THE FIRST 4,000 WORDS LIST AND PLACEMENT TESTS: 1-30-09

Michael F. Graves
University of Minnesota, Emeritus

Gregory C. Sales
Seward Incorporated

The Purpose, Origin, and Development of the First 4,000 Words List

The First 4000 Word List and accompanying placement tests were constructed for a project designed to ensure that first-fourth grade students have mastered a basic listening vocabulary. Although we call the list The First 4000 Words, it is actually a list of 3,541 words. In creating the list, we relied primarily on Zeno et al.'s *Educator's Word Frequency Guide* (1995) and the work of Hiebert (2005). In identifying a set of words that children need to master by the end of fourth grade, Hiebert selected the 5,586 words that occurred at least 10 times per million running words in the Zeno et al. corpus. She then clustered regular inflected forms with their roots and arrived at a set of 3,913 word families. Finally, she divided these words into four sets: the first ~300, the next ~600, the next ~1600, and the last ~1900.

In creating The First 4000 Words list, we took these ~3,900 words, culled a few non-words, and deleted the 100 most frequent words, another 116 function words, and 148 proper nouns. The remaining 3,541 words are the words we worked with. We then rank ordered these words from the most frequent word "little" to the least frequent "arteries."

The First 4000 Word Placement Tests

We created two group-administered tests to assess students' knowledge of these words, a listening test and a reading test. The listening test is modeled after the Peabody Picture Vocabulary Test, and the reading test is modeled after the Gates-MacGinitie Reading Test Vocabulary Test Level BR. Each test contains 40 words, with different words being used in each test. The 40 words for each test are a stratified random sample from The First 4,000 Words list. For the most part, one word represents each of the 36 100-word frequency blocks, but in a few cases two words represent a single frequency block and a few blocks are not represented. For the listening test, the teacher reads a word aloud and the students select a picture from one of four alternatives. For the reading test, students view a single picture and select a word from one of four written alternatives.

Technical Information on the Placement Tests

Reliability

Coefficient alpha statistics for each grade level on the listening test were as follows: .79 for 1st grade, .77 for 2nd grade, .70 for 3rd grade, and .65 for fourth grade. Mark Davison, our assessment expert, recommended revising 3 items on which students scored very high or very low, and we have revised those items.

Coefficient alpha statistics for each grade level on the reading test were as follows: .88 for 1st grade, .89 for 2nd grade, .87 for 3rd grade, and .88 for fourth grade. Our assessment expert, again recommended revising 3 items on which students scored very high or very low, and we have done so.

Validity

We assessed validity with three measures.

First, correlations with teacher rating were computed for each grade. For the listening test, those correlations were .635, .557, .635, and .496 for grades 1 – 4 respectively. All of these correlations are significant at < .001. For the reading test, the correlations were .626, .618, .630, and .495 for grades 1 – 4 respectively. Again, all of these correlations are significant at < .001.

Second, for the fourth graders, who had taken the Minnesota Comprehensive Assessment (MCA) reading test the previous year, reading and listening scores were correlated with the MCA reading scores. The correlations are .628 for listening and .629 for reading. Both of these correlations are significant at <.001.

Third, mean scores were computed by grade to see if those means increased steadily. On the listening test, the number correct means were 23.99, 27.59, 29.54, and 31.59 for grades 1–4, respectively. A one-way ANOVA indicated that the differences among these means were significant at <.001. Similarly, for the reading test, the number correct means were 16.23, 25.76, 30.90, and 34.23 for grades 1–4 respectively. Again, a one-way ANOVA indicated that the differences among these means were significant at < .001.

Results of the Placement Tests

We administered the listening test and the reading test to approximately 500 1st-4th grade students in three schools.

The sample included approximately 370 native English speakers, 130 English language learners, and 50 special needs students.

The test produced a rich array of information on primary grade students' listening and reading vocabularies, only a few highlights of which are given here.

Many children know a lot of these very frequent words. For example, of the 517 students who took the reading test, 20 students got all 40 items correct, 19 got 39 correct, 24 got 38 correct, 23 got 37 correct, and 35 got 36 correct.

Some children, on the other hand, know few of these words. For example, on the listening test, the bottom 10% of the first graders knew fewer than 4 of the 10 words representing the 1,000 most frequent words. Similarly, on the reading test the bottom 10% of the first graders knew fewer that 3 of the 10 words representing the first 1,000 most frequent words.

English language learners know fewer words than their native English speaking counterparts. For example, on the listening test, English language learners knew about 7 of the 10 words representing the 1,000 most frequent words, while native English speakers knew about 9 of the 10 words.

Clearly, there are many students who are likely to profit from instruction on the first 4,000 most frequent words.

In the tables below Q(1-10) refers to the most frequent 1,000 words, Q(11-20) the next 1,000 most frequent and so on.

1. Mean Scores on the *Listening* and *Reading* Tests

Test	Q(1-10)	Q(11-20)	Q(21-30)	Q(31-40)	All Questions
Listening (n=539)	8.29	7.14	6.88	6.07	28.24
Reading (n=517)	7.92	7.53	6.01	5.57	27.03

2. Mean Scores on the *Listening* Test by Grade Level

Grade	Q(1-10)	Q(11-20)	Q(21-30)	Q(31-40)	All Questions
1 (n=131)	6.93	5.93	6.10	5.03	23.99
2 (n=140)	8.16	7.02	7.04	5.87	27.59
3 (n=122)	8.84	7.48	7.21	6.34	29.54
4 (n=146)	9.30	8.16	7.26	7.07	31.59

3. Mean Scores on the *Reading* Test by Grade Level

Grade	Q(1-10)	Q(11-20)	Q(21-30)	Q(31-40)	All Questions
1 (n=125)	5.33	4.76	2.98	3.17	16.23
2 (n=127)	7.98	7.44	5.45	4.87	25.76
3 (n=119)	8.87	8.61	7.10	6.30	30.90
4 (n=145)	9.31	9.11	8.19	7.61	34.23

4. Mean Listening Test Scores of the Bottom 5%/10% of Students by Grade

Grade	n	Mean
1	7/13	12.57/14.15
2	7/14	14.57/17.64
3	6/12	11.17/21.50
4	7/15	23.29/25.07

5. Mean Reading Test Scores of the Bottom 5%/10% of Students by Grade

Grade	n	Mean
1	6/13	2.00/4.00
2	6/13	9.67/12.38
3	6/12	16.17/19.17
4	7/15	17.17/21.87

6. Mean Scores on the *Listening* Test by Grade Level for ELL students

Grade	Q(1-10)	Q(11-20)	Q(21-30)	Q(31-40)	All Questions
1 (n=39)	5.76	4.68	5.00	4.15	19.95
2 (n=41)	6.90	6.14	6.36	5.31	25.05
3 (n=27)	7.74	6.37	6.89	5.52	26.30
4 (n=29)	8.43	7.23	6.83	6.23	29.69

7. Mean Scores on the *Reading* Test by Grade Level for ELL Students

Grade	Q(1-10)	Q(11-20)	Q(21-30)	Q(31-40)	All Questions
1 (n=37)	3.65	3.57	2.16	2.24	11.62
2 (n=35)	7.11	6.80	4.57	4.06	22.60
3 (n=27)	8.04	7.41	5.67	4.37	25.52
4 (n=28)	8.71	8.61	7.61	6.54	31.46

8. Mean Scores on *Listening* Test by Grade Level for Special Needs Students

Grade	Q(1-10)	Q(11-20)	Q(21-30)	Q(31-40)	All Questions
1 (n=8)	5.38	4.75	5.88	3.88	19.13
2 (n=20)	6.95	6.10	6.35	5.10	24.50
3 (n=12)	8.17	6.83	7.00	5.50	26.83
4 (n=18)	8.79	7.21	6.89	6.16	30.39

9. Mean Scores on *Reading* Test by Grade Level for Special Needs Students

Grade	Q(1-10)	Q(11-20)	Q(21-30)	Q(31-40)	All
1 (n=7)	3.42	3.29	1.29	2.29	10.29
2 (n=16)	6.38	5.94	4.38	4.31	21.06
3 (n=11)	8.00	7.10	5.82	4.18	25.18
4 (n=18)	8.44	8.06	6.67	6.06	29.22