DOCUMENT RESUME

ED 044 240 RE 002 910

AUTHOR Durr, William K.
TITLE A Computer Study of High-Frequency Words in Popular Trade Juveniles.
PUB DATE May 70
NOTE 18p.; Paper presented at the conference of the International Reading Association, Anaheim, Cal., May 6-9, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$1.00
DESCRIPTORS Beginning Reading, *Childrens Books, Computers, Independent Reading, *Primary Grades, *Recreational Reading, *Word Frequency

ABSTRACT

Word frequency was determined for library books that primary-grade children selected for free reading. A survey of librarians determined which books these children selected. This list was reduced to 80 books through evaluations by elementary school teachers. A computer analysis of each word in these books revealed 105,280 running words. When proper names, onomatopoeic words, and easily recognizable inflected forms and compounds were omitted, there were only 3,220 different words in all of these books. A frequency count of these different words revealed that just 10 words account for almost one-fourth of all running words, 25 words account for over one-third of all running words, and 188 words account for almost seven out of 10 of all running words. It was suggested that systematic teaching of these high-frequency words would help insure that children have the background needed to read library materials of their own choosing at an early age. References and tables are given. (Author/DE)

# A COMPUTER STUDY OF HIGH-FREQUENCY WORDS IN POPULAR TRADE JUVENILES



WILLIAM K. DURR

Michigan State University

This paper reports the results of a study determining word frequency in library books selected for reading by primary grade children.

## Purpose

Word frequency studies have shown a void in using, as a source of data, library books that primary grade children select to read when they make their own free choices. This study was designed to fill that void.

The results of several studies that have previously explored word frequency have been used in preparing developmental reading materials for primary grade children. The Thorndike-Lorge (1944), the Rinsland (1945), and the Dolch (1950) studies are probably among the best known. The word list for the Spache (1958) formula and the Stone (1956) revision of this list have been widely used also.

The Thorndike-Lorge (1944) word frequency list is based on a summary of four different counts, none of which concentrates on literature for primary grade children. This list shows frequencies only in general ranges; for example, its most refined breakdown is into words found in the first five hundred in frequency.

The Rinsland (1945) list does break word frequency for the first thousand down into hundreds. However, it does not break down the list to show more specific frequencies and it is based on words children used in writing, not in reading.

Dolch (1950) has presented a basic sight vocabulary of 220 words and a list of 95 common nouns which is often used as a basis for word frequency. Dolch's list was compiled by taking words common to three

other lists: the International Kindergarten Union list which included words orally known by children before first grade, the Gates list, and the Wheeler and Howell list containing words common to first grade readers published in the 1920's. Dolch's list was not based on word counts of books children select for reading, it is not broken down into specific frequencies, and it was limited by materials published nearly half a century ago.

Spache (1958) has used the Dale list of 769 words in his readability formula for primary reading materials and Stone (1956) has suggested a revision based on an examination of materials prepared to teach primary reading. Neither list is broken down into frequencies below the total list of 769 and neither list concentrates on materials that children select to read.

There is no intention here to deny the value of previously compiled lists. They serve the purposes for which they were compiled. It does seem, however, that there is need for an up-to-date primary vocabulary list which concentrates on books that primary grade children select from libraries for free reading and which is broken down into specific frequencies for each word. Previous lists do not meet this need.

Finding the specific words that the young free-reader would most likely meet should be helpful. First, it would enable authors of developmental reading programs to concentrate on relatively high-frequency words in building reading vocabulary for the beginning reader. Further, if some few words are found to occur a large percentage of the time, developing such words to instantaneous recognition would insure that the young reader could read a large percentage of his selected books with ease. With a

large percentage of the words instantly recognizable, the young reader would be better prepared to use contextual clues together with grapheme-phoneme correspondences to decode those words that he did not recognize instantly.

This study was undertaken with the expectation and hope that such words of sufficiently high frequency would be found.

## Procedure

The initial step in this study determined which books are presently popular with primary-grade children. It should be emphasized that no effort was made to determine which books adults told children to read. The question was, "What books do primary-grade children read when they have a free choice of books in libraries?"

To answer this question, children's librarians were contacted in a number of communities with populations ranging in size from about ten thousand to several million. These librarians were not randomly selected. They were carefully picked to represent communities from a variety of geographic areas and to represent communities of varying socio-economic levels.

Because the open-ended question asked of these librarians required a great deal of their time, each librarian was told that she would receive one hundred dollars worth of free books when she returned the completed questionnaire. Fifty per cent of the questionnaires were returned.

The composite list of several thousand popular titles compiled from these reports was found to be so long that a thorough analysis of all of them would have been economically infeasible. Inordinately time consuming,

and, from subjective examination, no more profitable in reaching the final goal of the study than would an examination of some of these materials.

There seemed to be three logical procedures for reducing this cumbersome quantity of titles to a manageable size. Books could be randomly selected from the several thousand titles without taking cognizance of the librarians' frequencies of listing. Or, selection could be made by taking those books that were listed most frequently until a pre-determined number was reached. Or, using this list as a base, additional expert advice could be sought to determine which books were most often selected for free reading in classroom settings.

It was decided that a combination of the last two procedures would be used. No arbitrarily pre-determined number would be selected, but those which were analyzed would be books which were frequently chosen by children. Accordingly, the librarians' list was submitted to experienced teachers who were well acquainted with children's free reading interests. These teachers, taking into account the frequency of choice by librarians, selected the final eighty titles that were analyzed.

Almost one year after the study began, the chosen eighty selections were sent for computer analysis. This analysis consisted of recording each and every word that appeared in all eighty selections and of making a frequency count determining the number of times each word appeared in the total number of selections.

## Results

The total number of <u>running</u> words in all eighty selections was 105,280. The total number of different words, however, was markedly less. There were only

5,791 different words in the total of more than one hundred thousand running words. Stated differently, despite the large total number of words, it was found that each word reappeared, on the average, more than eighteen times.

Closer analysis revealed that 2,571, or about forty-five per cent of the nearly 6,000 different words, fell into one of several special categories. Some, as might be expected, were proper names which were more or less unique to a particular story. Some, such as mooo and buzzz, were onomatopoeic. Many others were made by adding common bound suffixes to base words, such as apples and runs, and many were compounded forms of other words, such as fireplace and birthday.

Proper names were omitted since they are relatively unique to any given selection and since their precise pronunciation is frequently not determinable without outside assistance. For example, although we may be fairly sure that B-o-b in print is pronounced Bob, if Bob's last name is spelled S-m-i-t-h, we cannot be sure without checking if it is Bob Smĭth or Bob Smīth.

Onomatopoeic words were omitted because they, too, are relatively unique to the content of a story and because, by definition, the phoneme-grapheme correspondences are specifically designed to sound like and, hence, call to mind their referents.

Only the base words of common inflected forms and compound words were counted. When the base word is known, the correct reading of forms made by adding s, es, ing, or ed is frequently as much a product of the pupil's grammar as it is his recognition vocabulary or decoding skill. For example, in the sentence, "Johnny ate two apples," the pupil who is

accustomed to hearing and using the plural apples in the second noun slot will not be likely to read it in the singular form. If he is accustomed to the singuler form apple in that slot he is likely to read it in the singular form but still gain appropriate meaning for the sentence.

Omitting proper names and onomatopoeic words and counting only the base words of common inflected forms or compound words, left a total of only 3,220 truly different words in the total eighty selections.

Let's examine the frequency of occurrence of these 3,220 words.

Table 1. A Summary Analysis of One Hundred Eighty-Eight Words of More Than Eighty-Eight Frequencies in a Computer Study of Vocabulary in Popular Trade Juveniles.

| Frequencies | No. of Words | Percentage of Different Words | Total Occurrences | Percentage of Running Words |
|---|---|---|---|---|
| 1,500 or more | 10 | .31% | 25,010 | 23.76% |
| 565 or more | 25 | .78% | 37,218 | 35.35% |
| 376 or more | 47 | 1.46% | 47,137 | 44.77% |
| 190 or more | 94 | 3.23% | 59,892 | 56.89% |
| 121 or more | 141 | 4.38% | 67,160 | 63.79% |
| more than 88 | 188 | 5.84% | 72,027 | 68.41% |

Table 1 shows that the ten words with the highest frequencies accounted for only about three-tenths of one per cent of the over 3,000 different words, but these ten words occurred a total of 25,010 times in all running words. Each of these ten words occurred more than 1,500 times. Viewed differently, on the average, the young reader could expect to meet one of

these ten words in nearly every four words he read in the entire eighty books. The young reader who would develop instant recognition of just these ten words, then, would be equipped, on the average, to deal instantly with nearly one-fourth of all words in the books he is most likely to select from the library for free reading.

Although frequencies decline rapidly as we go beyond the first ten, the count of words still maintains frequencies high enough to warrant careful consideration by those preparing children for free reading. Let's examine what happens as we continue exploring word frequencies as shown in Table 1.

Twenty-five words, including the ten previously discussed, account for less than eight-tenths of one per cent of the different words found; however, they occurred over 37,000 times in all running words. The young reader who had instant recognition of just these twenty-five words would be equipped, on the average, to deal instantly with over one-third of all words in the library books he is most likely to select for free reading.

Forty-seven words, including the twenty-five just discussed, account for less than one and one-half per cent of all the different words found; however, they occurred over 47,000 times in all running words. The young reader who had instant recognition of just these forty-seven words would be equipped, on the average, to deal instantly with nearly forty-five per cent, almost half, of all words in the library books he is most likely to select for free reading.

Ninety-four words, including those previously discussed, account for less than three and one-fourth per cent of all the different words found; however, they occurred nearly 60,000 times in all running words. The young

reader who had instant recognition of just these ninety-four words would be equipped, on the average,to deal instantly with nearly fifty-seven per cent of all words in the library books he is most likely to select for free reading.

One hundred forty-one words, including those just discussed, account for less than four and one-half per cent of all the different words found; however, they occurred over 67,000 times in all running words. The young reader who had instant recognition of just these one hundred forty-one words would be equipped, on the average, to deal instantly with nearly two out of every three words in the library books he is most likely to select for free reading.

One hundred eighty-eight words, including those previously discussed, account for less than six per cent of all the different words found; however, they occurred over 72,000 times in a total of slightly more than 100,000 running words. The young reader who had instant recognition of just these one hundred eighty-eight words would be equipped, on the average,to react instantly to nearly seven out of every ten words in the library books he is most likely to select for free reading.

Let's look now at what some of these high-frequency words are and the number of times that each occurs in the total of eighty books as shown in Table 2.

The ten words that account for nearly one-fourth of all running words are the, more than 6,200 times; to, over 3,000 times; and, nearly 3,000 times; he, over 2,500 times; a, over 2,400 times; I, nearly 1,700 times; and you, it, of, and in, over 1,500 times each.

It is interesting to note that, although there are only about three hundred structure words as opposed to over half a million words that Lefevre

(1964) calls "full words" in the English language, six of the ten words of highest frequency were words that signified structural relationships: the, to, and, a, of, and in. Just three of these six, the, to, and and, account for nearly twelve per cent of all running words in the study. The remaining four in the top ten were all pronouns.

As we have indicated, the highest frequency twenty-five words accounted for more than one-third of all running words. Let's look at these twenty-five words and their individual frequencies. Adding fifteen words to our original ten adds only one verb, said, and two words, was and had, that may serve as verbs or verb makers. Of the remaining twelve, seven are structure words: that, for, on, but, at, with, and up. Four are pronouns: his, she, they, and him. One word, all, would usually be used as a structure word by children although it may be used as a noun in such sentences as, "He gave his all."

To summarize, our top twenty-five contains fourteen words that are used as structure words by children, eight pronouns, one verb, and two words that may serve as verbs or verb makers.

## Limitations

Before comparing these results with the results of other frequency studies and discussing some implications of these findings, let's take cognizance of some possible limitations in the study.

First, the librarians who submitted the original list of books were deliberately picked to provide geographic and community socio-economic representations rather than being randomly selected from a total population of all possible librarians.

Second, from the thousands of title that were submitted, a workable number of eighty books was selected through the opinions of experts: classroom teachers.

These two limitations might have greater importance if the objective of the study was to arrive at a title list to be distributed as books which children usually select from libraries for free reading. However, this was not the final objective of the study and there seems to be no sound reason to believe that the frequencies of specific words in the selections would be markedly different if alternate procedures had been used. Hence, the author believes that the final frequency count is representative of that found in books which primary grade children select from libraries for free reading.

## Comparisons With Other Studies

It may be of some interest to compare the results of this study with the results of some of the other better-known vocabulary studies.

One of the best known is the Thorndike-Lorge (1944) which is a summary of four different counts. As might be expected, most of the words on our list are found in the first five hundred of the Thorndike-Lorge. Words which are on our list of one hundred eighty-eight that only appear in their second five hundred are dog, eat, oh, and ran. The word grand which appears on our list appears in the second thousand of the Thorndike-Lorge.

In the Dolch (1950) list of the two hundred twenty most common words other than nouns and the ninety-five most common nouns, twenty-one, or over eleven per cent, of the words in our count did not appear. The words on our list which did not appear on Dolch's list are began, cry, enough, even, friend, grand, hard, king, next, oh, other, people, place, side,

than, thought, told, took, turn, wait, and while.

Many people have preferred the Rinsland (1945) vocabulary list, perhaps because it provides a greater degree of specificity concerning placement than most others. Rinsland's list was taken from children's writings, not their reading, but it does break down the list according to the first hundred that are most frequent, the second hundred that are most frequent, and so forth. Again, as expected, there was much agreement between our list and Rinsland's, but there was also considerable disagreement. Twenty-nine per cent of the words in our list of one hundred eighty-eight did not appear in Rinsland's top two hundred. In fact, fourteen of the words on our list did not even make Rinsland's top five hundred and four words on our list (began, eye, grand, and king) did not make his top thousand. We can console ourselves with the fact that began, eye, and king did make the top five hundred in the Thorndike-Lorge. Perhaps the assumed high correlation between the words children write and the words they are called upon to read is not quite as high as we have believed. Or, perhaps it is simply that Rinsland's study of children's writings was published twenty-five years ago.

Spache's (1953) formula which is widely used to measure readability of primary level materials uses the Dale list of 769 words as an important part of the computation. Dale's list, in turn, was made from words found on both the International Kindergarten Union List and the first one thousand of the Thorndike Teacher's Word Book. Only two words on our list of one hundred eighty-eight are not found in the list of 769 used by Spache. They are grand and don't. It should be noted that the list Spache uses does not contain contractions; hence, the word don't would not appear and on our list grand was counted when it appeared as a part of grandfather, grandparents, grandson, and so forth.

Stone (1956) disagreed with the list used by Spache and presented a revised list which did include contractions among other substitutions to Spache. Stone's list of 769 words does not include king, which does appear on our list and it does not include grand which also appears on our list. It does, however, include grandfather and grandmother which would have added to the frequency for the word grand on our list.

Let's look at just one more comparison, an analysis of a collection of texts known as the Standard Corpus of Present-Day Edited American English that was used in compiling The American Heritage Dictionary of the English Language (1969). This count which shows the ten words of highest frequency agrees with ours in seven instances: the, and, to, he, a, of, and in. The others in their top ten were was, eleventh on our list; that, fourteenth on our list; and is, twenty-seventh on our list. They noted that the word he occurred almost three and one-half times as often as the word she, which is perhaps something for the ladies to contemplate.

## Implications

As we explore implications, let's first understand that, although this study concentrates on a count of individual words, the important communication in reading takes place through the flow of language with all of its syntactic, semantic,and intonational beauty and relationships -- not through individual, unrelated words. However, normal communication through the flow of printed language requires the pupil to deal with words -- the bricks which build the house of language.

The very high frequency of a relatively few words would seem to imply an advantage to having children decode these words early in their reading experiences. Teachers should then provide extensive practice to

insure that these words become permanently fixed to the level of instant recognition in the child's mind. Beginning with words of highest frequency and then progressing through the remainder would most likely insure earliest independence in free reading for the child together with a growing ability to read more and more advanced materials.

It would not appear, however, that such structuring could simply begin with the highest-frequency word and then continue through the list if we want children to have early reading experiences with stories rather than isolated words. Since there is only one verb and two words that may be verbs or verb makers in the first twenty-five words, it seems desirable to pick from the highest-frequency words and then to go selectively beyond these top words to words like is (the 27th word in frequency), go (35th), not (49th), and will (59th) in order to construct early printed context that will be interesting. Such selectivity at the very beginning of reading instruction will insure that the child has mastered very high-frequency words which he will meet in early free reading at the same time that he is reading interesting content in his developmental reading program.

There is no implication here that children should simply be told what these words are, or that they should be memorized in isolation. It has been proven (Goodman, 1965) that first grade pupils decode much more readily when words appear in context than they do when words appear in lists. Consequently, when introducing the very first high-frequency words, the teacher can supply oral context which the pupil can use in conjunction with his knowledge of grapheme-phoneme correspondences to decode these words for himself. As soon as he has built a sufficient vocabulary of words that are instantly recognizable, he can then meet other new words in printed context.

## Summary

This study reveals which words appear with high frequency in library books that primary grade children select for free reading. A survey of librarians determined which books these children select. This list was reduced to eighty books through evaluations by elementary school teachers. A computer analysis of each word in these books revealed 105,280 running words. When proper names, onomatopoeic words, and easily recognizable inflected forms and compounds were omitted, there were only 3,220 different words in all of these books.

A frequency count of these different words revealed that just ten words account for almost one-fourth of all running words, just twenty-five words account for over one-third of all running words, and just one hundred eighty-eight words account for almost seven out of ten of all running words.

It is suggested that systematic teaching of these high-frequency words would help insure that children will have the background needed to read library materials of their own choosing at an early age.

## References

Dolch, E.W. Teaching Primary Reading. (2nd ed.). Champaign, Illinois: Garrard Press, 1950.

Goodman, K.S. A Linguistic Study of Cues and Miscues in Reading. Elementary English, 1965, 42, 639-43.

Kučera, H. Computers in Language Analysis and in Lexicography. In The American Heritage Dictionary of the English Language. Morris, W. (ed.). Boston: Houghton Mifflin Co., 1969, xxxviii-xl.

Lefevre, C.A. Linguistics and the Teaching of Reading. New York: McGraw-Hill, 1964.

Rinsland, H.D. A Basic Vocabulary of Elementary School Children. New York: Macmillan, 1945.

Spache, G. A Readability Formula for Primary Reading Materials. In Research in the Three R's. Hunnicutt, C.W. and Iverson, W.J. (eds.). New York: Harper and Brothers, 1958. 177-84.

Stone, C.R. Measuring Difficulty of Primary Reading Material: A Constructive Criticism of Spache's Measure. Elementary School Journal, 1956, 57, 36-41.

Thorndike, E.L. and Lorge, I. The Teacher's Word Book of 30,000 Words. New York: Bureau of Publications, Teachers College, 1944.

Table 2. One Hundred Eighty-Eight Words of More Than Eighty-Eight Frequencies in a Computer Study of Vocabulary in Popular Trade Juveniles (Listed in Descending Order of Frequencies)

| | | | |
|---|---|---|---|
| the.....6277 | so..... 375 | got..... 187 | say..... 120 |
| to.....3063 | see..... 371 | take..... 185 | tree..... 119 |
| and.....2916 | not..... 364 | where.... 184 | tell..... 119 |
| he.....2513 | were..... 351 | every.... 182 | school... 119 |
| a.....2451 | get..... 346 | dog..... 182 | still.... 117 |
| I.....1664 | them..... 340 | way..... 181 | much..... 117 |
| you.....1566 | like..... 340 | away..... 181 | keep..... 116 |
| it.....1555 | just..... 336 | man..... 179 | children. 114 |
| of.....1504 | this..... 326 | old..... 178 | give..... 113 |
| in.....1501 | my..... 320 | by..... 177 | work..... 112 |
| was.....1429 | would.... 319 | their.... 176 | king..... 112 |
| said.....1429 | me..... 318 | here..... 175 | first.... 112 |
| his.....1066 | will..... 315 | saw..... 173 | even..... 112 |
| that..... 981 | big..... 315 | call..... 173 | cry..... 112 |
| she..... 820 | mother... 313 | turn..... 172 | try..... 111 |
| for..... 752 | went..... 310 | after.... 172 | new..... 111 |
| on..... 750 | are..... 305 | well..... 170 | must..... 111 |
| they..... 723 | come..... 296 | think.... 169 | grand.... 111 |
| but..... 651 | back..... 293 | ran..... 168 | start.... 107 |
| had..... 631 | if..... 280 | let..... 165 | soon..... 107 |
| at..... 617 | now..... 279 | help..... 165 | made..... 104 |
| him..... 614 | other.... 275 | side..... 159 | run..... 103 |
| with..... 600 | long..... 271 | house.... 158 | hand.....103 |
| up..... 577 | no..... 266 | home..... 155 | began.... 103 |
| all..... 568 | came..... 263 | thought.. 153 | gave..... 102 |
| look..... 564 | ask..... 257 | make..... 149 | friend... 102 |
| is..... 531 | day..... 256 | walk..... 148 | next..... 100 |
| her..... 528 | very..... 253 | water.... 145 | open..... 98 |
| there.... 506 | boy..... 246 | two..... 145 | has..... 98 |
| some..... 490 | an..... 243 | or..... 145 | hard..... 98 |
| out..... 488 | over..... 240 | head..... 141 | enough... 98 |
| as..... 485 | your..... 235 | door..... 140 | wait..... 97 |
| be..... 483 | time..... 234 | before... 139 | Mrs..... 97 |
| have..... 468 | from..... 232 | more..... 137 | morning.. 97 |
| go..... 466 | good..... 228 | eat..... 133 | find..... 97 |
| we..... 455 | any..... 225 | oh..... 132 | only..... 96 |
| one..... 455 | about.... 214 | again.... 132 | us..... 93 |
| then..... 451 | Mr..... 213 | play..... 131 | three.... 93 |
| little... 429 | father... 208 | who..... 129 | our..... 93 |
| down..... 424 | around... 208 | been..... 129 | found.... 93 |
| do..... 402 | want..... 206 | may..... 129 | why..... 92 |
| can..... 392 | don't.... 204 | stop..... 128 | girl..... 91 |
| could.... 386 | how..... 199 | off..... 128 | place.... 90 |
| when..... 385 | know..... 195 | never.... 125 | under.... 90 |
| did..... 378 | right.... 191 | eye..... 122 | while.... 89 |
| what..... 377 | put..... 191 | took..... 121 | told..... 89 |
| thing.... 376 | too..... 190 | people... 121 | than..... 89 |

Table 3. One Hundred Eighty-Eight Words of More Than Eighty-Eight Frequencies in a Computer Study of Vocabulary in Popular Trade Juveniles

(Listed in Alphabetical Order)

| | | | |
|---|---|---|---|
| a.....2451 | friend... 102 | Mr..... 213 | that..... 981 |
| about.... 214 | from..... 232 | Mrs..... 97 | the.....6277 |
| after.... 172 | gave..... 102 | much..... 117 | their.... 176 |
| again.... 132 | get..... 346 | must..... 111 | them..... 340 |
| all..... 568 | girl..... 91 | my..... 320 | then..... 451 |
| an..... 243 | give..... 113 | never.... 125 | there.... 506 |
| and.....2916 | go..... 466 | new..... 111 | they..... 723 |
| any..... 225 | good..... 228 | next..... 100 | thing.... 376 |
| are..... 305 | got..... 187 | no..... 266 | think.... 169 |
| around... 208 | grand.... 111 | not..... 364 | this..... 326 |
| as..... 485 | had..... 631 | now..... 279 | thought.. 153 |
| ask..... 257 | hand..... 103 | of.....1504 | three.... 93 |
| at..... 617 | hard..... 98 | off..... 128 | time..... 234 |
| away..... 181 | has..... 98 | oh..... 132 | to.....3063 |
| back..... 293 | have..... 468 | old..... 178 | told..... 89 |
| be..... 483 | he.....2513 | on..... 750 | too..... 190 |
| been..... 129 | head.....141 | one..... 455 | took..... 121 |
| before... 139 | help..... 165 | only..... 96 | tree..... 119 |
| began.... 103 | her..... 528 | open..... 98 | try..... 111 |
| big..... 315 | here..... 175 | or..... 145 | turn..... 172 |
| boy..... 246 | him..... 614 | other.... 275 | two..... 145 |
| but..... 651 | his.....1066 | our..... 93 | under.... 90 |
| by..... 177 | home..... 155 | out..... 488 | up..... 577 |
| call..... 173 | house.... 158 | over..... 240 | us..... 93 |
| came..... 263 | how..... 199 | people... 121 | very..... 253 |
| can..... 392 | I.....1664 | place.... 90 | wait..... 97 |
| children. 114 | if..... 280 | play..... 131 | walk..... 148 |
| come..... 296 | in.....1501 | put..... 191 | want..... 206 |
| could.... 386 | is..... 531 | ran..... 168 | was.....1429 |
| cry..... 112 | it.....1555 | right.... 191 | water.... 145 |
| day..... 256 | just..... 336 | run..... 103 | way..... 181 |
| did..... 378 | keep..... 116 | said.....1429 | we..... 455 |
| do..... 402 | king..... 112 | saw..... 173 | well..... 170 |
| dog..... 182 | know..... 195 | say..... 120 | went..... 310 |
| don't....204 | let..... 165 | school... 119 | were..... 351 |
| door..... 140 | like..... 340 | see..... 371 | what..... 377 |
| down..... 424 | little... 429 | she..... 820 | when..... 385 |
| eat..... 133 | long..... 271 | side..... 159 | where.... 184 |
| enough... 98 | look..... 564 | so..... 375 | while.... 89 |
| even..... 112 | made..... 104 | some..... 490 | who..... 129 |
| every.... 182 | make..... 149 | soon..... 107 | why..... 92 |
| eye..... 122 | man..... 179 | start.... 107 | will..... 315 |
| father... 208 | may..... 129 | still.... 117 | with..... 600 |
| find..... 97 | me..... 318 | stop..... 128 | work..... 112 |
| first.... 112 | more..... 137 | take..... 185 | would.... 319 |
| for..... 752 | morning.. 97 | tell..... 119 | you.....1566 |
| found.... 93 | mother... 313 | than..... 89 | your..... 235 |