

Free Vocabulary Tests and Individualized Flashcards

lexicaTM word engine

Dr. Charles Browne
Professor of Linguistics
Meiji Gakuin University
browne@gol.com

Guy Cihi
President
Lexxica
gcihi@lexxica.co.jp

Outline of Today's Session

1. Review some key facts about vocabulary & reading comprehension.
2. Look at solutions for testing and teaching high frequency vocabulary.
3. Try some of the beta software in small groups.

Coverage

(Lexical comprehension)

Coverage describes the average percentage of known words in running text.

Unless you know 'huia', you have only 96 percent coverage of this slide.

(24 known / 25 total = 96% known)

Key coverage thresholds

Below **80 percent coverage**, reading comprehension is almost impossible
(Hu & Nation, 2001)

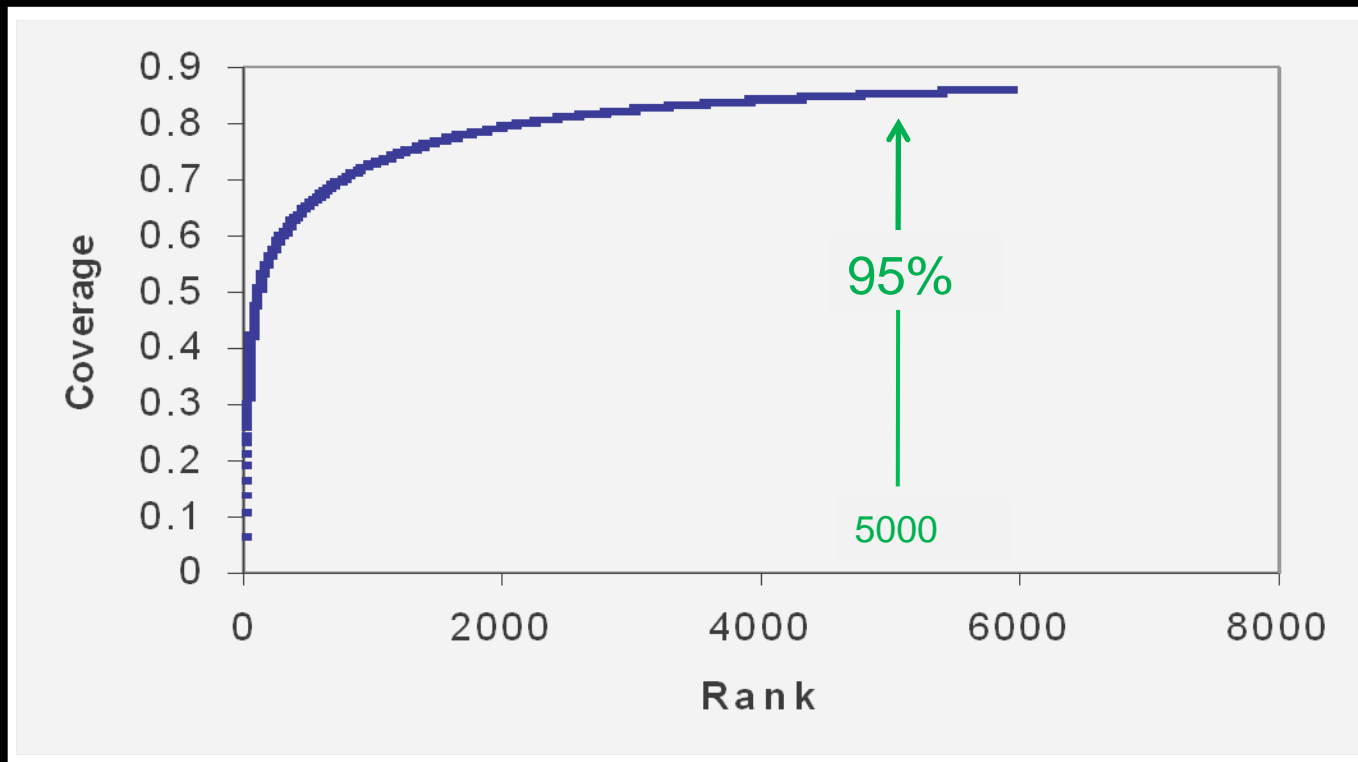
At **95 percent coverage**, it becomes possible to read without the help of dictionaries
(Laufer, 1989)

The relationship between coverage and high frequency vocabulary

<u>HF Words</u>	<u>Coverage</u>	<u>Research</u>
1	7%	West (53), Nation (90)
100	50%	West (53), Nation (90)
1000	75%	West (53), Engles (68)
2000	85%	West (53), Nation (90)
5000	95%	Hirsch & Nation (92)

British National Corpus coverage by the most frequent BNC words

(Leech, Rayson, & Wilson, 2001)



Coverage in action

Let's look at three versions of the exact same text about forestry.

The three slides simulate: 75, 85, and 95 percent coverage.

Only the first 1000 high frequency words.

Simulates 75% Coverage

If _____ planting rates are _____ with planting
_____ in each _____ and the forests
_____ at the earliest opportunity, the _____ wood
supplies could further increase to about 36 million
_____ meters _____ in the _____ 2001-
2015. The additional _____ wood supply should
greatly _____, even if much is
used for _____ production.

15 of 58 words are missing

Only the first 2000 high frequency words.

Simulates 85% Coverage

If _____* planting rates are maintained with planting targets satisfied in each _____* and the forests milled at the earliest opportunity, the _____* wood supplies could further increase to about 36 million _____* meters _____* in the period 2001-2015. The additional _____* wood supply should greatly exceed _____* _____*, even if much is used for _____* production.

9 of 58 words are missing

The first 2000 words, plus 836 words from the University Words List.

Simulates 95% Coverage

If current planting rates are maintained with planting targets satisfied in each region and the forests milled at the earliest opportunity, the available wood supplies could further increase to about 36 million _____* meters annually in the period 2001-2015. The additional available wood supply should greatly exceed domestic requirements, even if much is used for energy production.

1 of 58 words is missing

Key coverage thresholds - again!

Below **80 percent coverage**, reading comprehension is almost impossible.
(Hu & Nation, 2001)

At **95 percent coverage**, it becomes possible to read without the help of dictionaries.
(Laufer, 1989)

How many high frequency words do second language learners know?

Country	Vocab. Size	Hours of Instruction	Reference
Japan (University)	2200	800-1200	Shillaw (95), Barrow (99)
China (English Majors)	4000	1800-2400	Laufer (99)
Indonesia (University)	1220	900	Nurweni & Read (99)
France (HS students)	1000	400	Arnaud et al (85)
Germany (age 15, HS)	1200	400	Milton & Meara (98)

Why the discrepancy?

Shillaw (95) and Barrow (99) found that Japanese university students know approximately 2200 high frequency words whereas Lexxica's V-Check test scores show that Japanese university students know closer to 3800 total words. The discrepancy is likely because Shillaw and Barrow's research focused on testing for knowledge of high frequency vocabulary whereas V-Check is designed to identify all vocabulary knowledge. Research on the content of Japanese high school textbooks reveals a heavy emphasis on low frequency vocabulary. Japanese students knowledge of low frequency vocabulary accounts for the discrepancy.

Japanese high school textbooks coverage by high frequency words.

Popular Textbooks	Coverage obtained from 2000 high frequency words
Spectrum	71%
Milestone	76%
Unicorn	78%
Native-speaker texts	87%

The chronically low percentages indicate that too many low frequency vocabulary words are being used.

University entrance exams coverage by high frequency words.

Entrance Examinations	Coverage obtained from 2000 high frequency words
Tokyo Univ.	79%
Kyoto Univ.	77%
Waseda Univ.	72%
Sophia Univ.	72%
Keio Univ.	69%
Nagoya Univ.	68%

Again, the chronically low percentages indicate that too many low frequency vocabulary words are being used.

Here are just a few examples of low frequency words that are part of the standard Japanese curriculum:

- Anemone
- syncopate
- cockcrow
- neigh
- insularism
- cataclysm
- seedle
- prodigious
- whin
- pichblende
- mazurka
- polonaise

Diagnosis

Japanese students learn many low frequency words that native speakers do not use.

Japanese students are not meeting and learning nearly enough high frequency vocabulary.

Our recommendations

Make learning the 2000 high frequency words each student's highest priority.

Encourage students to learn as many of the 5000 high frequency words as possible.

Help students quickly learn missing high frequency vocabulary by:

1. Identifying *which* high frequency words are not known
2. Utilizing a spaced repetition learning system

Our solutions

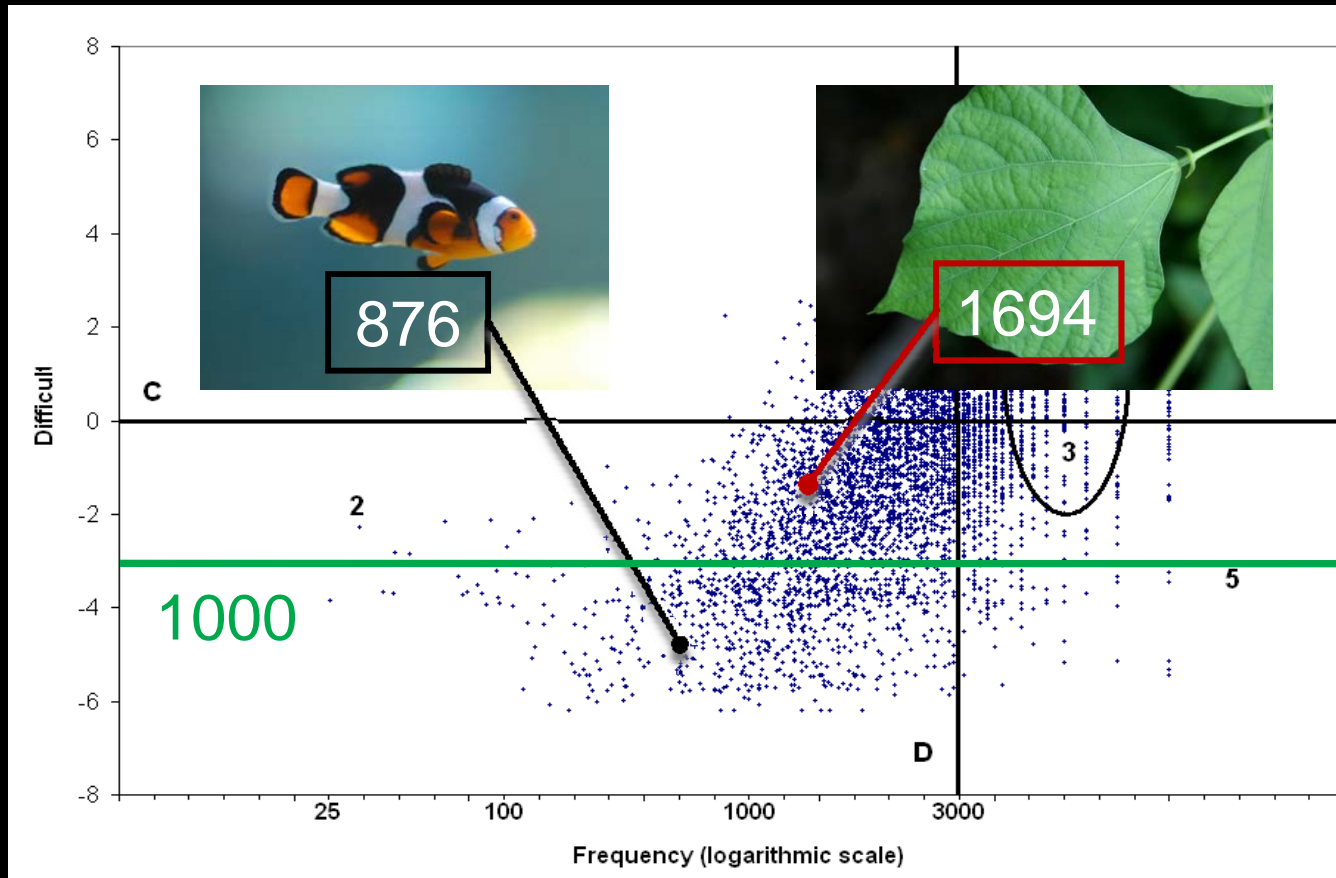
Corpus analysis to determine a valid lexicon of high frequency vocabulary

Lexical assessments to determine known/unknown vocabulary

Flashcards and games to help students rapidly learn their unknown high frequency vocabulary

Extensive graded reading to help students organize and control their new vocabulary

Lexxica works with a proprietary 2.2 billion word web-derived corpus to determine token frequencies



The V-Check test presents learners with simple lexical decision tasks.



legacy

Do you know this word?

yes no

The image shows a digital interface for a V-Check test. It features a central yellow rounded rectangle containing the word "legacy". Below this rectangle, the question "Do you know this word?" is displayed in white text on a blue background. At the bottom, there are two blue rounded buttons with white text: "yes" on the left and "no" on the right.

V-Check identifies which words are known, and which words need to be learned.

SCORE SHEET

User name: demo

Test date: 2007-05-18

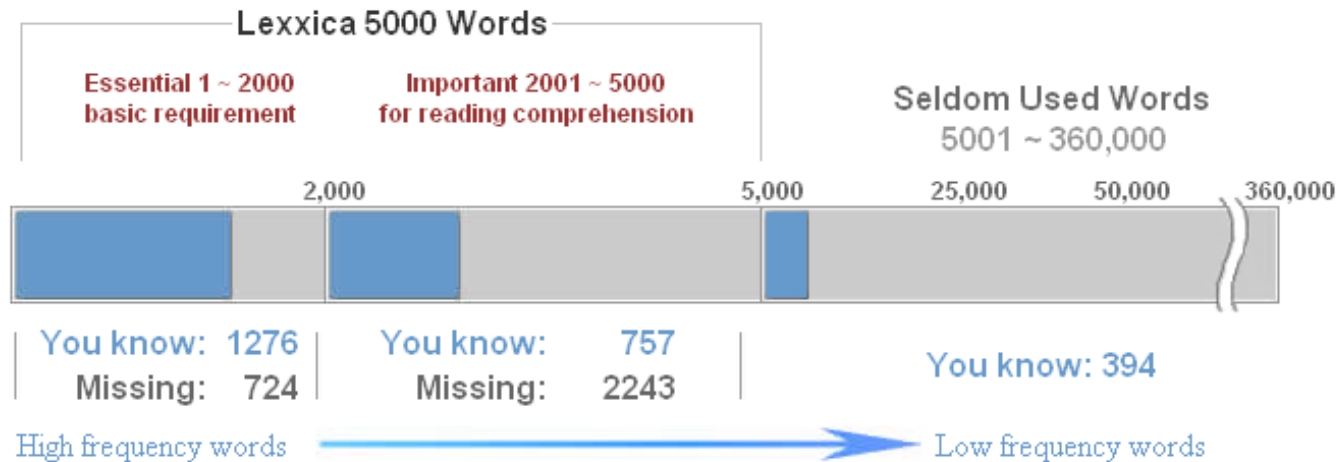


Part 1

Your Vocabulary Size: **2,427** This is the total number of words that you know.

Your Lexica 5000 Words: **2,033** This is the number of Lexica 5000 words that you know.

Your Missing Words: **2,967** This is the number of Lexica 5000 words that you do not know. You can quickly learn these words with V-Flash games and flashcards.



Flashcards and Games

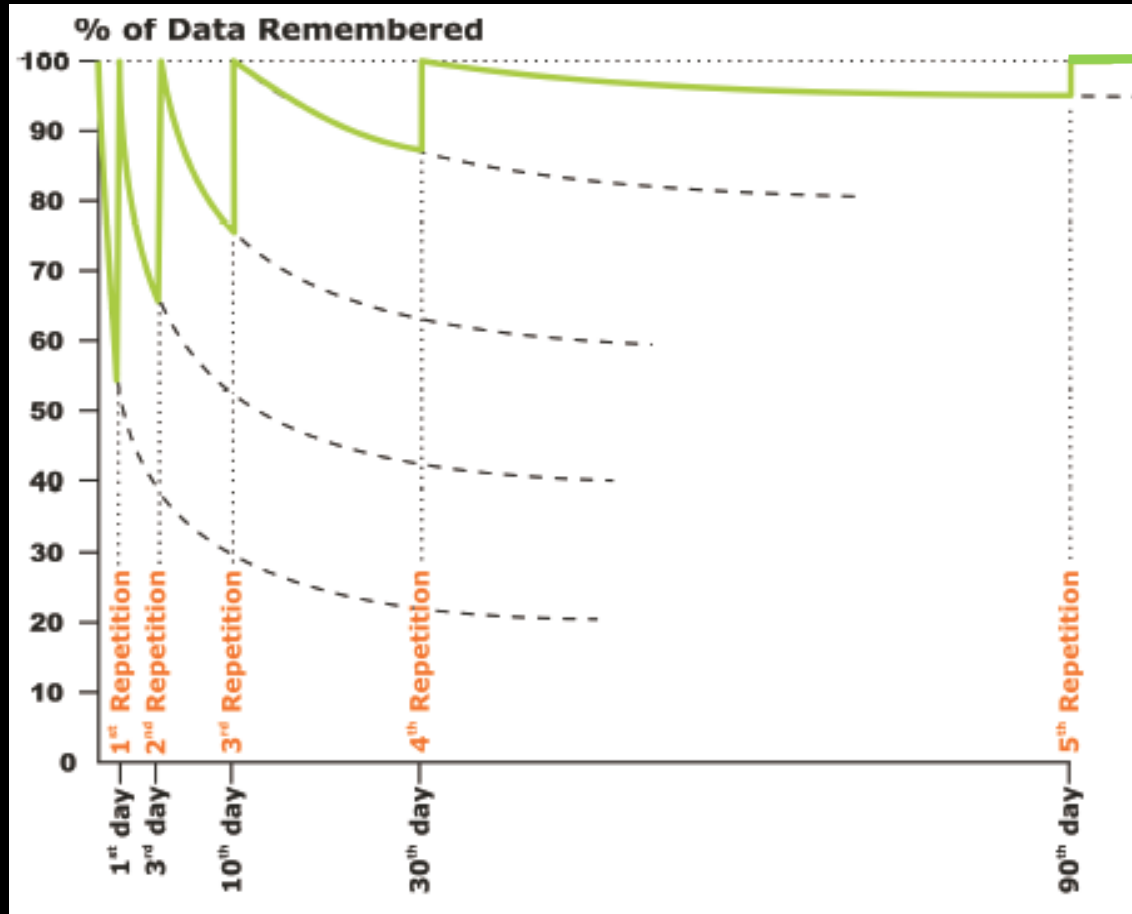
Flashcards and games deliver a personalized target list of high frequency words for each learner.

Our flashcards and games use an advanced spaced repetition system.

Words are repeated at increasing time intervals until the student demonstrates long-term retention.

Spaced repetition

Ebbinghaus (1885), Leitner (1972), Pimsleur (1967), Mondria, (1994)



Long-term memory

Flashcards teach comprehension



thermal
熱の

collocation
thermal springs, thermal insulation, thermal underwear, thermal properties

Did you know it?

yes no

The image shows a digital flashcard interface. The main card is light blue with a white border and rounded corners. It displays the word 'thermal' in a large, bold, dark blue font. Below it is the Japanese characters '熱の' (Atsu no) in a smaller, dark blue font. Underneath that, the word 'collocation' is written in a smaller, dark blue font, followed by a list of examples: 'thermal springs, thermal insulation, thermal underwear, thermal properties'. At the bottom of the card is a small, rounded rectangular button with a white circle and two dots inside, likely for audio playback. Below the card, on a darker blue background, is the question 'Did you know it?' in white text. At the very bottom are two rounded rectangular buttons, one labeled 'yes' and one labeled 'no', both in white text.

SightWords! teaches visual automaticity

The screenshot shows the game interface for the word "contract". At the top left, there are two boxes: "Level 1" and "Lives 10". In the center, the word "contract" is displayed in a large, white, sans-serif font. To the right, a "Time 22" box is visible. Below the word, there are three definition boxes: "legal agreement" (orange), "animal with six legs" (white), and "a viewpoint" (white). The background is a blue sky with a white cloud on the left and a brown ground area at the bottom.

Level 1 Lives 10

contract

Time 22

legal agreement

animal with six legs

a viewpoint

SoundBubbles! teaches aural automaticity



WordCheck! teaches visual automaticity

SoundCheck! teaches aural automaticity

POINTS **0** *Word Check!* ITEMS **1** of 30

perpetual

- 1 小数の
- 2 まじめな
- 3 永遠に続く

* QUIT

POINTS **0** *Sound Check!* ITEMS **1** of 30



1 小数の

2 まじめな

3 永遠に続く

* QUIT

Graded reading helps improve comprehension

Graded podcasts help improve listening

Known Words ▶ 500 Words 1000 Words 2000 Words 5000+ Words

Stomp The Yard by Alan Smithee

Stomp the Yard is a movie about **friendship**, love and dancing. In this movie, a young street dancer named DJ joins a College and finds he is trapped between two different groups of students. A traditional African American dance called "stepping" is at the center of their struggle. One of the groups wants to use DJ to help them win the school's big dance contest. DJ does not really like **stepping** and he is not very **interested** in social life at his new college. DJ likes his own kind of hip hop street dancing. But in order to win a **beautiful** girl's heart, DJ de-

interested
US having an interest in something; concerned;
JP 興味のある
more

... groups and enter the big stepping dance con-
... the some new things about dancing. At the
... something about the true meaning of friendship



Did you read it? Yes, I read it You have now read 34,902 words !

Buy the CD soundtrack to "Stomp the Yard" at Amazon.co.jp!

Buy the CD soundtrack to "Stomp the Yard" at Amazon.co.jp!

Buy the "Stomp the Yard" DVD at Amazon.co.jp!

Send this article to your friends

Read this article in your email.

Read this article on your mobile phone via email, SMS or mobile website.

Conclusion

Lexxica is developing free supplemental materials designed to quickly teach each student their missing high frequency general purpose vocabulary, and to promote extensive reading and listening.

Lexxica's supplemental programs are focused and effective because they were designed by top vocabulary and extensive reading experts who are also experienced classroom teachers.

Contact:

Dr. Charles Browne
Professor of Linguistics
Meiji Gakuin University
browne@gol.com

Guy Cih
President
Lexxica
gcihi@lexxica.co.jp