Journal of Experimental Psychology:
Learning, Memory, and Cognition
2004, Vol. 30, No. 1, 278–282

0278-7393/04/$12.00 DOI: 10.1037/0278-7393.30.1.278

REPLY TO EARLIER COMMENTARY

What Determines the Speed of Lexical Access: Homophone or Specific-Word Frequency? A Reply to Jescheniak et al. (2003)

Alfonso Caramazza and Yanchao Bi
Harvard University

Albert Costa
University of Barcelona

Michele Miozzo
Columbia University

A. Caramazza, A. Costa, M. Miozzo, and Y. Bi (2001) reported a series of experiments showing that naming latencies for homophones are determined by specific-word frequency (e.g., frequency of *nun*) and not homophone frequency (frequency of *nun* + *none*). J. D. Jescheniak, A. S. Meyer, and W. J. M. Levelt (2003) have challenged these studies on a variety of grounds. Here we argue that these criticisms are not well founded and try to clarify the theoretical issues that can be meaningfully addressed by considering the effects of frequency on homophone production. We conclude that the evidence from homophone production cannot be considered to provide support to 2-layer theories of the lexical system.

Levelt, Roelofs, and Meyer (1999) argued that data on the production of homophones can be used to adjudicate between models of lexical access that assume that two lexical layers—lemma and lexeme—mediate between the semantic and phonological contents of words (e.g., Dell, 1990; Levelt et al., 1999) and models that assume only one lexical layer (e.g., Caramazza, 1997). There are two ways in which homophones (e.g., *nun* and *none*) might be represented in the lexical system: They could either have independent lexical representations or share a lexical representation. If homophones share a representation, then there must exist another level of lexical representation in which they can be distinguished (so as to capture their distinct grammatical properties). If they do not share a representation, then the lexical system can have either one or two layers of lexical nodes.

These possibilities are represented schematically in Figure 1, where it is apparent that although the shared representation (SR) hypothesis is incompatible with one-layer theories of the lexicon, the independent representation (IR) hypothesis is compatible with both one- and two-layer theories of the lexicon. Thus, experimental evidence in support of the SR hypothesis of homophones would, ipso facto, provide support for two-layer theories of the lexical system. However, experimental evidence in support of the IR hypothesis would not distinguish between one- and two-layer models of lexical access (see Caramazza, Costa, Miozzo, & Bi, 2001, for extensive discussion).

Levelt et al. (1999) cited the results of two experiments, which investigated the effects of word frequency on the production of homophones (Dell, 1990; Jescheniak & Levelt, 1994), as providing evidence in support of two-layer theories of the lexicon. The logic of the studies is straightforward. If homophones share a lexical representation, the frequency of the shared representation would be the cumulative frequency of the homophone members (*homophone frequency:* e.g., the frequency of *nun* plus *none*) and not the frequency of the word itself (*specific-word frequency:* e.g., the frequency of *nun*). Furthermore, if frequency affects access to the shared lexical representation (the lexeme layer), then homophone frequency, and not specific-word frequency, would explain performance in naming words such as *nun.*

Data apparently consistent with this prediction were obtained by Dell (1990). He used an error-inducing paradigm in which participants were required to produce as quickly as possible phrases such as *him/hymn to sing.* A post hoc analysis revealed that homophone frequency was a better predictor of error rates than was specific-word frequency. However, it is not clear that this paradigm informs theories of lexical selection; rather, it may reveal the operation of postlexical phonological processes. The task requires participants to prepare a phrase, which is to be produced upon the presentation of a cue. In such a task, one can assume that participants access the lexical items they intend to produce and store them in a working memory system in preparation for production. The errors that occur in this task are likely to reflect properties of the phonological representations stored in the working memory system, including the strength of association among phonemes and phoneme sequences in a word, rather than the process of lexical access.

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

Correspondence concerning this article should be addressed to Alfonso Caramazza, Cognitive Neuropsychology Laboratory, Harvard University, 33 Kirkland Street, Cambridge, MA 02138. E-mail: caram@wjh.harvard.edu

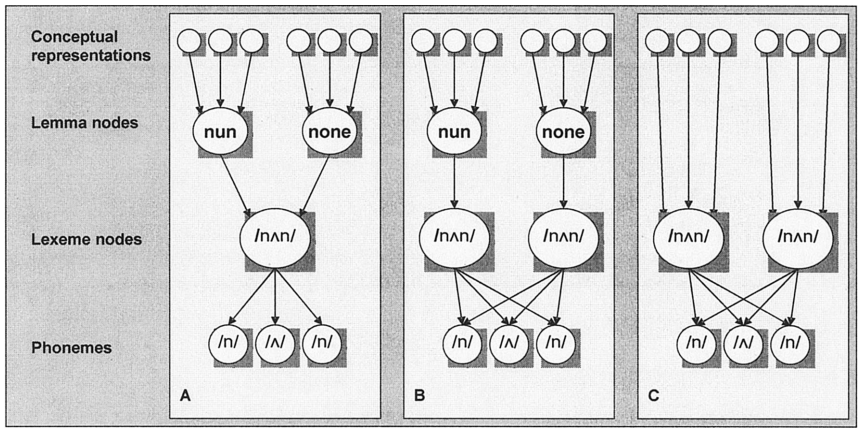

Figure 1. Schematic representation of the shared (SR) and independent representation (IR) hypotheses for two- and one-layer models of lexical access. A: The SR hypothesis within a two-lexical-layer model. B: The IR hypothesis within a two-lexical-layer model. C: The IR hypothesis within a one-lexical-layer model.

Jescheniak and Levelt (1994) used a word translation task to investigate homophone production. Dutch–English bilinguals were asked to translate English words into Dutch as fast as possible. Participants named homophones as fast as control words matched on homophone frequency and faster than control words matched on specific-word frequency. Jescheniak and Levelt interpreted this pattern of results as consistent with a model of lexical access that makes the following assumptions: (a) Homophones share a representation, and (b) the frequency effect in speech production is located at the level of the shared representation—the lexeme level in their model.

A different pattern of results was reported by Caramazza, Costa, et al. (2001) in three experiments that were carried out in different languages with different experimental paradigms. In all cases, naming latencies were determined by specific-word frequency and not by homophone frequency. In two experiments (one with English speakers and one with Mandarin speakers) they used a simple picture-naming task with homophones (e.g., *nun*) and matched controls for homophone frequency (e.g., *tooth*) and specific-word frequency (e.g., *owl*). Homophones were named as fast as specific-word frequency controls and significantly more slowly than homophone frequency controls. Caramazza, Costa, et al. also carried out a translation experiment similar to the one reported by Jescheniak and Levelt (1994), except that they used Spanish–English bilinguals and a larger set of stimuli. As in the case of the simple naming experiments, they found that specific-word and not homophone frequency determined production latencies. They concluded that the putative homophone frequency effect is not a reliable phenomenon and therefore cannot be used to support the SR hypothesis or other two-layer theories of the lexical system.

What Are the Theoretical Issues at Stake?

Jescheniak, Meyer, and Levelt (2003) challenged the conclusions reached by Caramazza, Costa, et al. (2001), and presented new data that supposedly replicate the homophone frequency effect and therefore favor the SR hypothesis. Before considering the specific criticisms that were raised by Jescheniak et al., it is important to clarify the theoretical claims that are at stake in this exchange. Jescheniak et al. attributed to Caramazza, Costa, et al. the following two claims: (a) that Caramazza, Costa, et al.'s results undermine the SR hypothesis of homophones and (b) that this argues against two-layer models of the lexical system.

Jescheniak et al. (2003) were right in attributing to us the claim that our results appear to favor the IR hypothesis of homophones, but they were mistaken in attributing to us the claim that our results argue against two-layer models of the lexicon. Casting Caramazza, Costa, et al.'s (2001) position in the terms chosen by Jescheniak et al. conflates two distinct claims, only one of which corresponds to the position taken by Caramazza, Costa, et al. The distinction is between the strong claim that a set of results is inconsistent with a theory and the weaker claim that those results do not support the theory in question but do not disconfirm it either. Caramazza, Costa, et al. argued that their data on homophone production do not support the lemma–lexeme distinction. They never claimed that those data argue against two-stage models of lexical access. In fact, Caramazza, Costa, et al. discussed two ways in which a two-layer model of the lexicon could accommodate their results (but there may be others). One possibility is to abandon the assumption that homophones share a representation while retaining the assumption of a lemma–lexeme distinction (see Figure 1B and discussion introduction). Another possibility is that the frequency effect has its locus not at the lexeme but at the lemma level (e.g., Dell, 1990).

The distinction drawn here does not aim simply to correct a mischaracterization of our position but to set the stage for a proper assessment of the implications of the available empirical facts on homophone production and, especially of the new results reported by Jescheniak et al. (2003).

The Criticisms Raised by Jescheniak et al. (2003)

The first part of Jescheniak et al.'s (2003) criticism concerns the cross-language comparison we adopted to rule out the possibility that object recognition—rather than frequency—was responsible for the results obtained with the homophones. For this purpose, we compared how speakers of different languages (Italian for the English experiment, and English for the Mandarin experiment)

named the same sets of pictures. The logic of this control is that because the experimental items are homophones only in the experimental language, we would expect to find different patterns of naming latencies for experimental and control stimuli in the two languages only if there were a homophone frequency effect.

Jescheniak et al. (2003) objected that cultural differences between the participants in the experimental and control languages may have contributed to the obtained patterns of results. This is an intriguing claim, for one would have to suppose that such putative cultural differences between English and Italian (for Experiments 1A and 1B) and Mandarin and English (for Experiments 2A and 2B) were such that they would perfectly align with the experimental and control tasks so as to produce the observed specific-word frequency effects in the two language contrasts in the two experiments. For example, one would have to argue that the items chosen for the English experiment (Experiment 1A) were such that the homophone stimuli were in fact much harder to recognize than were the specific-word and homophone frequency controls, thereby masking any advantage that would have accrued from their homophone status. One would also have to argue that such differences in difficulty between recognizing the homophone and control stimuli were not present for Italian speakers because of presumed (and unspecified) cultural differences between our American and Italian participants. Such a patterning of cultural differences would also presumably have occurred for the Chinese–English contrast (Experiments 2A and 2B). Although we cannot exclude such possibilities on logical grounds, they do not appear to have prima facie plausibility.

Jescheniak et al. (2003) also noted that the contrast in frequencies used in Caramazza, Costa et al.'s (2001) Experiment 1 was smaller than that used in Jescheniak and Levelt's (1994) experiment. This is correct, but its significance is unclear. What matters is whether the frequency differences were large enough to have a measurable effect on performance. On this score, the answer is clear: The homophone words and specific-word (low) frequency controls were named significantly more slowly than were homophone (high) frequency controls. Jescheniak et al. also noted that there was a 10-ms difference in the naming time for high- and low-frequency words in one of the delayed naming control experiments used to rule out the possibility that any differences in naming latencies for the experimental and control stimuli were due to differences in the ease with which the stimuli could be articulated. However, as reported by Caramazza, Costa, et al. (p. 1437, 2001), this difference was far from significant ($p < .20$).

The New Data Reported by Jescheniak et al. (2003)

In their effort to bolster their claim that homophone frequency determines naming latencies, Jescheniak et al. (2003) reported two new experiments on homophone production. Experiment 1 was an exact replication (with new participants) of the translation experiment reported in Jescheniak and Levelt's (1994) article, and they again found that homophones are produced as fast as homophone controls and more slowly than specific-word frequency controls. Although these findings confirm the robustness of Jescheniak and Levelt's results, they do not help clarify the reason for the contrasting results.

More useful is Jescheniak et al.'s (2003) Experiment 2, which also used a word translation task, but this time with German–English bilinguals. The results of this experiment pattern neither with the other results reported by Jescheniak and Levelt (1994) nor with those reported by Caramazza, Costa, et al. (2001). Instead, the homophone word condition differed from both the specific-word and homophone frequency controls (see Jescheniak et al., 2003, Figure 2). This result is inconsistent with both the SR and the IR hypotheses of homophones, because the former hypothesis predicts that naming latencies for homophones should be the same as those for homophone frequency controls, whereas the latter hypothesis predicts that naming latencies should be the same for homophones and specific-word frequency controls. However, Jescheniak et al. (2003) reached a radically different conclusion on the basis of their results:

> Overall, the data pattern is similar to the pattern obtained by Jescheniak and Levelt (1994, Experiment 6) and in Experiment 1 of the present article, *providing a cross-linguistic replication of the basic finding* [italics added]. In all cases the (low frequency) homophones were named significantly faster than the LF controls, and this difference did not dissipate over repetitions. . . . In both experiments low-frequency target words with a high-frequency homophone twin were produced substantially faster than low-frequency control words without a high-frequency twin. *This pattern strongly suggests that homophonic twins share a representation in the lexical system, most likely the word form representation* [italics added]. (p. 436)

This is a surprising conclusion when considered in light of the fact that Jescheniak et al. (2003) found that homophones were named not only faster than specific-word controls but also slower than homophone frequency controls. The latter result is inconsistent with the shared representation hypothesis. Thus, Jescheniak et al.'s conclusion follows only if one considers half of their data—the half showing that homophones were named faster than specific-word frequency controls. If one considers the other half of their data—the half showing that homophones are named more slowly than homophone frequency controls—one would have to conclude that homophones do not share a representation in the lexical system. The problem here is that Jescheniak et al. seem to have cast the issue under consideration exclusively from the vantage point of the IR hypothesis while ignoring the predictions made by the SR hypothesis. When the results are considered from the perspective of the predictions made by both theories, it is clear that they are inconsistent with both sets of predictions.

Where Does All This Lead?

The use of the translation task has not led to a consistent pattern of results for homophone production: Some results are consistent with the IR hypothesis (data from the Spanish–English bilinguals), some are consistent with the SR hypothesis (data from the Dutch–English bilinguals), and some are not consistent with either hypothesis (data from the German–English bilinguals). Thus, no firm conclusion is possible from these data. There are, however, the data from picture naming, which show that specific-word, and not homophone, frequency predicts naming latencies for homophones in English and Mandarin.

This conclusion is reinforced by the results of experiments reported by Bonin and Fayol (2002). These authors compared French speakers' production latencies, both in written and spoken naming, for the low- (e.g., *ver*, *worm*) and the high- (e.g., *verre*,

glass) frequency member of heterographic homophones (/ver/). Both in the oral and in the written production tasks, participants named the higher frequency member of a homophone pair (*verre, glass*) much faster than they named the lower frequency member (*ver, worm*). This effect cannot be attributed to differences in the recognition of the pictures used for the high- and the low-frequency members of a homophone pair because, if anything, the high-frequency homophones were harder to recognize than were the low-frequency homophones. Thus, the results reported by Bonin and Fayol (2002) are inconsistent with the SR hypothesis of homophones proposed by Jescheniak and Levelt (1994), which predicts that the only determinant of lexical access of homophones is the frequency of the shared lexeme and, therefore, that naming latencies for the higher and lower specific-word frequency members should be the same. Jescheniak et al.'s (2003) objection that Bonin and Fayol's experiments do not allow an evaluation of the IR hypothesis is correct but does not change the fact that the results do not provide support for a homophone frequency effect and therefore are inconsistent with Jescheniak and Levelt's SR hypothesis of homophones.[1]

Jescheniak et al. (2003) cite the results of Schriefers, Jescheniak, and Hantsch's (2002) and Janssen and Caramazza's (2003) studies on determiner production as potentially relevant to the issue of homophone representation (see also Alario & Caramazza, 2002; Schiller & Caramazza, 2003). However, the implications of these data for theories of homophone representation of open class words are far from obvious. Consideration of this issue would require discussion of the complex problem of determiner production, which involves processes that are distinct from those implicated in the lexical access of open class words (for discussion see Caramazza, Miozzo, Costa, Schiller, & Alario, 2001). Whether grammatical morphemes are represented in the same manner as open class words is an unresolved empirical issue. Space limitations preclude further consideration of this issue.

Finally, there is a remarkable clinical phenomenon—modality-specific grammatical category deficits—that poses serious difficulties for the SR hypothesis. Brain damage can selectively impair the production of either the noun or the verb member of (homographic) homophone pairs (such as *the watch–to watch*) in only one modality of output—either only speaking or only writing (e.g., Caramazza & Hillis, 1991; Rapp & Caramazza, 2002; for review see Caramazza & Shapiro, in press). Thus, for example, a patient might be able to write *watch* correctly both as a noun and as a verb, but in the oral modality he or she might only be able to produce the noun form of the pair "the watch–to watch" correctly, while producing semantic errors in naming verbs.

Within a theory such as that proposed by Levelt et al. (1999), which seems to assume a common lemma representation for phonological and orthographic processing (see also Roelofs, Meyer, & Levelt, 1996), such a pattern of performance rules out the lemma level as the locus of the oral production deficit for verb homophones. This is because the lemma level must be intact to support good performance in the spared modality. Furthermore, because the patient makes semantic errors in the impaired grammatical category, one can infer that the deficit involves a semantically-based form of lexical access, thereby excluding a dissociation deficit between lemma and lexeme levels in this type of architecture. Finally, because the patient is able to produce the phonological form of the homophone pair correctly, this rules out a peripheral phonological processing deficit, for otherwise the patient would have shown impaired performance for both the noun and verb forms of the homophone. By using this process of elimination one can locate the source of the patient's difficulty in accessing the lexical phonological form of the verb. But this implies that the noun form of the homophone pair must be represented independently of the verb form, for otherwise access to it should also have been impaired. Thus, the selective deficit in accessing one member of homophone pairs is prima facie evidence for the IR hypothesis of homophones.

Conclusion

Various sorts of data have been marshaled in support of the view that there are two lexical layers that mediate between the semantic and phonological contents of words. One such source of data is the putative existence of a homophone frequency effect in lexical access. The existence of this phenomenon has been called into question by studies that have investigated homophone naming latencies (Bonin & Fayol, 2002; Caramazza, Costa, et al., 2001). However, Jescheniak et al. (2003) challenged these studies on a variety of grounds. Here we have argued that these criticisms are not well founded. When all of the results concerning homophone production are considered together, including the evidence from neuropsychology, the preponderance of evidence is not consistent with the SR hypothesis of homophones proposed by Jescheniak and Levelt (1994). As a consequence, the evidence from homophone production cannot be considered to provide support for two-layer theories of the lexical system.

[1] One should note that the stimuli used by Bonin and Fayol (2002) were heterographic homophones (*ver/verre*), whereas the stimuli used by Jescheniak and Levelt (1994) were homographic homophones. However, this difference is irrelevant in the theoretical context under discussion here: The SR hypothesis proposed by Levelt et al. (1999) concerns claims about phonological lexemes only.

References

Alario, F. X., & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition, 82,* 179–223.

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology, 14,* 289–314.

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

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1430–1450.

Caramazza, A., & Hillis, A. (1991). Lexical organization of nouns and verbs in the brain. *Nature, 349,* 788–790.

Caramazza, A., Miozzo, M., Costa, A., Schiller, N., & Alario, F.-X. (2001). A cross-linguistic investigation of determiner production. In E. Dupoux (Ed.), *Language, brain and cognitive development: Essays in honor of Jacques Mehler* (pp. 208–226). Cambridge, MA: MIT Press.

Caramazza, A., & Shapiro, K. (in press). The representation of grammatical knowledge in the brain. In L. Jenkins (Ed.), *Variation and universals in biolinguistics.* Amsterdam: Elsevier.

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

Janssen, N., & Caramazza, A. (2003). The selection of closed-class words in noun phrase production: The case of Dutch determiners. *Journal of Memory and Language, 48,* 635–652.

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

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

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

Rapp, A., & Caramazza, A. (2002). Selective difficulties with spoken nouns and written verbs: A single case study. *Journal of Neurolinguistics, 15,* 373–402.

Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1996). Interaction between semantic and orthographic factors in conceptually driven naming: Comment on Starreveld and LaHeij (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 246–251.

Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language, 48,* 169–194.

Schriefers, H., Jescheniak, J. D., & Hantsch, A. (2002). Determiner selection in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 941–950.

Received April 11, 2003
Revision received June 5, 2003
Accepted June 5, 2003 ■