The purpose of this study was to compare student's tested recognition of the Dolch 220 words with their responses to the 220 highest frequency words found by Kucera and Francis in their "Computational Analysis of Present-Day American English." Subjects consisted of 155 second graders and 179 third graders from classrooms in four schools in a large suburban school division in Winnipeg, Manitoba. The factors of grade level, IQ, socioeconomic level, and sex were considered. From a statistical analysis, it was found that significant differences occurred at the .05 level between tests, between grades, and among IQ levels on the Dolch test. On the Kucera/Francis 220, only the difference between socioeconomic levels and sexes did not reach significance. Based on the findings, the following conclusions were reached: (1) the Dolch test appears to differentiate between the grade levels considered; (2) the Dolch 220 appears to differentiate between all three IQ levels considered; (3) the Dolch 220 test correlates higher to IQ scores than the other test considered; and (4) the Dolch 220 test correlates higher with a comprehension measure administered five months later than does the other test considered. (WR)
DOES THE DOLCH DO?

Considerable interest has been generated recently in word lists by Johnson (1971), Johns (1971), Otto and Chester (1972), Durr (1973), and Harris and Jacobson (1973-74). Consequently, the inveterate Dolch 220 has come under close scrutiny. Although the pervasiveness of this list is not precisely known, Jacobs (1967, p. 2) concluded from his interview of teachers that (1) the Dolch lists are familiar to most teachers in remedial and primary reading, (2) the use of the lists may be greatest in the area of remedial reading, and (3) few people are aware of the origins and ages of the lists.

The Dolch 220, also known as "The Basic Sight Word Test", was based on vocabulary studies of the 1920's and is still used in its original form. The current form was copyrighted by Garrard Publishing Company in 1942.

It seems logical to question the usefulness of this list simply because of its age and it is not surprising to find comments such as the following:
"The Dolch List, as a corpus, has outlived its usefulness and ... a more adequate substitute is available." (Johnson, 1971, pp. 449 and 451).

"Yet the fact is that in addition to becoming somewhat passe over more than three decades, the list [Dolch 220] had a questionable empirical base in the first place." (Otto and Chester, 1972, p. 435).

This paper describes a study which was designed to investigate the current usefulness of the Dolch 220. It was decided to test second and third grade children's actual recognition of the Dolch words and compare the results to scores on a test suggested by Johnson (1971). It is important to note, then, that this study deals with recognition of words and is not a study based on text demand (frequency counts, occurrence in basal readers, etc.) To provide some indication of the test's predictive validity, the Botel Word Opposites Test (1966) was administered five months later. This test was used since it correlated highly, from 0.82 to 0.90, with other comprehension measures. (Botel, 1970).

Before describing the results it is instructive to consider the research on word recognition tests. It is essential to keep in mind four factors influencing vocabulary study: (1) the source - whether from children or adults, (2) the modality - whether oral or written, (3) the method of collection - whether frequency counts, free-association, from basal readers, tested knowledge, adult judgment, or some combination, and (4) the period or time of collection - whether in the 1920's or late 1960's and early 1970's, etc.

Related Research

The importance of a basic sight vocabulary was clearly articulated by Dolch. His list contained the words which indicated the structural relationships between members of the form classes (nouns, verbs, adjectives, adverbs) as well as a number of "irregular" words, both being essential to all subject matter material. Furthermore, Dolch (1960) demonstrated that the basic sight
vocabulary comprised over one-half of the running words in the subjects of Reading, Arithmetic, Geography, and History.

Johns (1971) replicated this analysis for five current reading series and found that the Dolch words still comprised over one-half of the running words, although the percentage at each grade level (1-6) was somewhat lower. Johnson and Barrett (1972) reported similar findings for Johnson's list which contains many of the same words.

The similarity of these counts appears to indicate that publishers take into account the various studies of basic vocabularies or that they are governed by the fact that these words are necessary to all writing.

Vocabulary research has also influenced the measurement of reading difficulty, since according to Klare (1963), vocabulary accounts for the greatest amount of variance in readability formulas. And since readability formulas are also utilized to control basal reader material, the cycle of interdependent factors pointed out by Froese (1971) is completed.

A further problem related to the study of vocabulary is an environmental one—changes in technology, advancements in the sciences, and cultural influences have their concomitant language influences. Jacobs (1967) has documented one aspect of this change by replicating the Buckingham-Dolch Free-Association Word Study. He concluded that the 1926 Free-Association List is no longer representative of current students' vocabularies in terms of list content, grade-level assignment, and student performance."

Johnson (1971) also referred to this change when he stated: "Of the 220 words on the Dolch Primary Word List, representative of the 1920's, eighty-two words or 37 percent are not among the most frequently occurring 220 words in the Kučera and Francis corpus compiled in the 1960's." This statement must, however, be interpreted carefully since the Kučera and Francis
is based entirely on frequency of occurrence, includes nouns, and is derived from adult material—all of these are facets not common to the Dolch 220.

Lexicography also is concerned with language shifts and regular revisions of dictionaries are practically assumed. The motivation for the Word Frequency Book (1971) rests on these grounds but breaks new ground in applying computerized techniques to the vocabulary of children in grades 3 through 9 (and was likely inspired by Kučera's work with adult vocabulary).

A British study by Edwards and Gibbon (1964) examined the written vocabulary of children ages six to eight years and found that it expressed "exciting scientific wonders," that their lists overlapped but were not closely alike the American lists (of Rinsland, Thorndike, Gates and Dolch), and that the earlier lists were "out-of-date in respect of the scientific advances of recent years."

A further consideration to be taken into account when studying vocabulary is the shortcoming of frequency counts alone. Familiarity, meaning, and comprehension do not bear a one-to-one relationship to frequency as pointed out by Klare (1963). Dolch (1951) noted this and suggested that the factors of opportunity of use and emotional set are the causes of the discrepancy between frequency of use and tested word knowledge.

The Study

The purpose of this study was to compare students' tested recognition of the Dolch 220 words with their responses to the 220 highest frequency words found by Kučera and Francis in their Comprational Analysis of Present-Day American English (1967). The scores on the Dolch test were also compared to current student I.Q.'s and comprehension scores as tested five months later by the Botel Words Opposite Test.
Subjects for the study were 155 second graders and 179 third graders from classrooms in four schools in a large suburban school division in Winnipeg, Manitoba, Canada. Winnipeg is a city with a population of approximately 560,000. The factors of grade level, I.Q. level, socioeconomic level, and sex were considered.

The findings are presented in Table I.

**TABLE I**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Subjecs</th>
<th>Means &amp; Standard Deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Dolch 220</td>
</tr>
<tr>
<td>1. Grade 2</td>
<td>155</td>
<td>205(19.5)</td>
</tr>
<tr>
<td>3</td>
<td>179</td>
<td>217(5.1)</td>
</tr>
<tr>
<td>2. I.Q. Upper</td>
<td>11p</td>
<td>216(7.6)</td>
</tr>
<tr>
<td>Middle</td>
<td>15</td>
<td>212(14.6)</td>
</tr>
<tr>
<td>Lower</td>
<td>56</td>
<td>203(21.2)</td>
</tr>
<tr>
<td>3. Sex Middle</td>
<td>181</td>
<td>213(12.2)</td>
</tr>
<tr>
<td>Lower</td>
<td>154</td>
<td>210(17.7)</td>
</tr>
<tr>
<td>4. Sex Boys</td>
<td>164</td>
<td>210(16.0)</td>
</tr>
<tr>
<td>Girls</td>
<td>170</td>
<td>213(14.0)</td>
</tr>
</tbody>
</table>

From a statistical analysis, it was found that significant differences occurred at the .05 level between tests, between grades, and among I.Q. levels on the Dolch test. On the Kučera/Francis 220 only the difference between middle and lower I.Q.'s was significant. The differences between socioeconomic levels and sexes did not reach significance.
A further analysis consisted of correlating the I.Q.'s, Botel Word Opposite scores, and the scores on the two vocabulary tests. The results are presented in Table II.

TABLE II


<table>
<thead>
<tr>
<th></th>
<th>Dolch 220</th>
<th>Kucera/Francis 220</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 2</td>
<td>I.Q.</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>Botel</td>
<td>.51</td>
</tr>
<tr>
<td></td>
<td>Kucera/Francis</td>
<td>.84</td>
</tr>
<tr>
<td>Grade 3</td>
<td>I.Q.</td>
<td>.09 NS</td>
</tr>
<tr>
<td></td>
<td>Botel</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>Kucera/Francis</td>
<td>.82</td>
</tr>
</tbody>
</table>

All correlation coefficients except those marked NS (non-significant) are significant at the .01 level.

The fact that few grade three students made errors on these tests obviously resulted in a ceiling effect and hence the non-significant findings.

However, the correlations between the Dolch 220 and I.Q., as well as between the Dolch 220 and the delayed Botel test are considerably higher than the same comparison for the Kucera/Francis 220.
Conclusions

A consideration of the data leads to the following conclusions about the Dolch 220:

1. The Dolch 220 test appears to differentiate between the grade levels considered.
2. The Dolch 220 test appears to differentiate between all three I.Q. levels considered.
3. The Dolch 220 test correlates higher to I.Q. scores than the other test considered.
4. The Dolch 220 test correlates higher with a comprehension measure administered five months later than does the other test considered.

In consideration of the original question, the conclusion appears to be that the Dolch does do!
REFERENCES CITED


