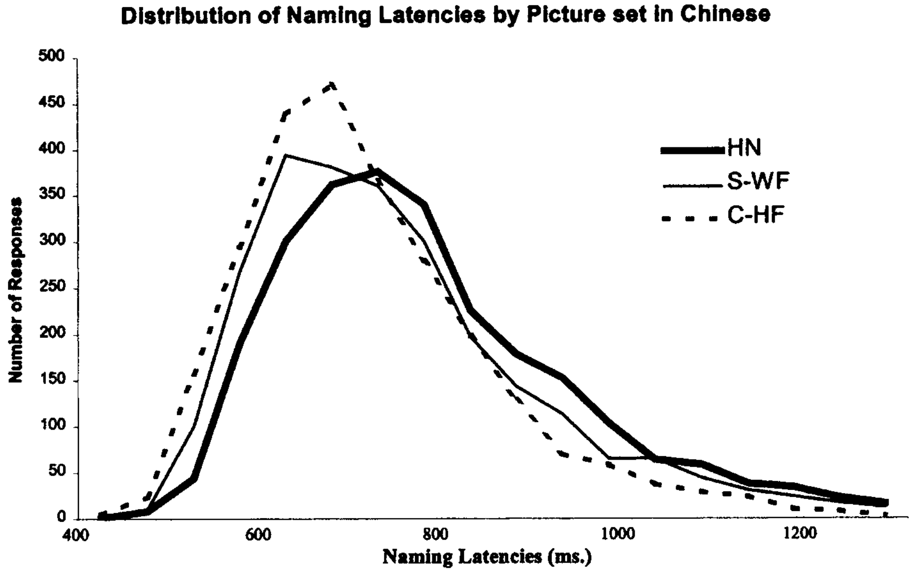

Figure 3. Distribution of naming latencies in Experiment 2A. The size of the interval is 50 ms. The lower interval begins at 400 ms, the highest interval ends at 1,300 ms. HN = homophone name; S-WF = specific-word frequency; C-HF = cumulative-homophone frequency.

Experiment 2B: Picture Naming in English

Method

Participants. Seventeen native English speakers who were students at Harvard University participated in Experiment 2B.

Materials. One picture used in the Chinese version of the experiment was excluded because there were two equally plausible alternative names in English (the picture *bean* could be named either *bean* or *peas*). The stimuli paired with it were also removed from the experimental set. Thus, the sets of HomName, specific-word frequency, and cumulative-homophone frequency pictures were each composed of 31 items. Because many of the English picture names are homophones, we examined the specific-word frequency and the cumulative-homophone frequency of all three sets of words (see Table 7). HomName and specific-word frequency pictures were comparable in terms of specific-word frequency and cumulative-homophone frequency ($ps > .3$). Cumulative-homophone frequency pictures were higher in both specific-word and cumulative-homophone frequency than the other pictures ($ps < .001$). The fillers used in Experiment 2A ($N = 21$) were also used in Experiment 2B. The procedure was identical to that of Experiment 2A.

Results and Discussion

Following the same criteria as in Experiment 1A, 1.9% of the data points were excluded from the analyses. The results of Experiment 2B are summarized in Table 8. As in the preceding experiments, we examined two variables: picture set (HomName vs. specific-word frequency vs. cumulative-homophone frequency) and presentation (first vs. second vs. third). The main effect of presentation was significant, $F1(2, 32) = 16$, $MSE = 26{,}367, p < .0001$; $F2\ (2, 180) = 45, MSE = 54{,}788, p < .0001$, a result reflecting a decrease of response latencies with repetition. The main effect of picture set was also significant, $F1(2, 32) = 45$, $MSE = 91{,}251$, $p < .0001$; $F2(2, 90) = 7.8$, $MSE = 172{,}461$, $p < .001$. There was no evidence of interaction between the two variables ($Fs < 1$). Pairwise comparisons showed that cumulative-homophone frequency pictures were named faster than both HomName pictures, $F1(1, 16) = 54$, $MSE = 33{,}009$, $p < .0001$; $F2(1, 60) = 18$, $MSE = 18{,}642$, $p < .001$, and specific-word frequency pictures, $F1(1, 16) = 16$, $MSE = 34{,}542$, $p = .0001$; $F2(1, 60) = 6.15$, $MSE = 19{,}190$, $p < .02$. The higher frequency of cumulative-homophone frequency pictures is the most likely explanation for this result. HomName pictures were named 35 ms slower than specific-word frequency pictures, $F1(1, 16) = 111$, $MSE = 47{,}693$, $p < .01$; $F2(1, 60) = 2.0$, $MSE = 28{,}275$, $p < .16$. This difference was similar to the one observed between the two sets of pictures in Experiment 2A (34 ms). As argued above, if the homophone status of the pictures' names were

Table 7
Mean Frequency (Specific and Cumulative) for the English Picture Names of Experiment 2B

Picture set	Example	Specific frequency	Cumulative frequency
Homophone name pictures	peach	23 (1–207)	51 (1–513)
Specific-word frequency matched pictures	brush	34 (1–147)	46 (1–293)
Homophone frequency matched pictures	bed	178 (6–717)	467 (7–8,925)

Note. Frequency ranges are shown in parentheses.

Table 8
Mean Naming Latencies (and Error Rates) for the Pictures of Experiment 2B

		Presentation			
Picture set	Example	1	2	3	Average
Homophone name pictures	peach	767 (3.0)	732 (2.1)	713 (1.3)	737 (2.1)
Specific-word frequency matched pictures	brush	726 (4.1)	703 (1.9)	676 (1.0)	702 (2.3)
Homophone frequency matched pictures	bed	673 (1.9)	649 (1.3)	632 (1.2)	651 (1.5)

to play a role in naming latencies, the difference between the HomName and the specific-word frequency controls would have been expected to be smaller in Experiment 2A than in Experiment 2B. The results show that the difference in naming latencies between the two sets of pictures is independent of their difference in cumulative-homophone frequency. Therefore, this result allows us to dismiss the hypothesis that the lack of a homophone frequency effect in Chinese is an artifact of uncontrolled differences in relative difficulties in recognizing the pictures in the three sets of stimuli.

Experiment 2C: Delayed Naming in Chinese

In Experiment 2C, we presented Chinese speakers with the written names of the pictures used in Experiment 2A and instructed them to name them when a cue appeared. This task served as a control for possible effects of articulation difficulty in naming the experimental pictures.

Method

Participants. Sixteen Mandarin Chinese native speakers who were students at Beijing Normal University, Beijing, China, participated in Experiment 2C.

Materials and procedure. The Chinese characters for the names of the HomName, specific-word frequency, cumulative-homophone frequency, and filler pictures used in Experiment 2A were included in the experiment (see Appendix B). The same procedure as in Experiment 1C was used. Participants were instructed to name the characters (in Mandarin) at the presentation of a cue (an asterisk). The stimulus–cue interval varied: It was set to 1 s for the characters corresponding to the HomName, specific-word frequency, and cumulative-homophone frequency pictures, and to 700 or 1,300 ms for the fillers. The Chinese characters were shown in 48-point Songti font, and were about 2.4 × 1.6 cm in size. The equipment and presentation software were the same as in Experiment 2A.

Results and Discussion

Following the same criteria as in Experiment 1C, 3.9% of the data points were discarded. Table 9 shows the mean naming latencies and error rates for HomName, specific-word frequency, and cumulative-homophone frequency characters. Naming latencies were not statistically different across stimuli sets ($F\text{s} < 1$), a result that suggests that all articulatory routines were similarly accessible for the three sets of Chinese words used in Experiment 2A.

Summary of Experiments 2A–2C

As in English, picture-naming latencies in Chinese are determined by specific-word frequency rather than cumulative-homophone frequency. This pattern of results is not due to uncontrolled differences in picture recognition or ease of articulation among stimulus sets. When we assessed these possibilities, we found no indications that they could account for the Chinese data. The fact that analogous results were obtained in two languages (English and Chinese) increases our confidence in the conclusion that homophone frequency does not affect picture naming. This conclusion is at variance with Jescheniak and Levelt (1994), who observed a homophone frequency effect with a word translation paradigm. In an attempt to clarify the source of this discrepancy, we carried out a replication of the translation experiment of Jescheniak and Levelt (1994).

Experiment 3A: Spanish–English Translation Task

In this experiment, English–Spanish bilingual speakers were instructed to translate Spanish words into English. We selected three sets of English words: HomName words (e.g., *hare*), controls matched for specific-word frequency (e.g., *plum*), and controls

Table 9
Mean Naming Latencies and Error Rates for the Words of Experiment 2C

Written word	Chinese word	Naming latency	Error rate
Homophone words	桃 (peach)	382	4.7
Specific-word frequency matched words	刷 (brush)	372	3.8
Homophone frequency matched words	床 (bed)	372	3.4

Note. English translations of Chinese words are shown in parentheses.

Table 10
Mean Frequency and Length for the English Words and Their Spanish Translations Used in Experiment 3A

Picture set	Examples	Specific frequency	Cumulative frequency	Familiarity
		English words		
Homophone name pictures	hare	21 (0–95)	1,478 (67–6,990)	4.4
Specific-word frequency matched pictures	plum	20 (0–95)	22 (0–127)	4.5
Homophone frequency matched pictures	tree	1,559 (54–8,996)	1,580 (54–9,362)	4.2
		Spanish translations		
Homophone name pictures	*liebre*	123		5.8
Specific-word frequency matched pictures	*ciruela*	99		6.1
Homophone frequency matched pictures	*árbol*	5,934		5.2

Note. Frequency ranges are shown in parentheses.

matched for cumulative-homophone frequency (e.g., *tree*).[8] The principal aim of the experiment was to examine whether the translation latencies for HomName words were comparable with those of the control stimuli matched for cumulative-homophone frequency, and faster than the translation latencies of the control stimuli matched for specific-word frequency.

Method

Participants. Twenty English–Spanish bilinguals participated in Experiment 3A. Participants were graduate or undergraduate students in one of the universities in the Boston area. They were native speakers of English with excellent knowledge of Spanish. Participants reported to have lived in a Spanish-speaking country for at least 1 year and to have studied Spanish for at least 6 years.

Materials. Three groups of 22 English monomorphemic words formed the sets of HomName, specific-word frequency, and cumulative-homophone frequency stimuli (see list in Appendix C). These words were selected according to the criteria used by Jescheniak and Levelt (1994) and we used them in Experiments 1A and 2A.[9] The means and ranges of specific-word and cumulative-homophone frequencies for the English words are reported in Table 10. Mean specific-word frequencies were the same for HomName and specific-word frequency words ($F < 1$), and significantly higher for cumulative-homophone frequency words ($ps < .01$). Mean cumulative-homophone frequencies were the same for HomName and cumulative-homophone frequency words ($F < 1$), but significantly lower for specific-word frequency words ($ps < .01$). (The means reported in Table 10 are for the 20 items retained for analysis. Two homophones and their associated controls were discarded because of problems in the selection of materials.) The three sets of English words were comparable in length (number of letters; $F < 1$), as were their Spanish translations, $F(1, 38) = 1.1$, $MSE = 2.0$, $p < .3$ (see means in Table 10). The Spanish translations of the cumulative-homophone words were higher in frequency than the Spanish translations of the other words ($ps < .03$; frequency norms are from Sebastian, Marti, Cuetos, & Carreiras, 1996). In Spanish, HomName and specific-word frequency words do not differ on frequency values ($F < 1$). To reduce the proportion of English homophonic words, we also showed 150 filler words (which were not included in any of the analyses). Thus, participants translated a total of 216 words. Words were printed in Geneva 20-point font.

Procedure. At the beginning of the experiment, participants read the printed list of Spanish words and their English translations. They were then instructed to produce the English words included in the list. Instructions were written in English and were read by the participants. Before the experiment proper, participants translated once the whole set of 216 words as fast as they could, without making mistakes. These words were shown a second time during the experiment. On each presentation, the words were divided into four blocks of 54 words. Words from the three sets were equally represented across blocks. The words were randomized, with the constraints that words from the same list would not appear in consecutive trials, and that only filler words were shown in the initial three trials of each block. Three randomizations were used (one for the practice and two for the experiment proper). Block order of presentation was randomized for each participant. Each trial had the following structure: Participants started the trial by pressing the space bar; a fixation point (a cross) was then shown in the center of the screen for 400 ms and was immediately followed by a Spanish word; the word remained on the screen for 500 ms. The equipment used was that described in Experiment 1A. Response accuracy was manually recorded by the experimenter. The entire experimental session lasted approximately 50 min. The procedure followed for analyzing the responses was the same as described in Experiment 1A (see *Method* section). One variable was examined: word set (HomName vs. specific-word frequency vs. cumulative-homophone frequency words), which was treated as a within- and between-subject variable in the $F1$ and $F2$ analyses, respectively. Two of the HomName words were discarded from analysis because of a problem in the selection of materials. Their paired control words were also excluded from analysis (see Appendix C).

[8] We retain the same terminology as that used for the picture-naming experiments, even though it is somewhat stilted. The reason for this use is that it makes comparisons across experiments more transparent.

[9] As already noted, however, one difference between the words used by Jescheniak and Levelt (1994) and the present experiment is that the Dutch homophones were also homographs, whereas the English homophones could be either homographs or heterographs.

Results and Discussion

Following the same criteria as in Experiment 1A, 8.6% of the data were discarded from the analyses. Mean translation latencies and error rates for the various word sets are presented in Table 11. There was a significant effect of word set, $F1(2, 38) = 51$, $MSE = 5{,}601, p < .001$; $F2(2, 57) = 15, MSE = 21{,}688, p < .001$. Pairwise comparisons revealed that naming latencies were faster for cumulative-homophone frequency words than for both specific-word frequency words and HomName words ($ps < .001$). There was no difference between the translation times for HomName words and specific-word frequency words ($Fs < 1$). Errors were unequally distributed across word sets, $F1(2, 38) = 15$, $MSE = 1.5, p < .001$; $F2(2, 57) = 5.6, MSE = 4.2, p < .002$. The latter result is in part explained by the fact that cumulative-homophone frequency words induced fewer errors than the words of the other conditions (for all Fs, $p < .05$). The difference in error rate between HomName (14.6%) words and specific-word frequency words (8.1%) was significant in the $F1$ analysis, $F1(1, 19) = 8.0$, $MSE = 1.9$, $p < .01$, but not in the $F2$ analysis, $F2(1, 38) = 2.7$, $MSE = 5.7$, $p < .1$.

As in the previous experiments, we carried out a regression analysis on naming latencies with specific-word frequency and cumulative-homophone frequency as predictors. Specific-word frequency is a better predictor ($R^2 = .48$) than cumulative-homophone frequency ($R^2 = .16$). Furthermore, when specific-word frequency was introduced first in the regression model, there was essentially no gain in explained variance by adding cumulative-homophone frequency (Model 1, specific-word frequency: $R^2 = .48$; Model 2, specific-word frequency and cumulative-homophone frequency, $R^2 = .49$).

We also analyzed the relation between AoA and translation latencies. As in Experiment 1A, we considered both specific and MinHomophone AoA. Specific-word AoA is a better predictor of naming latencies than MinHomophone AoA ($R^2 = .45$ vs. .16, respectively). In a stepwise regression analysis, the inclusion of MinHomophone AoA does not add appreciably to the variance accounted for by specific-word AoA (Model 1, specific-word AoA: $R^2 = .45$; Model 2, specific-word AoA and MinHomophone AoA: $R^2 = .46$). An additional stepwise regression analysis investigated whether AoA and frequency both contributed to the observed translation latencies. Both variables significantly contributed to the variance accounted for in production latencies (specific-word frequency: $R^2 = .486$; specific-word frequency and specific-word AoA: $R^2 = .553$).

The results of Experiment 3A contrast sharply with those obtained by Jescheniak and Levelt (1994). They found that HomName words (e.g., *hare*) were translated as fast as cumulative-homophone frequency words (e.g., *tree*), whereas we found that translation latencies for HomName words were not statistically different from words matched for specific-word frequency (e.g., *plum*). However, the complexity of the translation task is such that interpretation of our results must proceed cautiously, at least until we have ruled out the contribution of possible confounding factors. One factor relates to differences in word recognition. For example, if it took disproportionately longer to recognize the Spanish words for the HomName items, we might not be able to detect a homophone frequency effect, even if it were present. This possibility was examined in Experiment 3B, in which the Spanish words of Experiment 3A were presented to Spanish speakers for lexical decision. If the failure to observe a homophone frequency effect was because the Spanish stimuli for the HomName words were recognized relatively slowly, these words should produce slower decision latencies than their matched specific-word frequency controls. Alternatively, it could be that our failure to replicate the effect of homophone frequency is attributable to differences in the ease of articulation of the three word sets; namely, perhaps it is more difficult and it takes longer to articulate HomName words than their specific-word frequency controls. As in Experiments 1 and 2, we used a delayed-naming task (see Experiment 3C) to assess the possibility that differences in articulation difficulty are responsible for the effects obtained in Experiment 3A.

Experiment 3B: Lexical-Decision Task With Spanish Words

Method

Participants. Twenty native Spanish speakers who were students at the University of Barcelona, Barcelona, Spain, participated in the Experiment 3B in exchange for course credit.

Materials and procedure. The material of Experiment 3B included (a) the 66 Spanish experimental words used in Experiment 3A and (b) 66 legal nonwords. The latter were created by changing one letter in Spanish words (e.g., *brazo* [*arm*] → *blazo*). With two exceptions, the procedure was identical to that of Experiment 3A: (a) words were shown on the computer screen for 1,000 ms and (b) participants responded by pressing computer keys. *Word* and *non-word* responses were assigned to participants' dominant and nondominant hands, respectively. Participants were instructed to indicate as fast as they could, while preserving accuracy, whether the string of letters corresponded to a Spanish word. Participants were presented with a practice block of 15 words and 15 nonwords (these stimuli were not included in the experiment proper). Words and nonwords were shown only once. We excluded from analysis the 6 words that were also discarded in the analysis of Experiment 3A.

Table 11
Mean Response Latencies and Error Rates in the Spanish–English Translation Task (Experiment 3A)

Written word	Stimulus–response	Translation latency	Error rate
Homophone words	*liebre*–hare	1,058	14.6
Specific-word frequency matched words	*ciruela*–plum	1,060	8.1
Homophone frequency matched words	*árbol*–tree	852	3.2

Results and Discussion

Responses shorter than 200 ms, longer than 1,500 ms, or that exceeded a participant's mean by three standard deviations were excluded from the analyses (3.2 % of the data). Table 12 provides a summary of the results of Experiment 3B. ANOVAs revealed a significant effect of word set in the subject analysis but not in the item analysis, $F1(2, 38) = 11$, $MSE = 733$, $p < .001$; $F2(2, 57) = 2.3$, $MSE = 3{,}326$, $p < .11$. Pairwise comparisons revealed that decision latencies were faster for cumulative-homophone frequency words than for HomName words, $F1(1, 19) = 17$, $MSE = 825$, $p < .01$; $F2(1, 38) = 4.7$, $MSE = 2{,}796$, $p < .03$, and also faster than specific-word frequency words, $F1(1, 19) = 12$, $MSE = 849$, $p < .002$; $F2(1, 38) = 2.6$, $MSE = 3{,}713$, $p < .01$. Of particular importance here is that decision latencies that were not statistically different were found for the Spanish words of the HomName and specific-word frequency sets ($Fs < 1$). The latter result suggests that the Spanish words in the HomName set were recognized as easily as other Spanish words matched on specific-word frequency. By further inference, we can conclude that it is unlikely that the lack of a homophone frequency effect in Experiment 3A reflects differences in recognizing the Spanish words.

Experiment 3C: Delayed Naming in English

Method

For Experiment 3C, we used the English words of Experiment 3A (66 experimental items and 150 fillers). These words were shown along with the words of Experiment 1A. See the *Procedure* of Experiment 1A. The same participants as in Experiment 1C participated in this experiment.

Results and Discussion

Only the words analyzed in Experiment 3A were analyzed in this experiment. Following the same criteria as in Experiment 1C, 1.2% of the data were excluded from the analyses. Table 13 shows the mean naming latencies and error rates observed in Experiment 3C. Naming latencies were not statistically different across the three sets of English words tested in Experiment 3A (HomName, specific-word frequency, and cumulative-homophone frequency words; $Fs < 1$). This finding has immediate implications for the interpretation of the results of Experiment 3A: It rules out the possibility that the translation latencies observed for HomName words were due to features that slowed the articulatory processing of these words.

Summary of Experiments 3A–3C

The translation latencies of HomName words (e.g., *hare*) were comparable with those found for control words matched on specific-word frequency (e.g., *plum*), and were significantly slower than those found for control words matched on cumulative-homophone frequency (e.g., *tree*). This pattern of results is not merely a consequence of confounding characteristics of the stimuli. Experiments 3B and 3C tested and rejected the possibility that the absence of a homophone frequency effect in Experiment 3A is the result of differences in ease of recognition or differences in ease of articulation between the HomName and specific-word frequency word sets. In short, the results of Experiments 3A–3C show that translation times are affected by specific-word frequency, and not by cumulative-homophone frequency—the opposite finding of that reported by Jescheniak and Levelt (1994) with the same task.

General Discussion

In three sets of experiments, we investigated whether naming latencies for homophonic words (e.g., *nun*) are a function of specific-word frequency (i.e., the frequency of *nun*) or a function of cumulative-homophone frequency (i.e., the sum of the frequencies of *nun* and *none*). In Experiment 1A, English-speaking participants named three sets of pictures: (a) pictures whose names (HomName) have discrepant specific-word and cumulative-homophone frequencies; (b) pictures whose names match the specific-word frequency of the HomName pictures; and (c) pictures whose names match the cumulative-homophone frequency of the HomName pictures. The results of this experiment are clear: There was no difference between the naming latencies of HomName and specific-word frequency control pictures, but both sets of pictures were named slower than the cumulative-homophone frequency control pictures. These results show that naming latencies for homophonic words are determined by their specific-word frequencies and not by their cumulative-homophone frequencies. That is, no benefit accrues to a word's naming latency from having a homophone mate with higher frequency.

These results were fully replicated in Experiment 2A in a different language, Chinese. Further support for the observation that naming latencies are not a function of cumulative-homophone frequency was obtained in Experiment 3A, in which we used the translation task used by Jescheniak and Levelt (1994). In the latter experiment, English-Spanish bilinguals were asked to translate three sets of words from Spanish into English. As in Experiments 1A and 2A, one set of words had English translations with highly discrepant specific-word and cumulative-homophone frequencies, whereas the other two sets were matched either to the specific-word frequency or to the cumulative-homophone frequency of the first set. The results of this experiment replicated those obtained in Experiments 1A and 2A.

Table 12
Mean Response Latencies and Error Rates in the Lexical-Decision Task With Spanish Words (Experiment 3B)

Written word	Spanish word	Naming latency	Error rate
Homophone words	*liebre* (hare)	691	3.7
Specific-word frequency matched words	*ciruela* (plum)	686	3.4
Homophone frequency matched words	*árbol* (tree)	654	2.5

Note. English translations of Spanish words are in parentheses.

Table 13
Mean Naming Latencies and Error Rates in Experiment 3C

Written word	Example	Naming latency	Error rate
Homophone words	hare	375	1.0
Specific-word frequency matched words	plum	371	1.4
Homophone frequency matched words	tree	379	1.3

The effects observed in the factorial analyses of the data (ANOVAs) are supported by the results of regression analyses. In the latter analyses, it was consistently found that naming latencies are better predicted by the specific-word frequency than the cumulative-homophone frequency variable. Furthermore, the gain in explained variance was never significant when cumulative-homophone frequency was included as a factor in the regression model.

Several control experiments were carried out in order to assess the contribution of articulatory factors to the observed effects (Experiments 1C, 2C, and 3C). The results of the control experiments showed no difference among the three sets of words, suggesting that the effects observed in the picture-naming tasks are not due to differences among word sets in initiating and executing articulatory programs.

Two other control experiments were carried out to assess the relevance of the homophone status of words in picture naming (Experiments 1B and 2B). In Experiment 1B, the materials used in Experiment 1A were presented for naming to Italian speakers. This manipulation neutralizes the homophone/nonhomophone distinction between the HomName, specific-word frequency, and cumulative-homophone frequency picture sets. That is, while in English the HomName set has highly discrepant specific-word versus cumulative-homophone frequencies by comparison with the specific-word frequency and cumulative-homophone frequency control sets, this difference disappears in Italian because the translated HomName words are not systematically homophonic.[10] Therefore, the comparison among word sets across languages (English and Italian) allows us to further test the importance of a word's homophone status in determining naming latencies. The results for the two languages were very similar, which suggests that the homophonic status of the words in English does not affect naming latencies. A similar control experiment was carried out, with similar results, for the Chinese materials with English speakers. Finally, a control experiment (Experiment 3B) ruled out the possibility that the absence of a homophone effect in the translation task (Experiment 3A) was due to uncontrolled differences in the ease with which the Spanish words could be recognized across conditions.

In sum, the results of the three sets of experiments reported here present a clear and consistent picture: Naming latencies are affected by the word's specific frequency and not by the cumulative frequency of its homophonic cohort. In other words, we have failed to find any evidence for a homophone frequency effect in speeded-naming tasks. This conclusion also holds if we consider AoA instead of frequency of usage as the relevant variable in determining speed of lexical access. We consistently found that the better predictor of naming latencies is the AoA of the word and not the minimum AoA of a homophone cohort.

The results of the experiments reported here are in conflict with the observations made by Jescheniak and Levelt (1994), who found a homophone frequency effect in a single experiment with Dutch speakers. That result was obtained with a complicated translation task, and with an especially small number of items (11 items per set). In contrast, we have systematically failed to replicate this result with two different tasks, and with larger numbers of stimulus items (between 20 and 32). Furthermore, we have obtained converging evidence from two languages (English and Chinese). Finally, the absence of a homophone frequency effect in our experiments cannot be attributed to a lack of power in the experiments, because we obtained the classic word frequency effect (as well as an effect of AoA) in both the picture-naming and the translation tasks.

It is unclear what might be the cause(s) for the contrasting results. One possibility is that they are due to the use of different languages in the experiments—Dutch versus English. As already noted, Dutch has a transparent orthography, and therefore the homophones used in the experiment by Jescheniak and Levelt (1994) were homographs. In our experiments, we used English and Chinese, which have rather opaque orthographies. A consequence of the latter fact is that many of the homophones in our experiments have heterographic spellings (e.g., *nun/none*). It could be argued that this difference between stimulus materials is responsible for the contrasting results. We checked for this possibility by reanalyzing the results of Experiment 1A, where we had a substantial proportion of homographic homophones (15 out of 25). Table 14 reports naming latencies for the homographic homophones and for the combined heterographic and homographic homophones in each of the three experimental conditions: HomName pictures, specific-word frequency controls, and cumulative-homophone frequency controls. As is immediately apparent upon inspection of Table 14, the pattern of mean naming latencies for the homographic homophones is not different from the pattern obtained for homophones in general. That is, the orthographic form of a homophone—whether it is a homograph or a heterograph—does not appear to contribute to variation in naming latencies for HomName pictures. It is unlikely, then, that the discrepancy in results reported here and those reported by Jescheniak and Levelt is due to language differences in orthographic transparency.[11]

[10] Only one Italian word in the translated HomName set has a high-frequency homophone, *sole*, which means either *sun* or *only* (feminine, plural).

[11] Furthermore, it should be noted that if it were to turn out that the orthographic status of a homophone played a role in lexical access in speech production, the SR hypothesis in its current formulation would be undermined. As mentioned several times already, this hypothesis does not distinguish between heterographic and homographic homophones.

Table 14
Mean Naming Latencies for the Pictures of Experiment 1A

Picture set	Example	Presentation 1	2	3	Average
Homophone name pictures	nun (watch)	755 (743)	782 (764)	754 (746)	764 (751)
Specific-word frequency matched pictures	owl (piano)	748 (750)	768 (770)	741 (743)	752 (754)
Homophone frequency matched pictures	tooth (table)	709 (714)	724 (726)	708 (706)	714 (716)

Note. Numbers represent the naming latencies combining heterographic (e.g., *nun/none*) and homographic (e.g., *watch*) homophones in Experiment 1A; numbers in parentheses represent naming latencies for only the homographic homophones included in that experiment.

Another possible reason for the different results obtained in our Experiment 1A and the experiment reported by Jescheniak and Levelt (1994) is that we included in our study homonyms that are related in meaning (e.g., *the anchor/to anchor; the nurse/to nurse; seven out of twenty-five*). It might be argued that these words are somehow processed differently from other homophones. However, a reanalysis of the data excluding these items did not affect the pattern of results (HomName: 763 ms; specific-word frequency: 761 ms; cumulative-word frequency: 715 ms).

Our results also contrast with those reported by Dell (1990). Dell investigated the occurrence of sound errors for homophone pairs formed by high-frequency function words and low-frequency content words, such as *him/hymn* and *would/wood.* In a post hoc analysis, Dell found that homophone frequency predicted the rate of sound errors for the lower frequency members of the homophone pairs. The reason for the contrasting results in Dell's experiment and in ours is not clear. We can point to obvious differences between the two studies. For example, we measured naming latencies, Dell measured error rates; we focused on open-class words, Dell compared open- and closed-class words; we used simple picture-naming and translation tasks, Dell used an error-inducing task in which participants were required to produce "simple phrases" (e.g., *him/hymn to sing*) as quickly as possible. However, it is not clear why any of these differences would lead to the observed differences in patterns of frequency effects. Perhaps a more plausible reason for the different results is that what Dell measured is the effect of the frequency of "phoneme sequences" (as opposed to lexical frequency) on the preservation of the integrity of phoneme sequences in a disruptive situation. Be this as it may, the difference in results calls for further investigation.

As discussed in the introduction, resolution of the issue of whether there is a homophone frequency effect would have important implications for models of lexical access. We have argued that clear evidence against the existence of a homophone frequency effect would help determine the possible combinations of assumptions that one can entertain in a model of speech production. In particular, the absence of a homophone frequency effect has important implications for those models that assume that homophones share a common lexical–phonological representation—the SR models (Cutting & Ferreira, 1999; Dell, 1990; Jescheniak & Levelt, 1994; Levelt et al., 1999).

For example, consider Levelt et al's. (1999) discrete-stage activation model of lexical access (see also Jescheniak & Levelt, 1994). The model assumes that homophones are represented by distinct lemma representations that converge onto a single lexeme node for each homophone cohort (see Figure 1A). The model also assumes that the locus of the frequency effect in naming is at the level of lexeme representations. This combination of assumptions predicts that naming latencies are a function of cumulative-homophone frequency and not specific-word frequency. The results of our experiments, which show that naming latencies are not a function of cumulative-homophone frequency but instead are determined by specific-word frequency, indicate that at least one of the assumptions of the model may be incorrect. There are various ways in which the model could be modified to accommodate our results.

Jescheniak and Levelt (1994) pointed out that in their model there are at least three possible loci for the frequency effect in naming: the lemma level, the lexeme level, or the lemma–lexeme connections. They also noted that in a discrete-stage activation theory such as theirs, only those models that locate the frequency effects at the level of lexeme representations predict a homophone frequency effect. Those models that locate the frequency effect either at the lemma level or at the level of the lemma–lexeme connections do not predict a homophone frequency effect. Therefore, a discrete-stage activation model of lexical access that locates the frequency effect at one of these levels and retains the shared representation assumption for homophones can account for the results of our experiments.

Another way in which Levelt et al.'s (1999) model could be modified so as to accommodate the specific-word frequency effect is to drop the assumption that homophones share a common representation. In this new model, each lemma node would be connected to a distinct lexeme node, regardless of whether or not the word is a homophone. By dropping the shared representation assumption, the model becomes an independent representation (IR) model. In a model of this type, the locus of the frequency effect could be located at any of the three levels considered by Jescheniak and Levelt (1994): viz., the lemma level, the lexeme level, or the lemma-lexeme connections. A homophone frequency effect is not expected in any of these cases.

Along the same lines, the lack of a homophone frequency effect is problematic for interactive activation models, whether they assume that homophones share a common representation. Dell's (1990) model is silent about the effect of frequency on the speed with which lexical nodes are selected, as well as its further impact on naming latencies. However, in the measure to which the model accounts for the word frequency effect in naming by postulating some activation advantage to higher frequency words, the model would also then predict a homophone frequency effect. This is because any advantage that accrues to a high-frequency lexical node is shared by nodes that are connected to it (see the introduction). Thus, it is not unreasonable to argue that Dell's model predicts a homophone frequency effect on naming latencies, and that the absence of such an effect in our experiments shows that some aspect of the model needs to be modified.

We have already noted that without explicit simulation, predictions about the behavior of interactive models can only be made very tentatively. That is because the actual behavior of a model depends on the specific values chosen for the various parameters of the model. This point can be easily appreciated by considering the consequences of progressively increasing (or decreasing) the feedback connection strength in such a model. When the feedback value is very small, the effects of interactivity can be quite insignificant. As the feedback connection strength increases, the effects of interactivity become progressively more important. It is possible, therefore, to find parameter values for an interactive activation model that predict only very small and not easily detectable effects of homophone frequency. This model would then be able to account for the absence of a homophone frequency effect in our experiments. However, note that while this is certainly possible, we would then have to see whether a model with these characteristics could also account for other naming data. Thus, for example, we know that interactive models are able to account for the lexical bias effect in the speech-error data, because they assume feedback activation.[12] The question then is whether it is possible to find parameter values for feedback activation that allow the model to predict both the existence of a lexical bias effect in speech-error data and the absence of a homophone frequency effect in naming latencies.

In short, the absence of a homophone frequency effect creates difficulties both for those models that assume that homophones are represented by a shared lexeme node, and for those models that postulate strong interactivity between levels of representations. By contrast, the results fit quite well with cascaded activation models that assume IRs for homophones (Caramazza, 1997).

We have argued that the fact that specific-word frequency and not cumulative-homophone frequency predicts naming latencies undermines the SR hypothesis of homophones. This conclusion has implications for the functional architecture of the lexical-access system in language production, and, more specifically, for the number of levels of lexical representation that need to be postulated. In models where homophones are represented by a shared lexeme node, there must be another level of lexical representation where the homophones have distinct lexical/grammatical representations—the lemma level. The distinction between the two levels of lexical representation is unavoidable, if we assume that homophones share a common lexeme representation. Therefore, the presence of a homophone frequency effect would both support the SR hypothesis of homophones and the lemma–lexeme distinction. And, in fact, the homophone frequency effect reported by Jescheniak and Levelt (1994) has been cited by Levelt et al. (1999; see also Levelt, 2000; Roelofs et al., 1998) as evidence for the need to distinguish between lemma and lexeme strata in the lexicon (Figure 1A), and against the single lexical layer model proposed by Caramazza (1997; see also Caramazza & Miozzo, 1997; Figure 1B in the present article). However, our results cast serious doubt on the existence of a homophone frequency effect. Instead, the results we have reported provide clear evidence in favor of a specific-word frequency effect in lexical access. This effect undermines the empirical motivation for the SR hypothesis of homophones. If we give up the SR hypothesis, we also remove perhaps one of the strongest arguments cited in favor of the lemma–lexeme distinction. Of course, there are other grounds on which one may want to motivate this distinction (for extensive discussion of these other data, see Dell, 1986; Garrett, 1988; Levelt et al., 1999; but also see Caramazza, 1997, for an opposing view). The point here is simply that the homophone frequency effect cannot be counted in the ledger of those facts that require an assumption of a lemma–lexeme distinction in lexical representation and access.

To conclude, in three sets of experiments we have shown that naming latencies are determined by specific-word frequency rather than by cumulative-homophone frequency. The specific-word frequency effect documented in this study raises difficulties for interactive activation models of lexical access and for models of lexical access that assume shared representations for homophones (and locate the effect of frequency in naming at the level of the shared homophone representation). The results provide support for IR models of homophones, and therefore undermine arguments that use the assumption of shared representations for homophones to support the lemma–lexeme distinction.

[12] This effect refers to the observation that slips of the tongue result in word errors more often than would be expected by chance (e.g., Dell & Reich, 1981; Martin, Weisberg, & Saffran, 1989; but see Garrett, 1988).

References

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language, 24,* 89–106.

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50*(A), 560–585.

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

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the "tip-of-the-tongue" phenomenon. *Cognition, 64,* 309–343.

Caramazza, A., & Miozzo, M. (1998). More is not always better. A response to Roelofs, Meyer, and Levelt. *Cognition , 69,* 231–241.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information 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 & Cognitive Processes, 5,* 313–349.

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review, 98,* 604–614.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior, 20,* 611–629.

Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 515–523.

Forster, K. I., & Forster, J. C. (1990). *User's guide to the DMASTR display system.* Unpublished software instructions, Department of Psychology, University of Arizona.

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

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Language: Psychological and biological aspects* (Vol. 3, pp. 69–96). Cambridge, England: Cambridge University Press.

Goldinger, S. D., Azuma, T., Abramson, M., & Jain, P. (1997). Open wide and say "blah!": Attentional dynamics of delayed naming. *Journal of Memory and Language, 37,* 190–216.

Griffin, Z. M., & Bock, K. (1998) Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38,* 313–338.

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

Hodgson, C., & Ellis, A. W. (1998). Last in, first to go: Age of acquisition and naming in the elderly. *Brain and Language, 64,* 146–163.

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.

Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior, 23,* 39–66.

Levelt, W. J. M (2000). Models of word production. *Trends in Cognitive Sciences, 3,*223–232.

Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience, 10,* 553–567.

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.

Lewis, M. B., Gerhand, S., & Ellis, H. D. (2001). Re-evaluating age-of-acquisition effects: Are they simply cumulative-frequency effects? *Cognition, 78,* 189–205.

Martin, N., Weisberg, R. W., & Saffran, E. M. (1989). Variables influencing the occurrence of naming errors: Implications for models of lexical retrieval. *Journal of Memory and Language, 28,* 462–485.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43–71.

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997) Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50*(A), 528–559.

Morrison, C. M., & Ellis, A. W. (1995) Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 116–133.

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992) Age of acquisition, not word frequency, affects object recognition. *Memory & Cognition, 50,* 705–714.

Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1998). A case for the lemma/lexeme distinction in models of speaking: Comment on Caramazza and Miozzo (1997). *Cognition, 69,* 219–230.

Savage, G. R., Bradley, D. C., Forster, K. I. (1990). Word frequency and the pronunciation task: The contribution of articulatory fluency. *Language & Cognitive Processes, 5,* 203–236.

Sebastian, N., Marti, M. A., Cuetos, F., & Carreiras, M. (1996). *LEXESP: base de datos informatizada de la lengua espanola* [LEXESP: Computerized database of Spanish language]. Unpublished database, Universitat de Barcelona.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning Memory, and Cognition, 21,* 686–698.

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.

Wingfield, A. (1967). Perceptual and response hierarchies in object identification. *Acta Psychologica, 26,* 216–226.

Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *American Journal of Psychology, 18,* 226–234.

Xiandai Hanyu Pinlv Cidian [Institute of Language Teaching and Research]. (1986). *Modern Chinese frequency dictionary.* Beijing, China: Beijing Language Institute Press.

(Appendixes follow)

Appendix A

Pictures Shown in Experiment 1A (Along With Their Italian Names)

Picture set		
HomName	Specific-word frequency	Cumulative-homophone frequency
Nun (*Suora*)	Owl (*Gufo*)	Tooth (*Dente*)
Tower (*Torre*)	Apple (*Mela*)	Barrel (*Botte*)
Bear (*Orso*)	Lion (*Leone*)	Bone (*Osso*)
Tie (*Cravatta*)	Monk (*Frate*)	Box (*Scatola*)
Screw (*Vite*)	Bread (*Pane*)	Bus (*Autobus*)
Sun (*Sole*)	Dog (*Cane*)	Car (*Macchina*)
Deer (*Cervo*)	Goat (*Capra*)	Chain (*Catena*)
Swing (*Altalena*)	Eagle (*Aquila*)	Chair (*Sedia*)
Ark (*Arca*)	Sphynx (*Sfinge*)	Egg (*Uovo*)
Fire (*Fuoco*)	Tree (*Albero*)	Foot (*Piede*)
Train (*Treno*)	Bird (*Uccello*)	Horse (*Cavallo*)
Whistle (*Fischietto*)	Pumpkin (*Zucca*)	Lemon (*Limone*)
Well (*Pozzo*)	Doll (*Bambola*)	Money (*Moneta*)
Dam (*Diga*)	Crab (*Granchio*)	Moon (*Luna*)
Cross (*Croce*)	Shirt (*Camicia*)	Radio (*Radio*)
Safe (*Cassaforte*)	Scarf (*Sciarpa*)	Shoe (*Scarpa*)
Pear (*Pera*)	Cheese (*Formaggio*)	Soldier (*Soldato*)
Watch (*Orologio*)	Piano (*Pianoforte*)	Table (*Tavolo*)
Anchor (*Ancora*)	Ladder (*Scala*)	Tractor (*Trattore*)
Sail (*Vela*)	Maze (*Labirinto*)	Wall (*Muro*)
Whale (*Balena*)	Frog (*Rana*)	Wrist (*Polso*)
Mane[a]	Skunk[a]	Bed[a]
Nurse[a]	Pill[a]	Corn[a]
Stamp[a]	Skull[a]	Cow[a]
Bow[a,b]	Sword[a,b]	Pie[a,b]
Crack[a]	Pig[a]	Roof[a]

[a] These pictures were excluded in the control experiment carried out in Italian (Experiment 1B). [b] These pictures were excluded from the analyses in Experiment 1A.

Appendix B

Pictures Shown in Experiment 2A (Along With Their English Names)

Picture set		
HomName	Specific-word frequency	Cumulative-homophone frequency
锹 (Shovel)	勺 (Spoon)	窗 (Window)
帆 (Sail)	梯 (Ladder)	脑 (Brain)
鼠 (Mouse)	琴 (Guitar)	枪 (Gun)
糖 (Candy)	鼻 (Nose)	雪 (Snow)
桃 (Peach)	刷 (Brush)	床 (Bed)
铃 (Bell)	瓜 (Melon)	灯 (Lamp)
斧 (Axe)	塔 (Tower)	画 (Painting)
壶 (Pot)	舌 (Tongue)	桥 (Bridge)
蒜 (Garlic)	蛇 (Snake)	嘴 (Mouth)
碗 (Bowl)	耳 (Ear)	鱼 (Fish)
龟 (Turtle)	钩 (Hook)	手 (Hand)
桶 (Bucket)	兔 (Rabbit)	星 (Star)
裙 (Dress)	袜 (Sock)	脚 (Foot)
剪 (Scissors)	伞 (Umbrella)	衣 (Coat)
叶 (Leaf)	虎 (Tiger)	树 (Tree)
豹 (Leopard)	链 (Chain)	钱 (Money)
鹰 (Eagle)	筐 (Basket)	船 (Boat)
鸭 (Duck)	猫 (Cat)	马 (Horse)
瓶 (Bottle)	牙 (Tooth)	火 (Fire)
狮 (Lion)	鹅 (Goose)	线 (Thread)
梳 (Comb)	熊 (Bear)	花 (Flower)
旗 (Flag)	云 (Cloud)	书 (Book)
岛 (Island)	狗 (Dog)	表 (Watch)
鞭 (Whip)	扇 (Fan)	月 (Moon)
锁 (Lock)	猴 (Monkey)	门 (Door)
鹿 (Deer)	锤 (Hammer)	山 (Mountain)
驴 (Donkey)	钉 (Nail)	车 (Car)
锯 (Saw)	毯 (Rug)	路 (Road)
鲸 (Whale)	虾 (Lobster)	眼 (Eye)
剑 (Sword)	锅 (Pan)	水 (Water)
箭 (Arrow)	尺 (Ruler)	心 (Heart)
豆[a] (Bean)	鹤[a] (Crane)	虫[a] (Worm)

[a] These pictures were excluded in the control experiment carried out in English (Experiment 2B).

(*Appendixes continue*)

Appendix C

Spanish–English Translations Shown in Experiment 3A

Word set					
HomName		Specific-word frequency		Cumulative-homophone frequency	
Spanish	English	Spanish	English	Spanish	English
Freno	Brake	*Bufanda*	Scarf	*Perdido*	Lost
Monja	Nun	*Buho*	Owl	*Nueve*	Nine
Bruja	Witch	*Pulgar*	Thumb	*Quién*	Who
Haya	Beech	*Levadura*	Yeast	*Rueda*	Wheel
Harina	Flour	*Payaso*	Clown	*Cuchillo*	Knife
Nudo	Knot	*Grua*	Crane	*Para*	For
Abeja	Bee	*Mandíbula*	Jaw	*Con*	With
Colilla	Butt	*Cabra*	Goat	*Ella*	She
Ciervo	Deer	*Cordero*	Lamb	*Alto*	Tall
Caballero	Knight	*Ladrillo*	Brick	*Agua*	Water
Débil	Weak	*Desnudo*	Naked	*Habitación*	Room
Carne	Meat	*Leche*	Milk	*Dios*	God
Madera	Wood	*Viento*	Wind	*Todo*	All
Tomillo	Thyme	*Albahaca*	Basil	*Gente*	People
Ronco	Hoarse	*Podrido*	Rotten	*Río*	River
Mejillón	Mussel	*Grifo*	Faucet	*Nacimiento*	Birth
Crin	Mane	*Cangrejo*	Crab	*Facil*	Easy
Llanura	Plain	*Aguja*	Needle	*Esquina*	Corner
Aqujero	Hole	*Onda*	Wave	*Blanco*	White
Liebre	Hare	*Ciruela*	Plum	*Arbol*	Tree
Criada[a]	Maid	*Ala*[a]	Wing	*Allí*[a]	There
Rocío[a]	Dew	*Babero*[a]	Bib	*Dos*[a]	Two

[a] These words were excluded from the analyses in Experiment 3A.

Received March 22, 2001
Revision received March 28, 2001
Accepted March 28, 2001 ■