



# Threshold Vocabulary

---



# How many words?

---

- 200,000 words in English
- Educated native English adult knows about 20,000
- We learn about 1,000 a year
- We use about 1,000 in social talk
- English-only kinders from educated homes recognize about 4,000 when they start school (receptive “word bank”)



# ELs and vocab

---

- EL kinder needs to know 4,000 like his peers, plus 1,000 new words
- EL 4th grader needs to know 9,000 like his peers, plus 1,000 new words
- And so on....
- Where to begin?
  - Threshold vocabulary



# High frequency words

---

- Foreign students coming to the US for college
  - minimum 2,000 high frequency
  - plus 550 from the “academic word list”
- General Services List
  - 1st 1000---about 75% of any text
  - 2nd 1000--another 5-10%
- Content words--about 10%

# 1st Grade Expository

	Families	Types	Tokens	Percent	Words in text (tokens):	79
First 500:	...	...	(56)	(70.89%)	Different words (types):	47
<b>K1 Words (1 to 1000):</b>	40	43	<b>75</b>	<b>94.94%</b>	Type-token ratio:	0.59
Function:	...	...	(41)	(51.90%)	(Tokens per type:	1.68)
Content:	...	...	(34)	(43.04%)	Function-content ratio:	0.52
<b>K2 Words (1001 to 2000):</b>	2	2	<b>2</b>	<b>2.53%</b>	Onlist Tokens:	77
<b>AWL Words (academic):</b>	0	0	<b>0</b>	<b>0.00%</b>	Onlist Types:	45
<b>Off-List Words:</b>	?	2	<b>2</b>	<b>2.53%</b>	Onlist Type-Token:	0.58
	42+?	47	79	100%	Onlist Families:	42
					Onlist Family/token:	0.55
					Onlist Family/type:	0.93

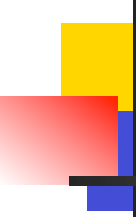
Output text: Today our class is taking a **trip** to the forest We will learn about the plants and trees that grow there We will also learn about the animals that make their home in the forest We must be very careful in the forest One **spark** can start a big forest fire the fires kill the trees and **hurt** the animals We like the forest It is very beautiful and is home for many interesting birds **bugs** animals and plants

# 1st Grade Narrative

	Families	Types	Tokens	Percent	Words in text (tokens):	75
First 500:	...	...	(65)	(86.67%)	Different words (types):	52
<b>K1 Words (1 to 1000):</b>	37	44	<b>67</b>	<b>89.33%</b>	Type-token ratio:	0.69
Function:	...	...	(47)	(62.67%)	(Tokens per type:	1.44)
Content:	...	...	(20)	(26.67%)	Function-content ratio:	0.63
<b>K2 Words (1001 to 2000):</b>	6	6	<b>6</b>	<b>8.00%</b>	Onlist Tokens:	73
<b>AWL Words (academic):</b>	0	0	<b>0</b>	<b>0.00%</b>	Onlist Types:	50
<b>Off-List Words:</b>	?	2	<b>2</b>	<b>2.67%</b>	Onlist Type-Token:	0.68
	43+?	52	75	100%	Onlist Families:	43
					Onlist Family/token:	0.59
					Onlist Family/type:	0.86

Output text: Today it is a bear He looks lost and afraid the tall buildings scare him And he has never seen so many people Do not worry I tell him The buildings will not hurt you and most of the people are friendly How did you get here I ask I climbed in to have a nap he explains and when I woke up I was lost I will help you Tell me where you live

## 2nd Grade Expository



	Families	Types	Tokens	Percent	Words in text (tokens):	114
First 500:	...	...	(89)	(78.07%)	Different words (types):	71
<b>K1 Words (1 to 1000):</b>	54	59	<b>95</b>	<b>83.33%</b>	Type-token ratio:	0.62
					(Tokens per type:	1.61)
Function:	...	...	(58)	(50.88%)	Function-content ratio:	0.51
Content:	...	...	(37)	(32.46%)		
<b>K2 Words (1001 to 2000):</b>	5	5	<b>7</b>	<b>6.14%</b>	Onlist Tokens:	102
					Onlist Types:	64
<b>AWL Words (academic):</b>	0	0	<b>0</b>	<b>0.00%</b>	Onlist Type-Token:	0.63
					Onlist Families:	59
<b>Off-List Words:</b>	?	7	<b>12</b>	<b>10.53%</b>	Onlist Family/token:	0.58
	59+?	71	114	100%	Onlist Family/type:	0.92

Output text: What is it like to have a new brother or sister **Tori** is going to find out soon Her mother is going to have a **baby** Is it going to be a boy or a girl **Tori** wonders Will **Mommy** love me as much when the **baby** comes **Jeremy** and **Jonathan** are **twins** We have each other to play with all the time says **Jeremy** **lan** and **Ryan** are **twins** too Dad says that we even sleep the same way says **Ryan** We like looking so much **alike** that we can **trick** people We are not **exactly** **alike** though says **lan** I am much better at drawing but my brother is better at **sports**

# 3rd Grade Narrative

	Families	Types	Tokens	Percent	Words in text (tokens):	175
First 500:	...	...	(118)	(67.43%)	Different words (types):	126
<b>K1 Words (1 to 1000):</b>	87	99	<b>143</b>	<b>81.71%</b>	Type-token ratio:	0.72
Function:	...	...	(78)	(44.57%)	(Tokens per type:	1.39)
Content:	...	...	(65)	(37.14%)	Function-content ratio:	0.45
<b>K2 Words (1001 to 2000):</b>	12	12	<b>17</b>	<b>9.71%</b>	Onlist Tokens:	164
<b>AWL Words (academic):</b>	4	4	<b>4</b>	<b>2.29%</b>	Onlist Types:	115
<b>Off-List Words:</b>	?	11	<b>11</b>	<b>6.29%</b>	Onlist Type-Token:	0.70
	103+?	126	175	100%	Onlist Families:	103
					Onlist Family/token:	0.63
					Onlist Family/type:	0.90

Output text: We stayed all **afternoon** waiting **anxiously** hoping to hear from the **rescue** organization We got to know one another better and we learned more about the **deer** **Peach** eyes were wide and bright Look how they **rotate** their big soft ears to the left and right she **exclaimed** **Clarence** said We studied **deer** in science Their hearing is very **sharp** and it helps them **detect** enemies **approaching** from far away Mr **Benny** **nodded** as he walked over to us I sometimes see this kind of **deer** at night in the **headlights** when I drive way past the city limits When they are **startled** by the **taxi** lights their **tails** go up like **flags** the **tails** are white underneath which means the animals are white **tailed** **deer** the **deer** **grazed** and slept **cautiously** always **alert** to danger they watched us with **curious** intelligent eyes I could see that the people made them **uncomfortable** and it helped me **appreciate** that these really were wild animals We tried to keep our distance and not make any **sudden** movements



## 4th Grade Expository

	Families	Types	Tokens	Percent	Words in text (tokens):	162
First 500:	...	...	(105)	(64.81%)	Different words (types):	108
<b>K1 Words (1 to 1000):</b>	81	87	141	87.04%	Type-token ratio:	0.67
					(Tokens per type:	1.50)
Function:	...	...	(70)	(43.21%)	Function-content ratio:	0.43
Content:	...	...	(71)	(43.83%)		
<b>K2 Words (1001 to 2000):</b>	11	11	11	6.79%	Onlist Tokens:	155
					Onlist Types:	101
<b>AWL Words (academic):</b>	3	3	3	1.85%	Onlist Type-Token:	0.65
					Onlist Families:	95
<b>Off-List Words:</b>	?	7	7	4.32%	Onlist Family/token:	0.61
	95+?	108	162	100%	Onlist Family/type:	0.94

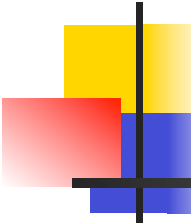
Output text: Heat waves can last for a few weeks or for as long as several weeks They are caused when large masses of hot air get **stuck** over a particular region and do not **budge** these hot **spells** can send the **thermometer** to over number degrees such hot **weather** can be dangerous even **deadly** People animals and **crops** can all be seriously **affected** for example too much heat can make you **sick** by forcing your body to work **extra** hard in the story you are about to read the **Kansas** farm girl tries to save her farm from a heat wave that may just prove to be too hot to **handle** **beating** the heat take it easy **avoid** working or playing too hard stay out of the sun or wear **sunscreen** wear light colors they **bounce** some of the sun **energy** away from you even if you are not **thirsty** drink **lots** of water your body needs water to **cool** down eat lightly



	Families	Types	Tokens	Percent	Words in text (tokens):	247
First 500:	...	...	(164)	(66.40%)	Different words (types):	131
<b>K1 Words (1 to 1000):</b>	79	91	<b>191</b>	<b>77.33%</b>	Type-token ratio:	0.53
Function:	...	...	(114)	(46.15%)	(Tokens per type:	1.89)
Content:	...	...	(77)	(31.17%)	Function-content ratio:	0.46
<b>K2 Words (1001 to 2000):</b>	11	11	<b>13</b>	<b>5.26%</b>	Onlist Tokens:	205
<b>AWL Words (academic):</b>	1	1	<b>1</b>	<b>0.40%</b>	Onlist Types:	103
<b>Off-List Words:</b>	?	28	<b>42</b>	<b>17.00%</b>	Onlist Type-Token:	0.50
	91+?	131	247	100%	Onlist Families:	91
					Onlist Family/token:	0.44
					Onlist Family/type:	0.88

Output text: The surest way to **popularize** an idea in song is to set new words to a familiar **melody** As early as number there was mention in **Philadelphia** of a **comic** song called **Yankee Doodle** When the word **yankee** first appeared in **print** people were not quite sure what it meant To this day there is some **confusion** about its **origin** Some people believe it comes from an **Indian** word Others think it is based on a **French** word The strongest possibility is that it comes from the **Dutch** name for the English colonists Jan **Kaas** or Jan **Kees** Jan **yan** is **Dutch** for **John Kees** means **cheese** **John Cheese** was not meant as a **compliment** Neither was **Doodle** which means **fool** **Yankee Doodle** first appeared in **print** in a **London broadside** in number its **subtitle** was The **Lexington** March The **British band** played it on the march to **Lexington** in those days **European** armies played **loud** music on the way into battle It **cheered** up the soldiers and gave them **courage** In this case the **strains** of the music let the **Minutemen** know **exactly** where the **British** were The **Minutemen** also realized that the **British** were trying to make **fun** of them by calling them **Yankee Doodles** In the true spirit of the times the familiar **melody** was taken up by the **Americans** with new words by a **Harvard** College student **Edward Bangs** and sung right back to them It is this **version** of **Yankee Doodle** that

## 5th Grade Expository



	Families	Types	Tokens	Percent	Words in text (tokens):	381
First 500:	...	...	(266)	(69.82%)	Different words (types):	191
<b>K1 Words (1 to 1000):</b>	120	143	<b>308</b>	<b>80.84%</b>	Type-token ratio:	0.50
Function:	...	...	(181)	(47.51%)	(Tokens per type:	1.99)
Content:	...	...	(127)	(33.33%)	Function-content ratio:	0.48
<b>K2 Words (1001 to 2000):</b>	14	15	<b>22</b>	<b>5.77%</b>	Onlist Tokens:	337
<b>AWL Words (academic):</b>	6	6	<b>7</b>	<b>1.84%</b>	Onlist Types:	164
<b>Off-List Words:</b>	?	27	<b>44</b>	<b>11.55%</b>	Onlist Type-Token:	0.49
	140+?	191	381	100%	Onlist Families:	140
					Onlist Family/token:	0.42
					Onlist Family/type:	0.85

Output text: Because Gary Soto had few books around when he was growing up and little encouragement to read he was not planning to become an author Instead he thought about being a priest a barber or maybe a scientist who studies dinosaurs In fact he went to college to become a geographer But when he found a book of poetry in the college library he was drawn to it Reading the poetry inside changed his life and inspired him to become an author himself Soto writes poetry novels and stories for adults and children He writes about what he knows mainly what it is like to be a Mexican American boy growing up poor in central California Soto often writes about enjoyable childhood experiences but his life was also filled with hardship He was born on April number number in Fresno which is in the heavily agricultural San Joaquin Valley His grandparents were born in Mexico Both his parents and grandparents worked as laborers in the fields and factories around Fresno They picked oranges grapes and cotton Soto himself grew up doing physical labor He picked grapes washed cars mowed lawns and painted house numbers on curbs When Soto was five he moved with his family to a Mexican American neighborhood or barrio on the edge of Fresno Tragedy struck the family not long after their move however Soto father was killed in an accident at the factory where he worked he was only twenty seven years old Soto mother had to raise her children alone

## Holt 7th Author Bio

	Families	Types	Tokens	Percent	Words in text (tokens):	428
First 500:	...	...	(263)	(61.45%)	Different words (types):	223
<b>K1 Words (1 to 1000):</b>	120	133	<b>316</b>	<b>73.83%</b>	Type-token ratio:	0.52
Function:	...	...	(198)	(46.26%)	(Tokens per type:	1.92)
Content:	...	...	(118)	(27.57%)	Function-content ratio:	0.46
<b>K2 Words (1001 to 2000):</b>	19	20	<b>25</b>	<b>5.84%</b>	Onlist Tokens:	354
<b>AWL Words (academic):</b>	13	13	<b>13</b>	<b>3.04%</b>	Onlist Types:	166
<b>Off-List Words:</b>	?	57	<b>74</b>	<b>17.29%</b>	Onlist Type-Token:	0.47
	152+?	223	428	100%	Onlist Families:	152
					Onlist Family/token:	0.43
					Onlist Family/type:	0.92

## Holt 9th Author Bio

Output text: **Saki** stories are notable for their **biting satire wit** and occasional moments of **horror** **Hector Hugh Munro** was born in **Burma** the third child of a **British** army officer When **Hector** was two his mother was killed by a **runaway cow** and the three children were taken back to **England** to live with their **paternal grandmother** and their two **aunts** **Munro** was raised in a large country **house—Broadgate Villa** in **Devonshire—that** included servants a **governess** gardens and access to the sea but **Munro recollections** of the time were of **frequent** illnesses and the **nattering constraints** of his **aunts** As a child he **invented elaborate fantasies** in which he **punished** his **aunts** for their **oppressive treatment** In one such **fantasy** a **wolf devours** his **aunts** This **wolf** appears in a different form in his stories along with his **aunts** who show up in various **unflattering guises** **Munro** did not receive much **formal education** his **aunts** thought his **health** was **fragile** but he was happy at boarding school in spite of its **discipline** He enjoyed writing in particular and was singled out for his **fastidiousness** and **facility** with language Nevertheless he was to look back on his boarding school days with growing **skepticism** about **traditional preparatory** schools and the values of the **British upper** class they **perpetuated** When **Munro** was twenty three his father **arranged** for him to return to **Burma** and take a post with the **military police** He went reluctantly and after three years of **policing** and a **bout** with **malaria** **Munro**

- 
- 
- Website for GSL word analysis

[www.er.uqam.ca/nobel/r21270/textool/  
web\\_vp.html](http://www.er.uqam.ca/nobel/r21270/textool/web_vp.html)



# Which words, which grade?

---

- Doesn't matter: 6-year old and 16-year old both need to recognize the threshold vocabulary (GSL 1,000)
- Houghton Mifflin covers about 650 of the 1,000 over 6 years, but still....
- Kids need to know the 1,000 within a year of school entry
- $1,000 / 36 \text{ wks} = 28 \text{ words a week}$



# Knowing

---

- Think of a 4-year old in your home....  
he hears words, and answers questions that contain the words...
- He can't yet retrieve the word to use it.
- He can't read it.
- But, when he does learn to decode the word, he'll already have it in his  
“receptive word bank”





## 28 words from the GSL list

---

- a, ability, able, above, abroad, absent, accept, accident, accuse, ache, act, across, actual, admire, adult, advantage, adventure, advice, afford, afraid, afternoon, again, agency, agree, airplane, allow, already, altogether





## Week #1 (GSL 1-500, random)

---

- at, better break, cause, control, dance general, give, history, hope, moment, more, near, need, open, probable, report, send, several, sure, talk, those, toward, treat, type, view, watch, work



## Week #19 (GSL 501-1000, random)

---

- animal, cent, check, cross, director, district familiar, hang, heat, language, length, medical, motor, opportunity, pattern, please, population, recommend, red, reduce, secretary, settle, share, song, stick, suppose, surprise, title



# What to do with the lists?

---

- Use the 28 words deliberately
- Have 3-4 standard activities
  - Make sure kids hear words in context several times; recognize their use of words
  - For older kids, use reading/writing of words
  - For second language kids, use translation
  - Include variants (plurals, tenses, adjectives, adverbs), synonyms, non-examples, collocations
  - For kids with strong receptive vocabularies, use higher level activities