

Researching Vocabulary acquisition in Extensive Reading

Rob Waring

AsiaTEFL Ha Noi August 7th 2010

Essential questions

- What kinds of vocabulary do they need to learn?
- How much language do students need to meet?
- How many times do they need to meet words to learn them?
- How much will an Extensive Reading program add to their vocabulary growth?

Types of vocabulary

Individual words: *book, table, life, chance, walk, airplane...*
 Affixes: *used, user, usefulness, user-friendly, disuse...*
 Multi-part words: *traffic jam, the day after tomorrow, lunch box...*
 Lexical phrases: *by the way, to and fro, a kind of,...*
 Idioms: *let the cat out of the bag, raining cats and dogs*
 Sentence heads: *Do you mind if I..., If I were you,.. Could you...?*
 Collocations: *High season, mild cheese, blonde hair...*
 Colligations: *agree to do x, agree on X, rely on someone, have an effect on x, x affects y...*
 Others: *SONY, Paul, twenty-seven, etc. , UNESCO...*

How much to learn: vocabulary

Learners need 8000-9000 words to know 98% of the vocabulary in native novels, magazines and most general reading
 Advanced learners will need 4000-6000 words
 An average high frequency word has about 8-15 common collocations
 There are 400-500 common phrasal verbs
 There are 1000-1500 common idioms
 There are hundreds of common sentences heads and formulaic phrases

Intentional vs. Incidental learning

Intentional learning		Incidental learning
Direct focus on learning when the focus is to learn words	FOCUS	Learning 'by accident' - as a result of focusing on something else
wordlists, word cards, vocabulary exercises, dictionary use	E.G.	from reading or listening, watching movies, listening to songs, casual conversation
<ul style="list-style-type: none"> •Can be learnt systematically •Meanings are learnt 16 times faster than with incidental learning •Retention high if learnt well •Decontextualized or 'local' learning level 	LEARNING	<ul style="list-style-type: none"> •Slow and fragile learning •Input tends to be random and unpredictable, unsystematized •Contextualized (chances for integrative learning)
Best for 'form-meaning' level learning	USE	Best for 'deeper aspects' of vocabulary learning

Occurrence rates

Frequency	Word	No.	% of corpus	No. of words to read to meet it x times				
				1	5	10	20	50
1st	BE	10,387	4.989%	20	100	200	401	1,002
2nd	THE	10,027	4.816%	21	104	208	415	1,038
Top 25	SOME	1,192	0.573%	175	873	1,747	3,493	8,733
Top 50	SENTENCE	606	0.291%	344	1,718	3,436	6,871	17,178
Top 100	OUT	323	0.155%	645	3,223	6,446	12,892	32,230
Top 500	PREDICT	50	0.024%	4,164	20,820	41,641	83,281	208,203
Top 1000	ORGANIC	18	0.009%	11,567	57,834	115,668	231,337	578,342
Top 1500	TIMETABLE	9	0.004%	23,134	115,668	231,337	462,673	1,156,683
Top 2000	COMMERCE	4	0.002%	52,051	260,254	520,507	1,041,015	2,602,538
Top 2500	BIOCHEMICAL	2	0.001%	104,102	520,508	1,041,015	2,082,030	5,205,075
Top 3000	REFUND	2	0.001%	104,102	520,508	1,041,015	2,082,030	5,205,075

Opportunity

In EFL environments:

- Natural opportunity is low
- Natural opportunity takes time, effort and commitment
- Opportunities have to be sought
- For many, creating language opportunities can only be done realistically through course requirements

Exposure

Number of words in a typical 5 level course by occurrence rate

	50+	30-49	20-29	10-19	5-9	1-4	Total
Function	84671						
1k	92390	6465	3986	2021	309		105171
2k	1942	1291	1434	4302	3414	189	12572
Other			42	175	160	2496	2873
Total	179003	7756	5462	6498	3883	2685	205287
Names	2106	616	812	1527	1131	953	7145

Function	41.25%						
1k	45.01%	3.15%	1.94%	0.98%	0.15%	0.00%	51.23%
2k	0.95%	0.63%	0.70%	2.10%	1.66%	0.09%	6.12%
Other	0.00%	0.00%	0.02%	0.09%	0.08%	1.22%	1.40%
Total	87.20%	3.78%	2.66%	3.17%	1.89%	1.31%	100.00%

Data from Sequences by Heinle Cengage

Exposure

Different words (types) in a typical 5 level course by occurrence rate

	50+	30-49	20-29	10-19	5-9	1-4	Total
Function	40						40
1k	434	167	163	131	42		937
2k	22	35	60	322	506	52	997
Other			2	13	27	1263	1305
Total	456	202	225	466	575	1315	3239
Names	14	16	34	109	177	367	717

Function	1.23%						
1k	13.40%	5.16%	5.03%	4.04%	1.30%	0.00%	28.93%
2k	0.68%	1.08%	1.85%	9.94%	15.62%	1.61%	30.78%
Other	0.00%	0.00%	0.06%	0.40%	0.83%	38.99%	40.29%
Total	15.31%	6.24%	6.95%	14.39%	17.75%	40.60%	100.00%

Uptake

What uptake can we expect from a typical course?

How much text do learners meet?

-Possibly 200-300,000 words over a typical 5 level course.

How frequently do learners meet words?

-Function words - very very frequently

-De-lexical words - (*have, be, do* etc.) - very very frequently

-800-900 words (25-27%) are met enough times for acquisition

-400-600 (10-15%) words will be partially known

-1300 words (over 65-70%) will probably not be learnt

Cautions

'Acquisition' assumes meeting the words enough times over 3 years is sufficient

The data do not include multiple meanings, collocations, idioms, multiple-meaning senses, multi-part words, grammar, etc.

The above data are for possible uptake rates from a typical course (individual results will vary)

The data are calculated on input frequency (receptive vocabulary)

Productive vocabulary size is typically 1/4 to 1/5th of the receptive. We can expect a productive vocabulary of 200-300 easy-to-access words and another 200-300 partial-access

What to do?

Courses in general tend not to recycle vocabulary enough to allow for deep acquisition

Most courses are linear in design – always teaching new things in each unit / lesson

Learners need a good balance of intentional vs. incidental learning

Start with lots of words, phrases, lexical chunks, sentence heads first. Focus on communicability. Grammar later

Continue the coursework to provide the framework and initial knowledge of words, grammar etc.

Massive text exposure to build incidental learning

Course book plus Extensive Reading

Uptake if they add 1 graded reader per week

	50+	30-49	20-29	10-19	5-9	1-4	
1k	707	93	44	68	47	32	
2k	223	107	90	162	125	168	
Other	93	83	116	309	398	1125	
Total	1023	283	250	539	570	1325	3990

	50+	30-49	20-29	10-19	5-9	1-4	
1k	17.72%	2.33%	1.10%	1.70%	1.18%	0.80%	
2k	5.59%	2.68%	2.26%	4.06%	3.13%	4.21%	
Other	2.33%	2.08%	2.91%	7.74%	9.97%	28.20%	
Total	25.64%	7.09%	6.27%	13.51%	14.29%	33.21%	100.00%

Take home:

We can learn lots of words at the form-meaning level quickly

Initial word knowledge decays quickly unless learnt well

Deeper aspects of word knowledge grow over time from exposure

We can expect students exposed to normal levels of input from a typical course to acquire less than a thousand words receptively and a few hundred productively

Adding a Extensive Reading almost doubles their vocabulary learning

Course book plus Extensive Reading

Significant improvement in 'learnt' vocabulary (800 --->1600 words)

More of the words in their course book reach the 'acquisition' level (27% ---> 40%)

They will have a better sense of how the vocabulary and grammar fit together

They will have a better sense of collocation, and other deeper aspects of vocabulary acquisition as well as picking up phrases and so forth.

Thanks for listening

This presentation is available from
www.robwaring.org/presentations/