The locus of the frequency effect in picture naming: When recognizing is not enough

JORGE ALMEIDA AND MARK KNOBEL
Harvard University, Cambridge, Massachusetts

MATTHEW FINKBEINER
Macquarie Centre for Cognitive Science, Sydney, New South Wales, Australia

AND

ALFONSO CARAMAZZA
Harvard University, Cambridge, Massachusetts
and University of Trento, Rovereto, Italy

The lexical frequency effect in picture naming is generally assumed to constitute a signature of lexical access. Lexical frequency, however, is correlated with other variables, like concept familiarity, that can produce effects similar to those of lexical frequency in picture naming tasks. In this study, a delayed picture naming task was employed to address the hypothesis that the frequency effect in picture naming is due to variables that affect processing in the perceptual and semantic identification stages (i.e., input stages). Despite the fact that all the input processing stages were completed prior to the presentation of the naming cue, a strong frequency effect was still obtained in this task. These results establish that the lexical frequency effect is independent of variables affecting the input stages of picture naming, and, hence, confirm the lexical frequency effect as a signature effect of lexical access.

It has been demonstrated that pictures whose names occur more frequently (e.g., *dog*) are named faster than pictures whose names occur less frequently (e.g., *deer*; Oldfield & Wingfield, 1965). The received view among cognitive psychologists is that these effects arise as a result of differences in the representation and/or processing of high and low frequency lexical representations (e.g., Caramazza, 1997; Levelt, Roelofs, & Meyer, 1999; Roelofs, 1997).[1] Lexical frequency, however, is strongly correlated with variables like familiarity, and visual complexity (Snodgrass & Vanderwart, 1980), and interacts with variables like structural similarity in picture naming experiments (Humphreys, Riddoch, & Quinlan, 1988). Importantly, each of these variables affects picture naming latencies, and so, it could be hypothesized that the so-called effects of lexical frequency in naming tasks may actually arise in recognition processes (Johnson, Paivio, & Clark, 1996). We will refer to this possibility as the *input hypothesis*. On this view, the provenance of the apparent lexical frequency effect is the perceptual identification stages of picture naming.

Consistent with the input hypothesis, Bates et al. (2003) proposed that much of the frequency effect obtained in picture naming is, in reality, an effect of conceptual accessi-

bility (i.e., concept familiarity). In a large study of picture naming performance across seven languages, Bates et al. found that a frequency measure, based on the other six languages' frequency ratings, explained unique variance in both name agreement and reaction times in all seven languages, even when the target language's frequency was included as a factor. Furthermore, in four of the seven languages, the target language frequency did not contribute significantly. They argued that cross-language frequency measures the general familiarity or conceptual availability of depicted objects, and, thus, the so-called lexical frequency effect actually reflects conceptual processes. If this view is correct, it would have major implications for how we interpret a long line of research wherein the lexical frequency effect has been used to constrain theories of lexical access in speech production (e.g., Alario, Costa, & Caramazza, 2002; Bonin & Fayol, 2002; Griffin, & Bock, 1998; Jescheniak, Meyer, & Levelt, 2003).

In addition to the results of the Bates et al. (2003) study, the input hypothesis has received support from two studies that have documented so-called lexical frequency effects in picture recognition and picture matching tasks. Kroll and Potter (1984) were interested in whether picture- and word-recognition processes accessed common conceptual

J. Almeida, jalmeida@wjh.harvard.edu

representations. To address this issue, they devised a series of experiments that contrasted object decisions with lexical decisions. In their object decision task, participants were presented with pictures and were instructed to indicate whether they depicted real objects or nonobjects. Post hoc analyses revealed an apparent frequency effect (i.e., real objects with high frequency (HF) names were identified faster than real objects with low frequency (LF) names). Similarly, Bartram (1976, Experiment 2) reported a frequency effect in a picture-to-picture matching task, wherein participants were presented with two pictures in succession and quickly indicated whether these pictures represented the same basic level concept. Picture matching was faster for pictures with HF names, relative to those with LF names. These findings suggest that effects which mimic the lexical frequency effect may be obtained in tasks that do not necessarily require lexical access.

Needless to say, support for the input hypothesis is highly problematic for theories that construe the frequency effect as an index of lexical access processes. In light of the support that has been marshaled in favor of the input hypothesis, it becomes vital to conduct a stringent test of this hypothesis. Previous researchers have examined the input hypothesis by designing tasks that tap into the input stages of picture naming processes but not the later stages of lexical access. Presumably, if frequency effects occur in these "input" tasks, the input hypothesis would be supported. Conversely, if no effects of lexical frequency are obtained, the input hypothesis can be dismissed. In one example, Jescheniak and Levelt (1994, Experiment 2) asked participants to indicate whether previously presented words denoted the objects depicted in target pictures. Jescheniak and Levelt did not find a reaction time advantage for objects with high-frequency names over objects with low frequency names. Based on this null result, these authors concluded that lexical frequency effects cannot be attributed to variables that affect object recognition. Although such null results are suggestive, they do not provide sufficient evidence to reject the input hypothesis. The null result may reflect a true lack of the hypothesized effect, or it may merely reflect insufficient experimental or statistical power. Given these possibilities, it is important to devise a task that can assess the input hypothesis without interpretation of null effects.

In the research reported here, we test the input hypothesis of the frequency effect, using an experimental paradigm (delayed picture naming) that isolates the lexical access stages by allowing for picture recognition processes to be fully completed prior to naming. We reasoned that interpolating a sufficiently long delay between the presentation of the to-be-named picture and the response cue would allow for the completion of picture recognition processes. Hence, if an effect of lexical frequency is obtained in this paradigm, it would be very difficult to attribute that effect to picture recognition processes. Note, however, that previous experiments using delayed picture naming have failed to reveal effects of lexical frequency (e.g., Barry, Hirsch, Johnston, & Williams, 2001). Previous delayed naming tasks have allowed participants to prepare their response all the way up to the point of articulation, meaning that

both recognition processes and lexical access processes were completed by the time the response cue was introduced. In the present experiment, we sought to prevent the preparation of naming responses by introducing response uncertainty; before the response cue, participants were unsure whether they would have to name the picture or categorize it. Furthermore, the prevalence of both tasks was unbalanced, so that subjects had to categorize pictures 75% of the time and name pictures only 25% of the time. Because participants primarily categorized the pictures, we reasoned that they would be more likely to prepare category names than picture names. Thus, although participants were given sufficient time to complete all recognition processes, we predicted that they would not select the lexical nodes corresponding to picture names.

Under these conditions, if the locus of the frequency effect is at the pre-lexical stages of picture recognition, no frequency effect should be obtained. If, however, lexical frequency effects arise in stages beyond early picture recognition, a frequency effect should emerge. Furthermore, the presence of a frequency effect under these delayed naming conditions would allow us to confidently state that frequency effects arising from input processing are neither sufficient nor necessary to obtain a lexical frequency effect. Two other experimental conditions were included: one without delay, where picture recognition, lexical access, and word selection are at play and, therefore, a frequency effect is expected under all hypotheses; and one without response uncertainty, where only post selection processes can be responsible for whatever effects obtained. If frequency is, in fact, a signature of lexical access, no frequency effects are expected in this experimental condition. These two experimental conditions were included to ensure that our results are not due to extraneous factors.

METHOD

Participants

Forty-eight native English speakers (16 per experimental condition) with normal or corrected-to-normal vision participated in the experiments. All the participants were Harvard University students who provided informed consent for this study. Participants were either paid or given course credit compensation for their participation.

Materials

One hundred and sixty pictures of common objects were selected evenly from two distinct semantic categories (animals and artifacts). All pictures were selected from the IPNP database (Szekely et al., 2004) or were created to look similar to the pictures from this database. From each category, twenty pictures were selected to be used as targets, while the remaining 60 pictures were selected to be used as fillers (i.e., pictures for use in categorization trials). The target pictures were selected so that for each category, 10 had low frequency names (LF; range 1–12 per million) and 10 had high frequency names (HF; range 23–235 per million). The two groups of high- and low-frequency target pictures were matched on concreteness, imageability, number of syllables, length in phonemes, phonological neighborhood size, average frequency of phonological neighbors, average positional segment frequency, and average position-specific biphone frequency (see Appendixes 1 and 2 for the materials and matching variables). In a pilot experiment ($N = 16$) using immediate naming with these pictures, a robust frequency effect was observed [$M_{HF} = 654$ msec, $SD_{HF} = 81$ msec; $M_{LF} = 733$ msec, $SD_{LF} = 76$ msec; $t_1(15) = -8.43$, $p < .001$; $t_2(38) = -3.27$, $p = .002$]. A

further pilot study was performed with the same pictures to ensure that there were no articulatory differences in our materials. Participants ($N = 16$) were asked to name the target pictures after a fixed delay. No frequency effect was obtained in this experiment [$M_{HF} = 382$ msec, $SD_{HF} = 82$ msec; $M_{LF} = 400$ msec, $SD_{LF} = 83$ msec; $t = -1.145$, $p = .27$], replicating Barry et al. (2001). Finally, 20 pictures (10 animals and 10 artifacts) were selected for the practice session, following the same constraints used for the experimental pictures.

Procedure and Design

Three filled circles were created: one gray, one blue, and one yellow. The blue and the yellow circles were used as response cues for the categorization and naming procedures, respectively, while the gray circle was used during the delay. The circles were large enough to contain the target stimulus, a black and white line drawing on a white square background.

Forty blocks of filler items were created with the number of fillers in each block being pseudorandomly selected, so that each block had no fewer than two and no more than seven fillers. Target items were pseudorandomly assigned to the end of each block of fillers. The resulting blocks were also pseudorandomly ordered. The experimental order of items met the following constraints: (1) No more than three items from the same semantic category were presented in a row, and (2) no picture name would share an onset with the preceding picture name.

Participants were tested individually in a sound-attenuating and dimly lit room. There were three phases in the experiment. In the first phase, participants were familiarized with the pictures and their names. On each trial a picture was presented centrally along with its name, and participants were required to read the picture name. The second phase consisted of 20 trials wherein the experimental task was practiced. The third phase was the experiment proper. The trial structure of the second and third phases was identical, and depended on the condition studied. Stimulus presentation was controlled by DMDX software (Forster & Forster, 2003).

Three conditions were created by manipulating response uncertainty and response delay as follows: (1) delayed naming with uncertainty in response (*delayed/uncertain*); (2) delayed naming with certainty in response (*delayed/certain*); and (3) immediate naming with uncertainty in response (*immediate/uncertain*). In the delayed/uncertain condition—our critical condition—participants were presented with the following sequence of events in the second and third phases of the experiment: In each trial, a fixation cross appeared for 500 msec and was replaced by a picture, superimposed on a gray circle. After the picture had been shown for 1,000 msec, the circle changed its color from gray to either blue or yellow. If the surrounding circle was blue, participants named the picture's category; if the surrounding circle was yellow, they named the picture itself. They were instructed to respond only after the circle changed color, and then as quickly as possible; response times were recorded. Each picture remained on the screen for 3,000 msec or until the participant's response. Trials were separated by an empty screen, displayed for 1,000 msec. Pictures were only presented once.

The trial structure of the two other conditions was derived from the delayed/uncertain condition. In the delayed/certain condition, participants withheld their response until a cue was given, as in the delayed/uncertainty condition, but were informed of the required response (naming or categorization) simultaneously with picture onset. In other words, when the picture was presented for the first time, it was already surrounded with the colored circle that signaled the response; the response cue was a change from the critical colored circle to the gray circle. In the immediate/uncertain condition, we wanted to assess the magnitude of frequency effects when the delays between picture onsets and response cues were eliminated. In this condition, participants were instructed to name or categorize the pictures as soon as they were presented. The pictures were circumscribed by the critical colored circle corresponding to whether they were to be named or categorized immediately at onset.

After the experiment proper was over, participants were thanked and debriefed.

Analyses

Recorded responses were scored as errors and discarded if (1) they differed from the target response, (2) they contained verbal disfluencies, (3) there were any problems in triggering the voice key, or (4) target responses were below 100 msec. Responses that were three standard deviations away from a subject's mean response time were counted as outliers, and were replaced with values equal to cutoffs established 3 SD units above and below the mean for each participant. We report F_1 and t_1 (based on subject means) and F_2 and t_2 (based on item means) analyses.

RESULTS AND DISCUSSION

A total of 162 errors (8.43%) were produced by participants in response to the critical items. Additional 25 (1.3%) responses were considered outliers. A 2 (frequency—HF vs. LF) $\times$ 3 (experimental condition—delayed/uncertain, delayed/certain, and immediate/uncertain) ANOVA with frequency as a repeated factor for participants, and experimental condition as a repeated factor for items, was performed. This analysis yielded a main effect of frequency [$M_{HF} = 673$ msec, $SD_{HF} = 262$ msec; $M_{LF} = 724$ msec, $SD_{LF} = 285$ msec; $F_1(1,45) = 28.6$, $MS_e = 2,174.6$, $p < .001$; $F_2(1,38) = 8.745$, $MS_e = 8,534$, $p = .005$] and a main effect of experimental procedure [$M_{DU} = 682$ msec, $SD_{DU} = 118$ msec; $M_{DC} = 414$ msec, $SD_{DC} = 91$ msec; $M_{IU} = 999$ msec, $SD_{IU} = 172$ msec; $F_1(2,45) = 86.6$, $MS_e = 31,741.2$, $p < .001$; $F_2(2,76) = 921.9$, $MS_e = 4,215.4$, $p < .001$]. Crucially, this analysis revealed a significant interaction effect between frequency and experimental condition [$F_1(2,45) = 3.6$, $MS_e = 2,174.6$, $p = .037$; $F_2(2,76) = 2.65$, $MS_e = 4,215.4$, $p = .084$].

Post hoc t tests were performed to assess the frequency effects in each experimental condition. As shown by the leftmost bars of Figure 1, in our critical condition—that is, delayed/uncertain condition—HF targets ($M = 651$ msec, $SD = 119$ msec) were named 61 msec faster than LF targets [$M = 712$ msec, $SD = 112$ msec; $t_1(15) = -4.65$, $p = .0003$; $t_2(38) = -2.34$, $p = .025$], suggesting that the manipulation of lexical frequency affected stages of processing beyond picture recognition, which was completed during the delay period. Participants also made more errors on LF ($M = 2.56$, $SD = 1.63$) than HF [$M = 1.125$, $SD = 1.26$; $t_1(15) = -2.926$, $p = .01$; $t_2(38) = -3.147$, $p = .003$] targets. The frequency effect observed in this experimental condition is, however, consistent with several interpretations. For instance, frequency effects may arise in postlexical (articulatory) processes (Vitevitch, Armbrüster, & Chu, 2004) or as a function of the different workloads associated with keeping high and low frequency items in a memory buffer (Goldinger, Azuma, Abramson, & Jain, 1997). Moreover, the frequency effect in the delayed/uncertain condition may reflect unique demands of the dual-task paradigm that we employed.

The two other experimental conditions were performed to address these issues. In the delayed/certain condition, participants withheld responding until a cue was given, but were informed of the required response (naming or categorization) simultaneously with picture onset. Thus, picture identification, response selection, and response preparation could all be completed prior to the response
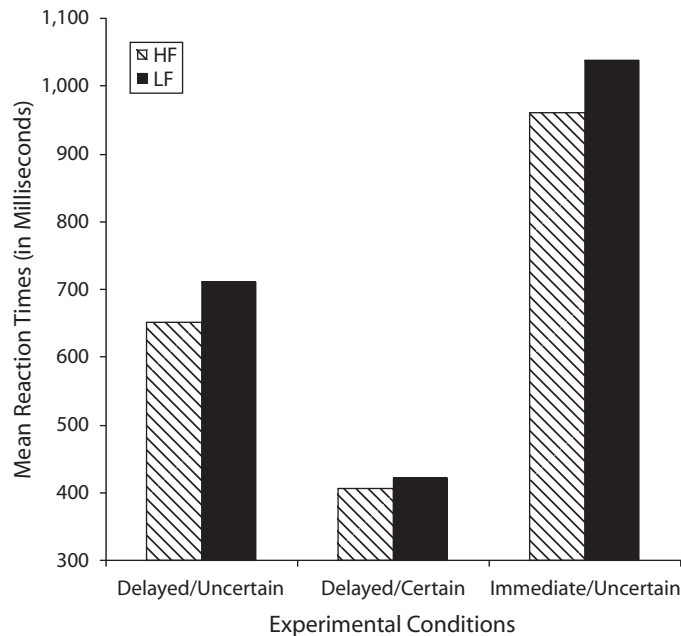
Figure 1. Mean reaction times for each frequency category in each experimental condition.

cue. As expected, no frequency effects were obtained in this task [M_{HF} = 406 msec, SD_{HF} = 90 msec; M_{LF} = 421 msec, SD_{LF} = 94 msec; $t_1(15)$ = −1.01, p = .327; $t_2(38) < 1$]. In the immediate/uncertain condition, participants had to name or categorize the picture as soon as they were presented on the screen. Hence, all potential effects were allowed in response time: Participants had to identify pictures, select the appropriate response type, and, in critical trials, select lexical entries for naming. As shown by the rightmost bars of Figure 1, HF targets (M = 961 msec, SD = 166 msec) were named 76 msec faster than were LF targets [M = 1,037 msec, SD = 175 msec; $t_1(15)$ = −3.84, p = .0016; $t_2(38)$ = −2.84, p = .0073].

Taken together, the experimental conditions reported here suggest that the lexical effect in picture naming cannot be wholly attributed to the recognition or to the articulatory stages. Moreover, our results do not appear to reflect memory load differences associated with keeping high- and low-frequency items in memory during the delay period (cf. Goldinger et al., 1997). Finally, the demands of our dual-task cannot account for the obtained frequency effects. By process of elimination, then, we suggest that the frequency effect serves as an index of lexical access processes.

GENERAL DISCUSSION

The research reported here establishes one important finding: Reaction time differences in naming performance between pictures with low- and high-frequency names obtained in picture naming experiments are driven by lexical processes, independently of the picture recognition and articulatory processes that take part in picture naming. A delayed picture naming task and a concurrent delayed

picture categorization task were employed to test this hypothesis. Despite the delay between picture presentation and naming cue, a strong frequency effect was obtained, such that pictures with HF names were named faster than pictures with LF names. The time allotted (1,000 msec) between picture presentation and response cue was sufficient for recognition processes to complete. There is no reason to assume that participants did not fully process the picture presented, given that both tasks (naming and categorization) required picture recognition. This demonstrates that variables that affect the processing stages completed during the delay period—at minimum, the picture recognition and semantic processing stages—do not exhaust the possible sources of the frequency effects observed in picture naming experiments. Please note that we are not claiming that input stages are completely insensitive to frequency manipulations. In fact, the numerical difference in frequency effects observed in the immediate/uncertain and delayed/uncertain experimental conditions may indicate that prelexical stages of processing are frequency sensitive (Bates et al., 2003). Our results suggest, however, that frequency effects reliably index lexical processes beyond such input processes.

The results reported converge with those obtained in previous experiments designed to test whether recognition stages are the primary source of frequency effects (e.g., Jescheniak & Levelt, 1994; Morrison, Ellis, & Quinlan, 1992; Wingfield, 1967). It is important to note, however, that our conclusions do not rest upon a null effect. Rather, by demonstrating that lexical frequency effects persist, even after all picture recognition processes have run their course, we are able to establish that the frequency effect does not reside exclusively in the picture recognition stage. Furthermore, by establishing that the frequency ef-

fect does not arise in articulatory processes, we conclude that the lexical frequency effect is in fact lexical.

In summary, the results obtained suggest that the greater part of the lexical frequency effect is not reducible to variables that affect the recognition stages; even after full completion of input processes, a robust frequency effect was obtained. These results confirm lexical frequency as a signature of lexical access, and in doing so, validate the use of the lexical frequency effect as means to constrain and test theories of lexical access.

AUTHOR NOTE

REFERENCES

Alario, F.-X., Costa, A., & Caramazza, A. (2002). Frequency effects in noun phrase production: Implications for models of lexical access. *Language & Cognitive Processes*, **17**, 299-319.

Barry, C., Hirsch, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of acquisition, word frequency, and the locus of repetition priming of picture naming. *Journal of Memory & Language*, **44**, 350-375.

Bartram, D. J. (1976). Levels of coding in picture–picture comparison tasks. *Memory & Cognition*, **4**, 593-602.

Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, **10**, 344-380.

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.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, **35**, 116-124.

Goldinger, S. D., Azuma, T., Abramson, M., & Jain, P. (1997). Open wide and say "blah!" Attentional dynamics of delayed naming. *Journal of Memory & 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 & Language*, **38**, 313-338.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, **5**, 67-103.

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, & 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, & Cognition*, **29**, 432-438.

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin*, **120**, 113-139.

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

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

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, **20**, 705-714.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, **17**, 273-281.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, **64**, 249-284.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 174-215.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory & Language*, **51**, 247-250.

Vitevitch, M. S., Armbrüster, J., & Chu, S. (2004). Sublexical and lexical representations in speech production: Effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 514-529.

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

NOTE

1. There is a growing debate on whether putative frequency effects truly reflect age of acquisition (AoA) or lexical frequency. Our use of the term *frequency* is not intended to prejudge this issue. It simply reflects our use of lexical frequency norms.

APPENDIX A
Target Pictures for Experiment

High-Frequency Items		Low-Frequency Items	
Animals	Artifacts	Animals	Artifacts
camel	bottle	ant	axe
cat	box	crab	couch
chicken	bus	deer	crib
cow	chair	dolphin	ladle
dog	clock	lizard	lock
elephant	cup	moose	rake
horse	lamp	penguin	razor
lion	mirror	squid	saw
pig	table	squirrel	screw
snake	train	swan	stool

APPENDIX B
Mean Values for Each Frequency Category
by Controlled Variables

	High-Frequency Targets	Low-Frequency Targets
Concreteness	605.26	597.00
Imageability	608.53	586.17
Number of syllables	1.40	1.30
Number of phonemes	3.95	4.05
Number of neighbors	15.55	14.50
Neighborhood frequency	110.51	169.34
Segments sum	0.24	0.23
Biphones sum	0.03	0.02
Segments average	0.06	0.05
Biphones average	0.01	0.01