Clinical Aphasiology, Vol. 23, 1995, pp. 107–119

Relationship of Word Frequency in Printed Materials and Judgments of Word Frequency in Daily Life to *Boston Naming Test* Performance of Aphasic Adults

Robert H. Brookshire and Linda E. Nicholas

The effect of word frequency on aphasic adults' word retrieval and production has been studied intermittently for many years (Newcombe, Oldfield, & Wingfield, 1965; Rochford & Williams, 1965; Wepman, Bock, Jones, & Van Pelt, 1956; Williams & Canter, 1982). The general conclusion is that the frequency with which words occur in a language has strong effects on aphasic adults' word retrieval and production; that is, less frequent words are more difficult for aphasic adults to retrieve and produce than are more frequent ones. This conclusion is based primarily on the results of studies of confrontation naming, in which aphasic adults were found to name pictures with high-frequency names more successfully than pictures with low-frequency names.

The *Boston Naming Test* (BNT) (Kaplan, Goodglass, & Weintraub, 1983) is widely used to assess the confrontation naming performance of adults with brain damage. The test contains 60 line drawings that depict objects to be named by examinees. The familiarity of the test items appears to decrease as the test progresses; however, the test manual does not say how the order of test items was established, and it provides no information about either the frequency of occurrence in English of test item names or their probable familiarity to adults in the United States.

In studies of the effects of word frequency on aphasic adults' confrontation naming, word frequency has been determined according to norms for printed materials, perhaps because adequate norms for spoken frequency of occurrence in United States English do not exist. Jones and Wepman (1966) published spoken word counts for a corpus of

136,000 words, but this corpus is not large enough to provide reliable data, and many relatively common words do not appear in the corpus.

Norms for printed materials, however, may not be valid for estimating frequency of occurrence in spoken language. Adult-to-adult conversations contain more high-frequency words and fewer low-frequency words than do printed texts (Hayes, 1989). Furthermore, norms based on printed frequency are likely to reflect sampling biases toward words that are associated with print media, but that occur infrequently in spoken language. Table 1 illustrates this point.

The three italicized words in Table 1 represent words that are likely to be associated with print media. The other words seem more likely to occur in daily-life speech. The word *page* is extremely frequent in print materials; it occurs about 430 times per million and is the 227th most frequent word in the corpus. In contrast, the words *gas* and *salt*, which one would expect to be more common in daily life, hold much lower ranks in print. The frequencies in Table 1 suggest that one is about 5 times more likely to hear the word *page* in daily life than to hear the words *gas* or *salt*, and about 30 times more likely to hear the word *page* than to hear the words *sack* or *dime*.

In this paper, we report a study of the relationships among aphasic adults' confrontation naming, word frequency in printed materials, and judgments of word frequency in daily life. The study addressed five questions:

Table 1. Examples of Bias in Word Frequency Norms that Are Based on Printed Materials

Word	*U-Value*	*Rank*
Page	429.70	227
Gas	86.23	1009
Salt	75.28	1117
Chapter	53.17	1527
Toes	28.30	2502
Hospital	25.22	2734
Author	33.10	2419
Sack	14.49	4005
Dime	14.46	4015

Note: The italicized word in each word set is likely to be more frequent in printed materials than in daily-life interactions. The other two words in each set occur more frequently in spoken language than in printed materials.

1. What is the relationship between the order of items in the BNT and aphasic adults' accuracy in naming them?
2. What is the relationship between the order of items in the BNT and word frequency in printed materials?
3. What is the relationship between word frequency in printed materials and aphasic adults' performance on the BNT?
4. What is the relationship between the order of items in the BNT and judgments of word frequency in daily life?
5. What is the relationship between judgments of word frequency in daily life and aphasic adults' performance on the BNT?

METHOD AND RESULTS

Subjects

We tested 20 aphasic adults (6 with nonfluent aphasia, 14 with fluent aphasia) with the BNT, using administration and scoring procedures described by Nicholas, Brookshire, MacLennan, Schumacher, and Porrazzo (1989). The aphasic subjects were at least 3 months postonset of a single left hemisphere cerebrovascular brain injury, and ranged in age from 51 to 77 years (M = 64.9, SD = 6.8) and in education from 10 to 16 years (M = 13.1, SD = 1.7). Their aphasia severity, as measured by a four-subtest shortened version (SPICA) (Disimoni, Keith, & Darley, 1980) of the *Porch Index of Communicative Ability* (Porch, 1971), ranged from the 40th to the 80th percentile (M = 63.7, SD = 14.5).

Evaluation of Experimental Questions

Question 1. *What is the relationship between the order of items in the BNT and aphasic adults' accuracy in naming them?* The accuracy with which the 20 aphasic adults named each item in the BNT is given in Table 2 and is shown graphically in Figure 1. The percentage of aphasic subjects who correctly named each BNT item diminished across the test. However, there was great variability, with some early items (e.g., "octopus") being disproportionately difficult to name and some late items (e.g., "funnel") being disproportionately easy. A Pearson correlation coefficient calculated on the relationship between the order of items in the BNT and the accuracy of naming yielded r = .74, which suggested a relatively strong relationship.

Table 2. U-Values, Subjective Ratings, and 20 Aphasic Adults' Naming Accuracy for *Boston Naming Test* (BNT) Items

BNT Item No.	*Stimulus Word*	*U-Value*	*Subjective Rating*	*Accuracy (percent)*
1	Bed	166.03	94.55	95
2	Tree	562.36	86.60	95
3	Pencil	62.58	89.20	90
4	House	659.19	94.35	75
5	Whistle	30.56	47.30	95
6	Scissors	9.82	73.40	75
7	Comb	11.94	84.95	90
8	Flower	176.04	73.70	90
9	Saw	523.58	47.15	85
10	Toothbrush	1.10	81.55	75
11	Helicopter	16.94	40.25	60
12	Broom	14.22	73.00	85
13	Octopus	3.53	21.10	25
14	Mushroom	9.98	53.75	50
15	Hanger	0.83	64.50	65
16	Wheelchair	0.31	46.75	70
17	Camel	21.43	28.70	75
18	Mask	12.61	42.95	85
19	Pretzel	0.35	57.30	80
20	Bench	23.41	60.10	80
21	Racquet	0.22	39.55	50
22	Snail	12.66	25.55	75
23	Volcano	26.54	26.85	30
24	Sea horse	0.41	10.65	25
25	Dart	5.36	29.45	70
26	Canoe	35.85	37.50	85
27	Globe	35.25	35.60	40
28	Wreath	2.92	32.05	63
29	Beaver	25.25	28.00	58
30	Harmonica	2.30	28.30	37
31	Rhinoceros	2.09	19.95	44
32	Acorn	5.05	34.95	56
33	Igloo	0.73	13.05	56
34	Stilts	1.34	13.45	72
35	Dominoes	0.12	23.95	50
36	Cactus	11.37	31.90	44
37	Escalator	0.50	52.60	29
38	Harp	2.51	19.25	76
39	Hammock	2.56	28.40	35
40	Knocker	0.23	18.65	41
41	Pelican	3.01	15.10	47
42	Stethoscope	0.68	26.85	18

(Continued)

Table 2. (continued)

BNT Item No.	*Stimulus Word*	*U-Value*	*Subjective Rating*	*Accuracy (percent)*
43	Pyramid	15.69	22.30	41
44	Muzzle	3.52	18.65	47
45	Unicorn	0.54	9.80	12
46	Funnel	7.49	38.05	65
47	Accordion	1.54	19.55	53
48	Noose	0.65	12.30	35
49	Asparagus	0.75	31.40	41
50	Compass	30.05	25.35	35
51	Latch	1.12	25.70	35
52	Tripod	0.96	20.55	47
53	Scroll	2.40	12.90	65
54	Tongs	1.37	28.30	59
55	Sphinx	0.75	7.60	24
56	Yoke	2.38	7.20	24
57	Trellis	0.31	16.20	31
58	Palette	0.16	12.55	13
59	Protractor	4.03	17.20	6
60	Abacus	3.20	4.10	0

Question 2. *What is the relationship between the order of items in the BNT and word frequency in printed materials?* To determine the frequency of occurrence in printed materials for the items in the *Boston Naming Test,* we obtained the U-value for each BNT item from the *American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971), which provides norms for a corpus of more than 5 million words obtained from printed materials. The U-value for a word represents the word's estimated frequency of occurrence per million tokens. The U-values for BNT test item names are given in Table 2 and are shown, in BNT test order, in Figure 2. There was no systematic decrement in U-values across BNT items, and there was also substantial variability in U-values from early to late items, with many early items occurring infrequently in printed materials (e.g., *scissors*) and some late items occurring relatively frequently (e.g., *compass*). A Pearson correlation coefficient calculated on the relationship between the order of items in the BNT and word frequency in printed materials yielded $r = .43$, suggesting that BNT test order does not strongly reflect word frequency in printed materials.

Question 3. *What is the relationship between word frequency in printed materials and aphasic adults' performance on the BNT?* To evaluate this

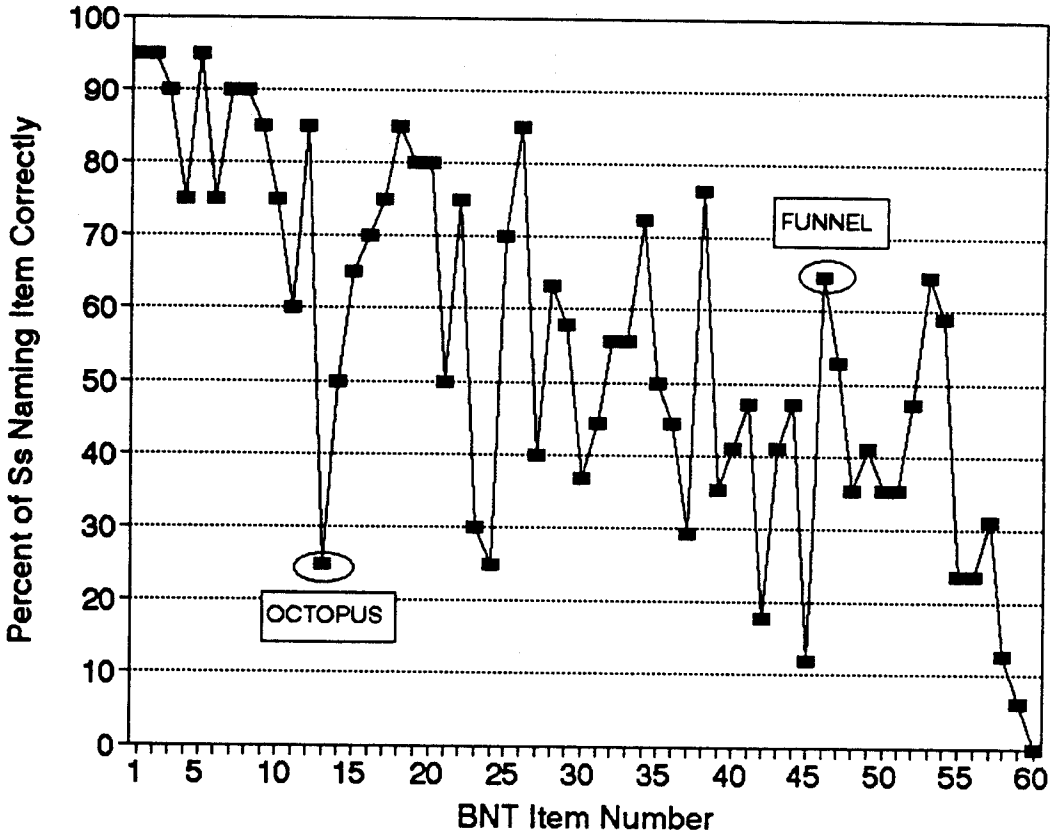

Figure 1. Order of items in the *Boston Naming Test (BNT) and* aphasic subjects' naming accuracy for the items.

question, we plotted the BNT performance of the 20 aphasic adults according to the U-values of the BNT test items. The results are illustrated in Figure 3. There was little relationship between an item's frequency of occurrence in print materials and the probability that it would be correctly named by aphasic adults. There was a general decrement in naming performance as U-values decreased, but the variability was great, and in many cases, items with higher U-values were less likely to be named correctly than items with lower U-values. For example, *abacus*, with a U-value of 3.2, was never correctly named, whereas *pretzel*, with a U-value of 0.35, was named correctly 80% of the time. A Pearson correlation coefficient, calculated on the relationship between U-values and aphasic subjects' performance, yielded $r = .37$, suggesting that an item's frequency of occurrence in printed materials was not strongly related to the probability that it would be correctly named by aphasic adults.

Question 4. *What is the relationship between the order of items in the BNT and judgments of word frequency in daily life?* Data for adult-to-adult conversations reported by Hayes (1989) suggest that the BNT item

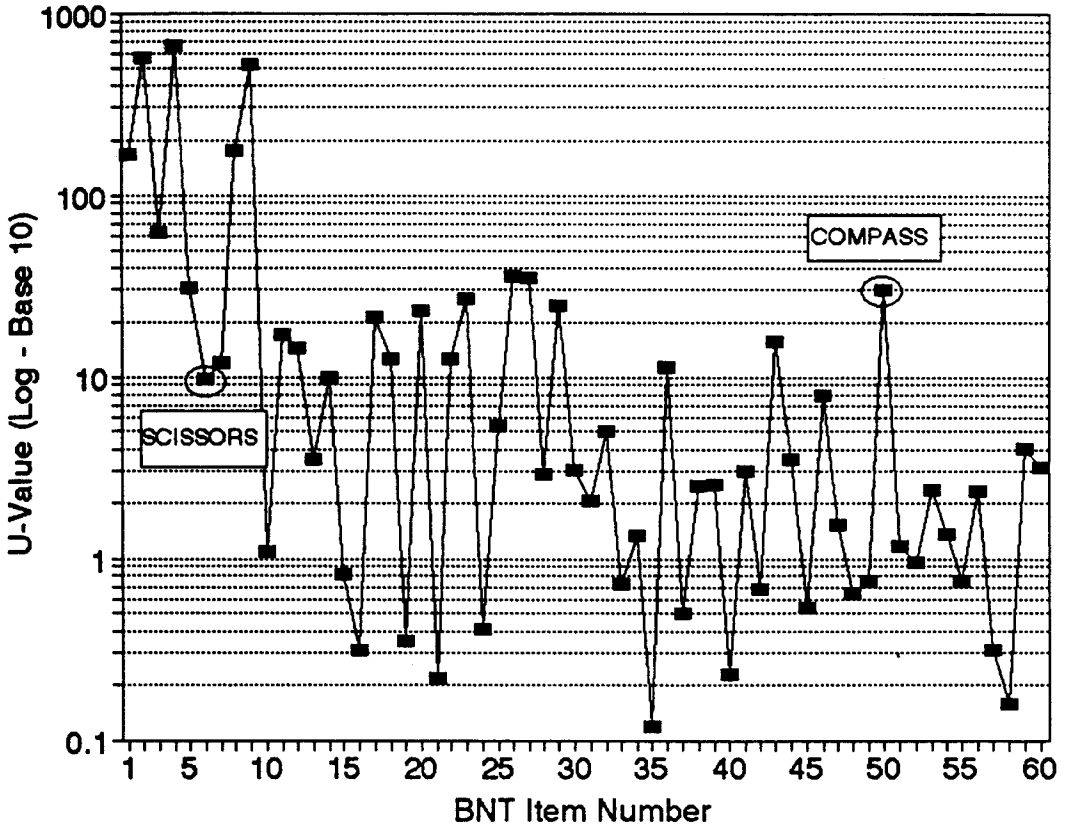

Figure 2. Order of items in the *Boston Naming Test* (BNT) and their U-values (frequency of occurrence per million tokens).

names are not representative of words that are likely in daily-life adult-to-adult conversations. Hayes reported frequency-of-occurrence values for adult-to-adult conversations, using the Carroll et al. (1971) norms. He reported that about 83% of the words in adult-to-adult conversations are within the 1,000 most frequently occurring words in English. To be among the 1,000 most frequent words, an item must have a U-value of 87 or greater. The U-values in Table 2 show that most of the words tested by the BNT are not likely to be encountered in adult-to-adult conversations. Only 5 of the 60 BNT test items fall within the 1,000 most frequent words (*bed, tree, house, flower,* and *saw*).

According to Hayes (1989), about 94% of the words in adult-to-adult conversations are within the 5,000 most frequent words in English, having U-values of 10 or greater. Only 26 (or 43%) of the 60 BNT items fall within the 5,000 most frequent words. Figure 4 shows how the distribution of items in the BNT differs from that reported by Hayes for adult-to-adult conversations. It shows that conversations are heavily weighted with high-frequency words, whereas the BNT is heavily weighted with low-frequency words. Hayes's data for conversations are based on frequency-of-occurrence norms for printed mate-

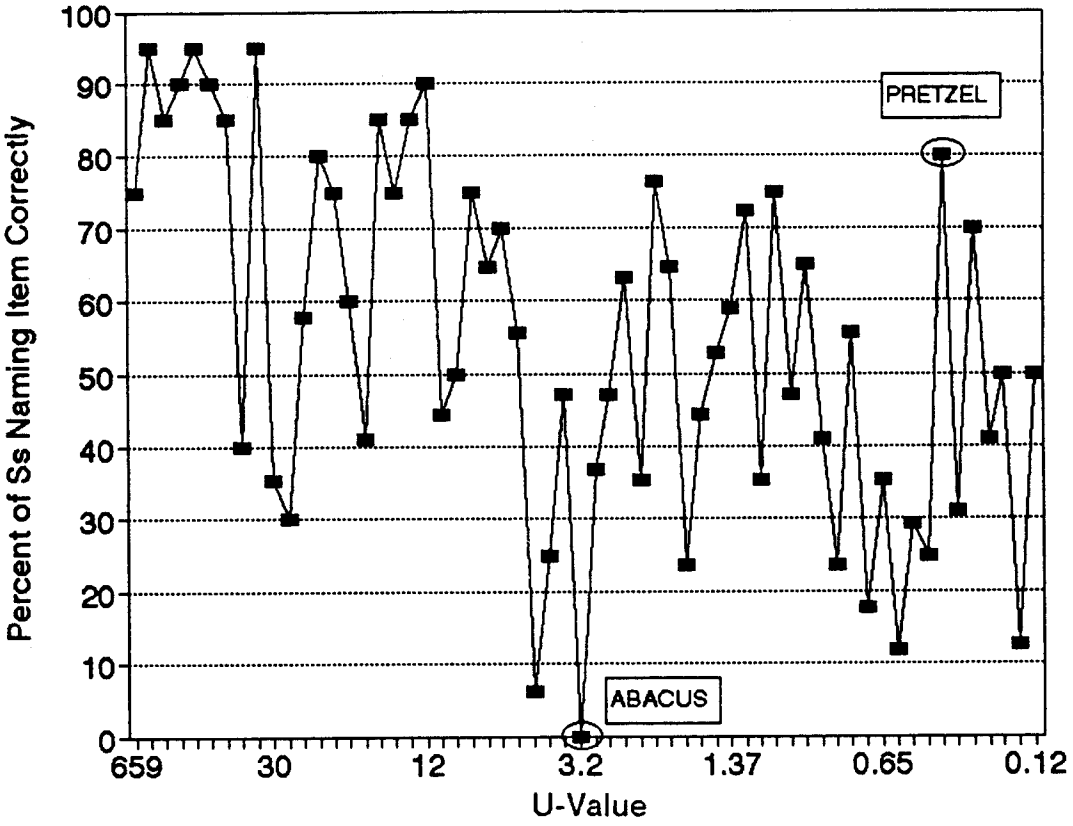

Figure 3. U-values (frequency of occurrence per million tokens) for *Boston Naming Test* items and aphasic subjects' naming accuracy for the items.

rials. To evaluate the validity of using norms based on printed materials to estimate word frequency in daily life, we asked 20 normal adults to rate the BNT item names in terms of their frequency of occurrence in daily-life speech. The judges' ages ranged from 19 to 60 years (M = 31.9, SD = 12.94) and their education ranged from 14 to 21 years (M = 16.6, SD = 2.18). Each judge was given a printed form that listed the BNT item names in random order. Then they were asked to estimate on a 0 (*low*) to 100 (*high*) scale how frequent they thought each word would be in daily-life speech. The results are summarized in Figure 5.

There was a gradual decline in the subjective frequency of item names across the BNT, but with substantial item-to-item variability, especially for items early in the test. The gradual decline in subjective frequency estimates across BNT items suggested that the order of items in the BNT reflects, to some extent, their potential familiarity to adults. To evaluate the strength of this relationship, we calculated a Pearson correlation coefficient between the order of items in the BNT and the subjective judgments of frequency in daily life. We obtained a corre-

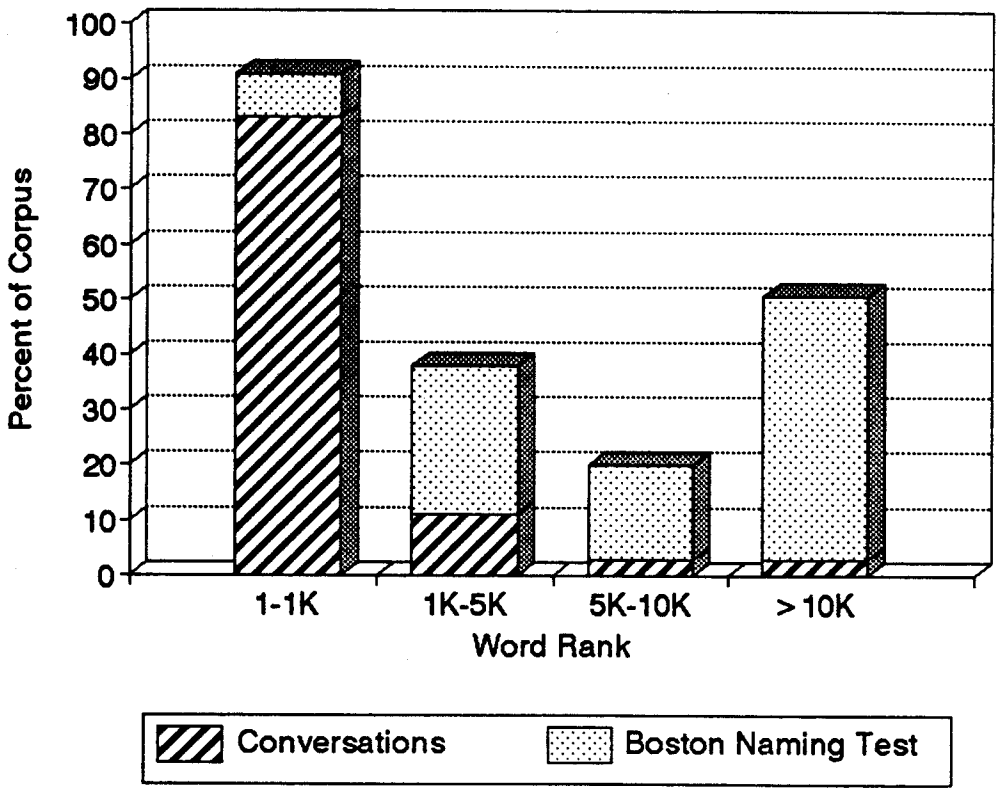

Figure 4. Word ranks for *Boston Naming Test* items and for words in adult-to-adult conversations (Hayes, 1989).

lation of $r = .80$, which suggested that BNT test order represents, in general, decreasing daily-life familiarity of item names for normal adults.

Question 5. *What is the relationship between judgments of word frequency in daily life and aphasic adults' performance on the BNT?* To address this question, we plotted the BNT performance of the aphasic subjects with the test items arranged in decreasing order of judged daily-life frequency of occurrence. Figure 6 illustrates the results. Aphasic adults' naming accuracy gradually declined as subjective frequency decreased, but once more there was substantial variability from item to item, with some frequent items (e.g., *escalator*) yielding low accuracy, and some infrequent items (e.g., *stilts*) yielding high accuracy. A correlation coefficient calculated on the relationship between aphasic subjects' accuracy of naming and subjective frequency of occurrence yielded a correlation of $r = .71$, suggesting that aphasic adults' success at naming BNT items was quite strongly related to their judged frequency of occurrence in daily life.

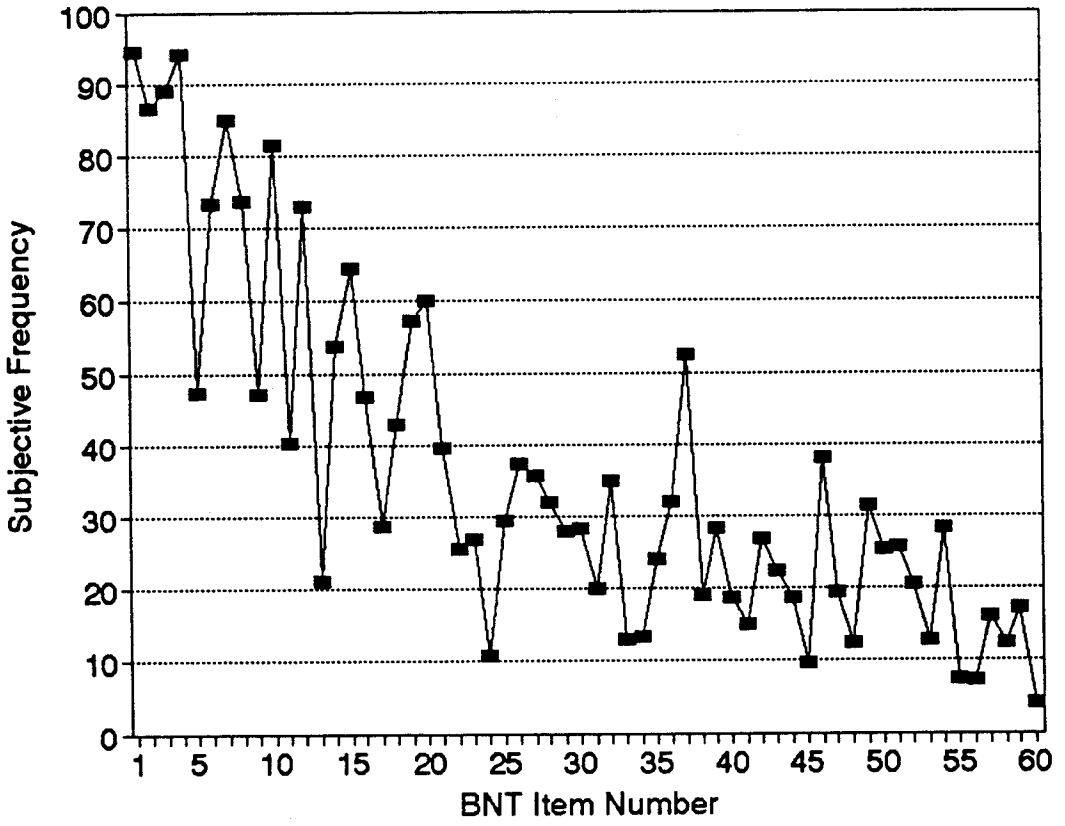

Figure 5. Order of items in the *Boston Naming Test* (BNT) and their judged frequency of occurrence in daily life.

DISCUSSION

The results of this study show that there is no dependable relationship between a word's frequency of occurrence in printed materials and aphasic adults' success in naming it in a confrontation naming task. Consequently, using printed frequency-of-occurrence values to manipulate stimulus difficulty in clinical or experimental word retrieval tasks appears inadvisable, if one is concerned with performance on individual stimuli. This is especially true for words occurring only a few times per million tokens. For such low-frequency words, small perturbations in frequency of occurrence caused by sampling artifacts can have large effects on their rank. If one is concerned not with the frequency of occurrence of individual stimuli, but with overall frequency of occurrence for large sets of words, use of norms based on printed frequency of occurrence (as Hayes did in analyzing adult-to-adult conversations) may be appropriate, because distortions intro-

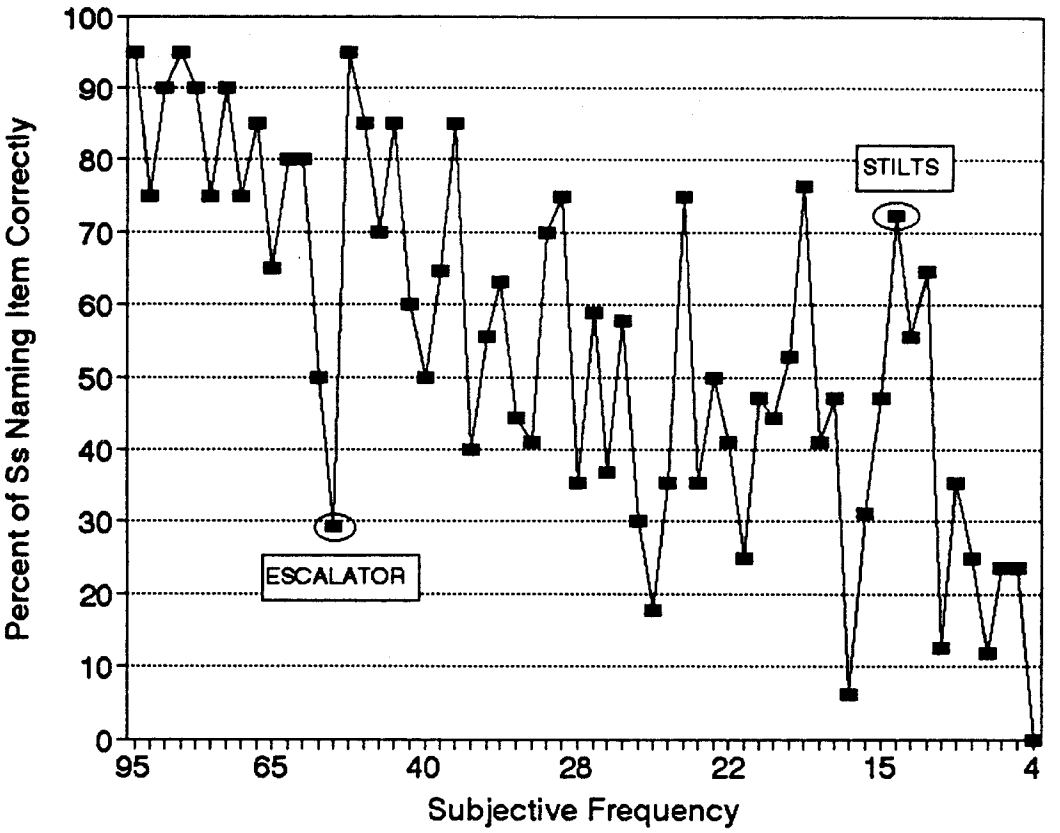

Figure 6. Subjective frequency of occurrence in daily life for *Boston Naming Test* items and aphasic subjects' naming accuracy for the items.

duced by sampling biases are likely to occur as random effects within sets of words.

Our results suggest that normal adults' subjective judgments of daily-life word frequency are reasonably good predictors of accuracy of word retrieval for aphasic adults. This conclusion is also supported by the results of several studies with college students, in which subjective judgments of word frequency were shown to be strong predictors of normal subjects' speed and accuracy in lexical decision tasks (Gernsbacher, 1984; Gordon, 1985). In the absence of published norms for word frequencies in daily-life speech, such subjective judgments may be a valid way to estimate the daily-life frequency of words. Our judges were younger and may have been better educated than the average aphasic adult. This may limit, to some extent, the generalizability of these results to the adult aphasic population. That our judgments of word familiarity were obtained from a relatively small group of judges may account for some of the item-to-item variability in judgments of familiarity reported herein. We are currently designing a study

with a larger sample of judges who better represent the age range for aphasic adults, and with a larger corpus of words.

Our results suggest that the order of test items in the BNT represents, in general, decreasing levels of daily-life frequency of occurrence and increasing difficulty of word retrieval. However, users should be aware that most of the BNT test items are words that are unlikely to be encountered in daily-life conversations. Therefore, the BNT may provide information about the size of an individual's vocabulary, or the person's ability to retrieve infrequently occurring words in a confrontation naming task, but it may provide little dependable information about word retrieval in daily-life speech.

ACKNOWLEDGMENTS

This research was supported by the Department of Veterans Affairs Medical Research Service and by the Medical Research Service, Minneapolis Veterans Medical Center.

REFERENCES

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage word frequency book.* New York: Houghton Mifflin.

Disimoni, F. G., Keith, R., & Darley, F. L. (1980). Prediction of PICA overall score by short versions of the test. *Journal of Speech and Hearing Research, 23,* 511–516.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology, 113,* 256–281.

Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory and Language, 24,* 631–645.

Hayes, D. P. (1988). Speaking and writing: Different patterns of word choice. *Journal of Memory and Language, 27,* 572–585.

Hayes, D. P. (1989). *Guide to the lexical analysis of texts* (Tech. Rep. Series 89-6). Ithaca, NY: Cornell University, Department of Sociology.

Jones, L. V., & Wepman, J. W. (1966). *A spoken word count.* Chicago: Language Research Associates.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test.* Philadelphia: Lea & Febiger.

Newcombe, F., Oldfield, R. C., & Wingfield, A. (1965). Object-naming by dysphasic patients. *Science, 207,* 1217–1218.

Nicholas, L. E., Brookshire, R. H., MacLennan, D. L., Schumacher, J. G., & Porrazzo, S. (1989). Revised administration and scoring procedures for *The Boston Naming Test* and norms for non–brain-damaged adults. *Aphasiology, 3*, 569–580.

Porch, B. E. (1971). *The Porch Index of Communicative Ability.* Palo Alto, CA: Consulting Psychologists Press.

Rochford, G., & Williams, M. (1965). Studies in the development and breakdown of the use of names: Part IV. The effects of word frequency. *Journal of Neurology, Neurosurgery, and Psychiatry, 28*, 407–413.

Wepman, J. M., Bock, R. D., Jones, L. V., & Van Pelt, D. (1956). Psycholinguistic study of aphasia: A revision of the concept of anomia. *Journal of Speech and Hearing Disorders, 21*, 468–477.

Williams, S. E., & Canter, G. J. (1982). The influence of situational context on naming performance in aphasic syndromes. *Brain and Language, 17*, 92–106.