A Close Examination of Vocabulary in Japanese EFL Textbooks

Shusaku Nakayama, Meiji Gakuin University

JALT
THE JAPAN ASSOCIATION FOR LANGUAGE TEACHING
全国語学教育学会

REFLECTIONS AND JALT2021 NEW PERSPECTIVES
ONLINE
NOVEMBER 12~15, 2021

Contents

1. **Introduction**

2. **Literature Review**

3. **Research Question**

4. **Methods**

5. **Results**

6. **Discussion & Conclusion**

7. **Suggestions for Further Research**

References

MEIJI GAKUIN UNIVERSITY

MEIJI GAKUIN UNIVERSITY

2. Literature Review

Vocabulary has been long acknowledged its essentialness in language learning (Alqahtani, 2015; Schmitt, Schmitt, & Clapham, 2001; Zimmerman, 1997).

▪95% of the running words in a text should be known for reading a text (Laufer, 1989; Liu & Nation, 1985) and guessing from context (Liu & Nation, 1985).

▪98%-99% of the running words in a text should be known for reading a text for pleasure (Hu & Nation, 2000).

MEIJI GAKUIN UNIVERSITY

Research into Vocabulary Occurring in Japanese EFL Textbooks

Chujo (2004): around the most frequent **3,000** word families ← SH

Wongsarnpigoon (2018): the most frequent **2,000** word families ← JH
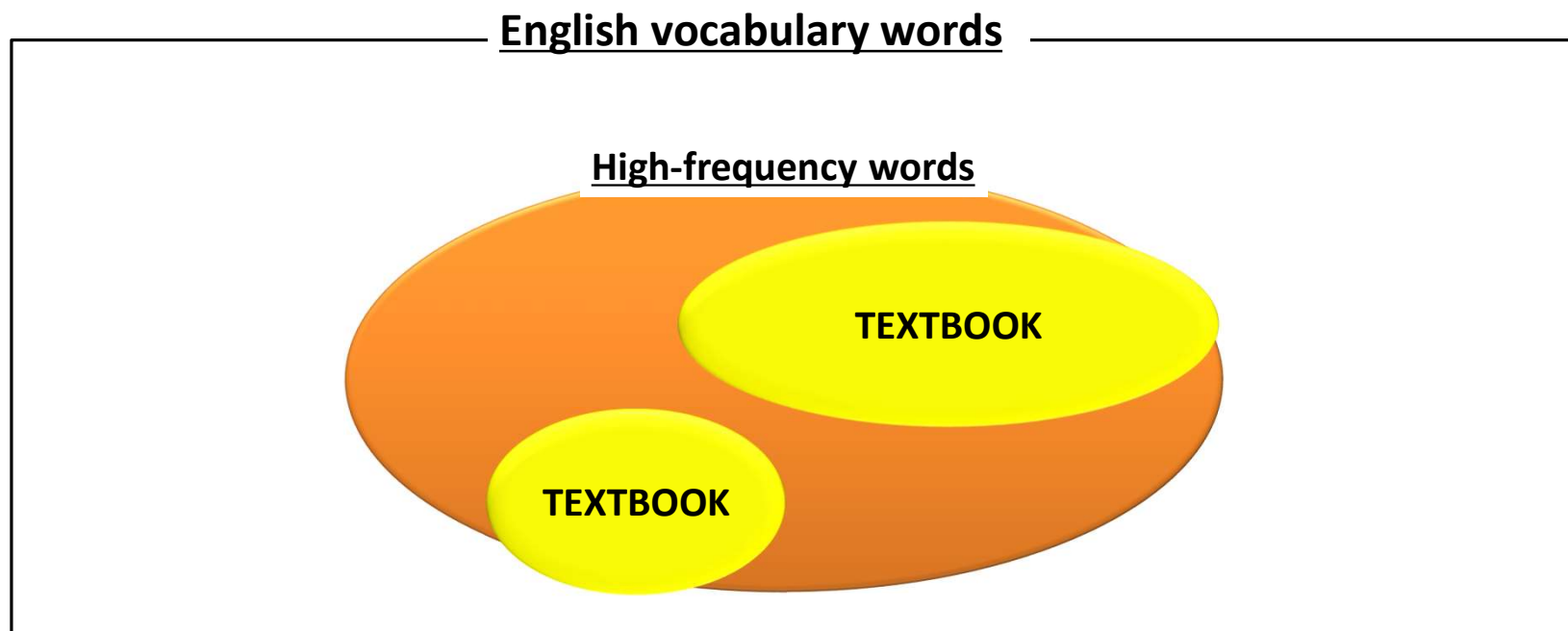
3. Research Question

1. How much vocabulary in MEXT-approved textbooks comprises what learners are likely to meet in the real world?

2. If Japanese students master all of the vocabulary words taught in MEXT-approved EFL textbooks, is that enough for them to read authentic texts?

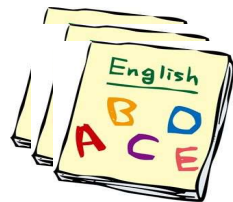MEIJI GAKUIN UNIVERSITY

4. Methods

Textbooks Under Analysis

- The top three best-selling MEXT-approved textbooks in Tokyo in 2021 (Tokyo Metropolitan Board Education, 2020)

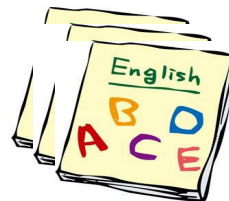- Textbooks are used in Komyunikeisyon Eigo (directly translated as English Communication) course.

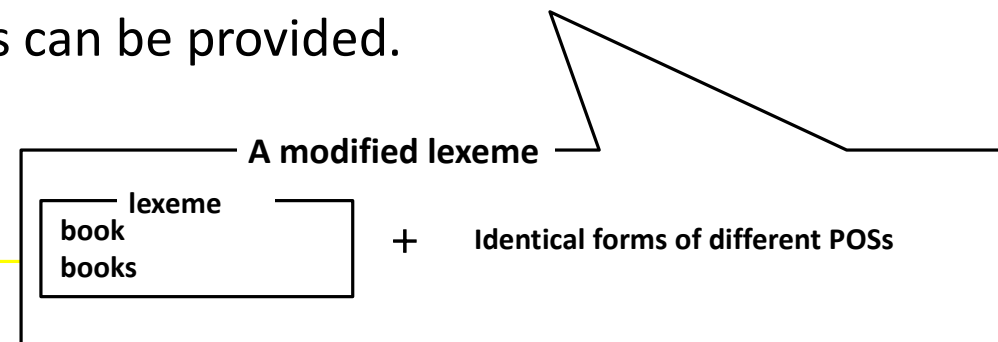SH1 × 3   SH2 × 3   SH3 × 3

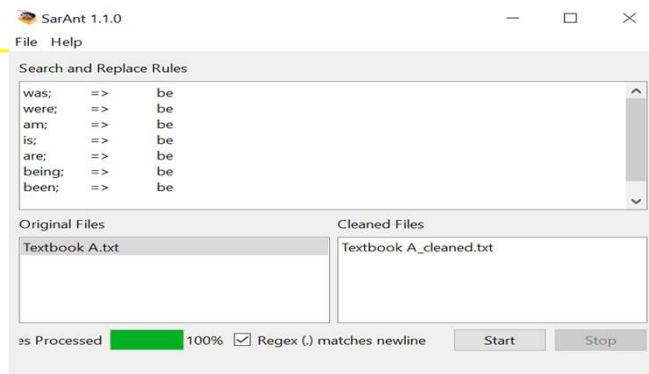Based on the website of Tokyo Metropolitan Board Education, 2020

MEIJI GAKUIN UNIVERSITY

Benchmark word list

The New General Service List ver. 1.01. (NGSL; Browne, Culligan, & Phillips, 2013)

- 273 million words within the Cambridge International Corpus

- the top **2,801 high-frequency words** in general English plus **52 supplemental words** such as days of the week, months of the year, and numbers where as

- proper nouns, abbreviations, slang, and other noise are excluded.

- As a word counting unit, **a modified lexeme approach** was adopted.

- around **92% coverage** of most general texts can be provided.

A modified lexeme

| lexeme
book
books | + | Identical forms of different POSs |

Anthony's SarAnt (version 1.1.0)

Text Lex Compare ver. 4.2. (2021)

SarAnt 1.1.0
File Help

Search and Replace Rules

was;	=>	be
were;	=>	be
am;	=>	be
is;	=>	be
are;	=>	be
being;	=>	be
been;	=>	be

Original Files
Textbook A.txt

Cleaned Files
Textbook A_cleaned.txt

es Processed 100% ☑ Regex (.) matches newline   Start   Stop

New words in second/last text   Units = tokens + types   Index-Edit-Area at bottom

First text(s): (2853 tokens/2853 types)
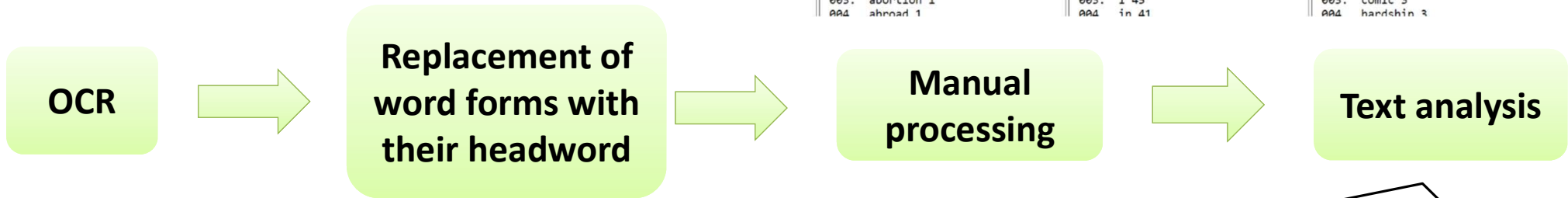Second text: (1464 tokens/451 types)

TYPES ANALYSIS
Using the word type as unit of comparison means that if *cat* is in Text 1 and *cats* in Text 2 then this is **not** considered a recurrence of the same word.

new! TOKEN Recycling Index: 1813 shared / 4317 total = 42.00%
TYPES Recycling Index: 410 shared / 2894 total = 14.17%

Unique to first(s)		Shared		Unique to second/last	
2443 tokens		1813 tokens		61 tokens	
2443 types	Extract	410 types	Extract	41 types	Extract
					VP novel items
001. abandon 1		001. the 94		001. penguin 5	
002. ability 1		002. be 72		002. emperor 4	
003. abortion 1		003. i 45		003. comic 3	
004. abroad 1		004. in 41		004. hardshin 3	

OCR ➡ **Replacement of word forms with their headword** ➡ **Manual processing** ➡ **Text analysis**

$$\frac{(\textbf{Number of words in the textbook} - \textbf{Number of words unique to the textbook})}{\textbf{Number of words in the textbook}}$$

= lexical coverage of textbooks

$$\frac{(\textbf{2,801} - \textbf{Number of words unique to the NGSL})}{\textbf{2,801}}$$
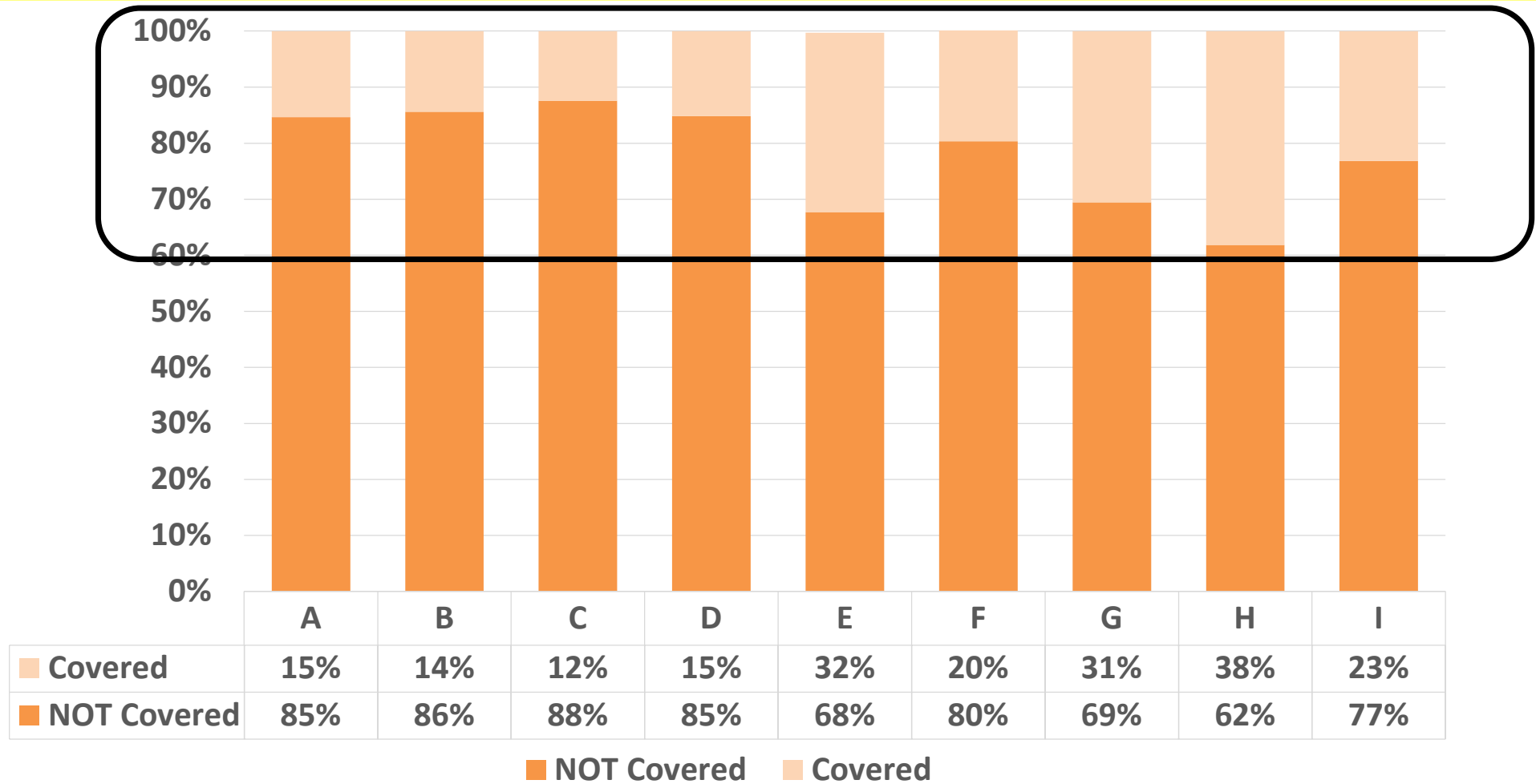
= coverage of the NGSL

5. Results

Results of Lexical Coverage Analysis

Grade	Textbook	Number of tokens/ (tokens unique to the textbook)	Number of word types/ (types unique to the textbook)	Coverage
SH1	A	1633(118)	512(74)	92.8%
	B	1464(61)	451(41)	95.8%
	C	1257(50)	384(30)	96%
SH2	D	1755(74)	486(53)	95.8%
	E	6149(395)	1153(232)	93.6%
	F	2475(144)	644(84)	94.2%
SH3	G	4923(335)	1064(192)	93.2%
	H	7457(487)	1388(300)	93.5%
	I	3164(217)	786(125)	93.1%

MEIJI GAKUIN UNIVERSITY
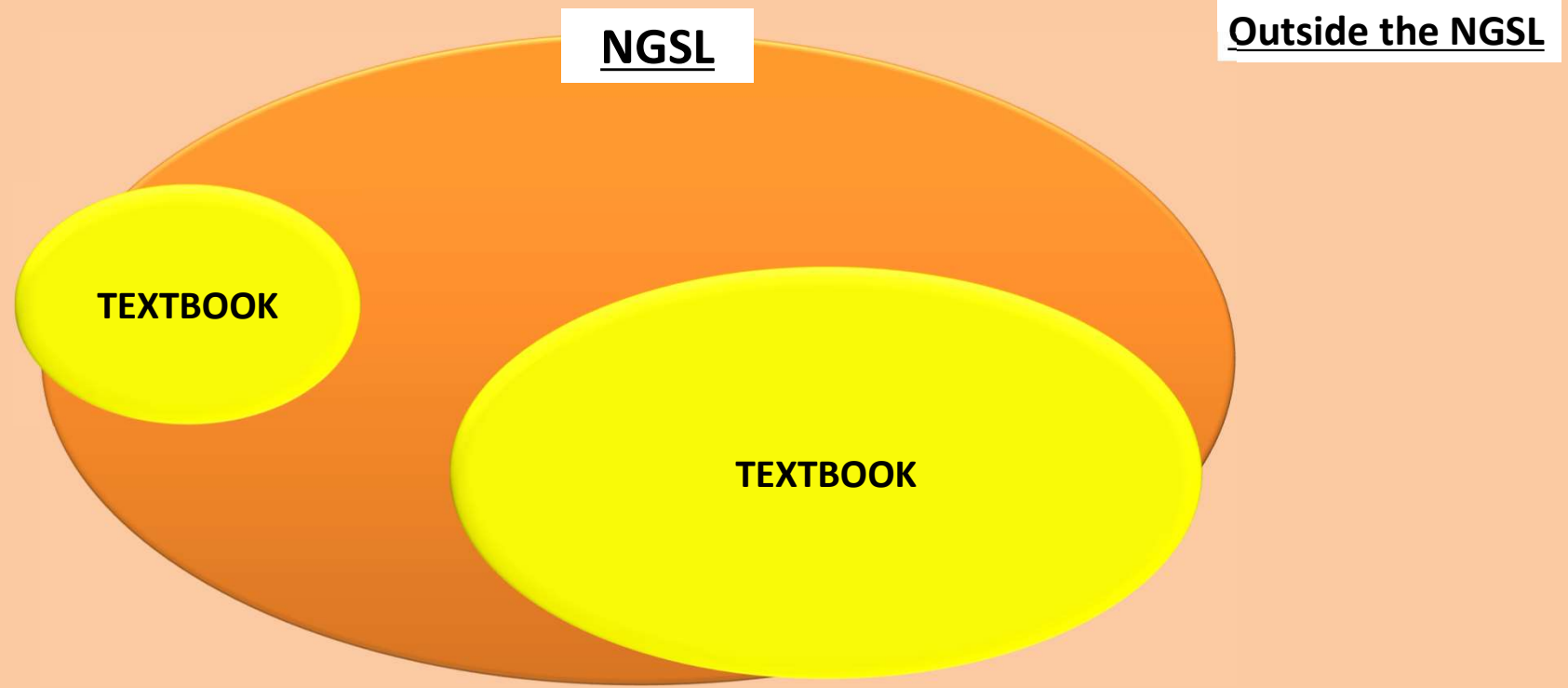
Proportion of NGSL Words Covered by the Textbooks to NGSL words not Covered by Them

	A	B	C	D	E	F	G	H	I
Covered	15%	14%	12%	15%	32%	20%	31%	38%	23%
NOT Covered	85%	86%	88%	85%	68%	80%	69%	62%	77%

■ NOT Covered ■ Covered

MEIJI GAKUIN UNIVERSITY

The Distribution of Words Toward Three Frequency Bands

Grade	Textbook	NGSL 1 (1st-1000)		NGSL 2 (1001st-2000)		NGSL 3 (2001st-2801)	
		(a)	(b)	(a)	(b)	(a)	(b)
SH1	A	84%	33%	6%	7%	2%	3%
	B	86%	31%	7%	7%	3%	3%
	C	88%	29%	5%	4%	2%	2%
SH2	D	87%	32%	5%	7%	3%	4%
	E	85%	62%	6%	22%	2%	10%
	F	84%	40%	6%	9%	3%	7%
SH3	G	83%	57%	7%	20%	3%	11%
	H	84%	69%	7%	28%	2%	13%
	I	84%	48%	6%	13%	2%	6%

Note. (a) = the percentage of words occurring at the frequency band; (b) = the coverage of the frequency band by the textbook.

MEIJI GAKUIN UNIVERSITY

6. Discussion & Conclusion

RQ 1: How much vocabulary in MEXT-approved textbooks comprises what learners are likely to meet in the real world?

Grade	Textbook	Coverage
SH1	A	92.8%
	B	95.8%
	C	96%
SH2	D	95.8%
	E	93.6%
	F	94.2%
SH3	G	93.2%
	H	93.5%
	I	93.1%

Vocabulary words taught in the textbooks were largely composed of words in the NGSL with greater than 92% lexical coverage.
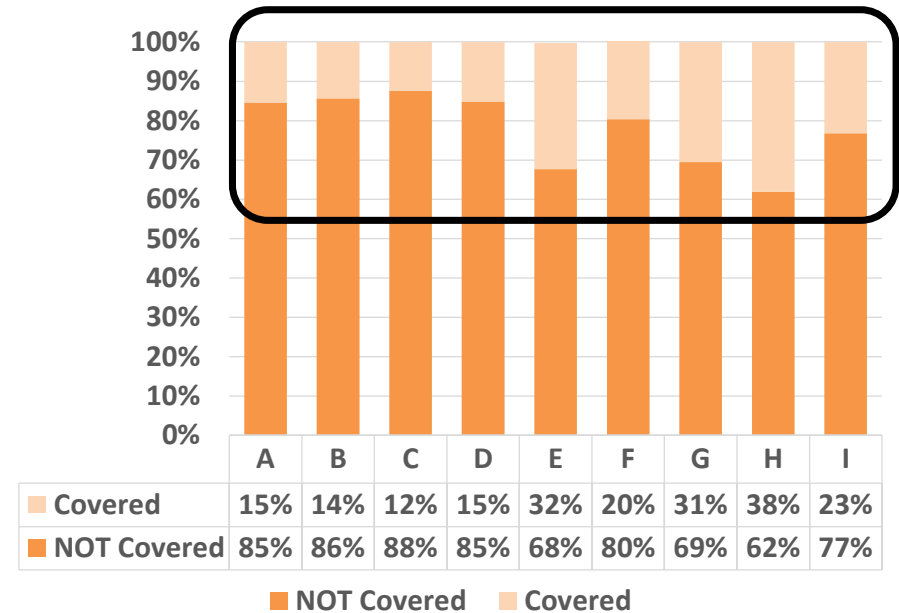
MEIJI GAKUIN UNIVERSITY

More than 80% of NGSL words occurring in the textbooks were found to be occupied by the first 1,000 most frequent words.

Vocabulary that learners frequently see in textbooks is extremely likely to appear in the real world.

Grade	Textbook	NGSL 1 (1-1000)	
		(a)	(b)
SH1	A	84%	33%
	B	86%	31%
	C	88%	29%
SH2	D	87%	32%
	E	85%	62%
	F	84%	40%
SH3	G	83%	57%
	H	84%	69%
	I	84%	48%

MEIJI GAKUIN UNIVERSITY

RQ 2: If Japanese students master all of the vocabulary words taught in MEXT-approved EFL textbooks, is that enough for them to read authentic texts?

NO

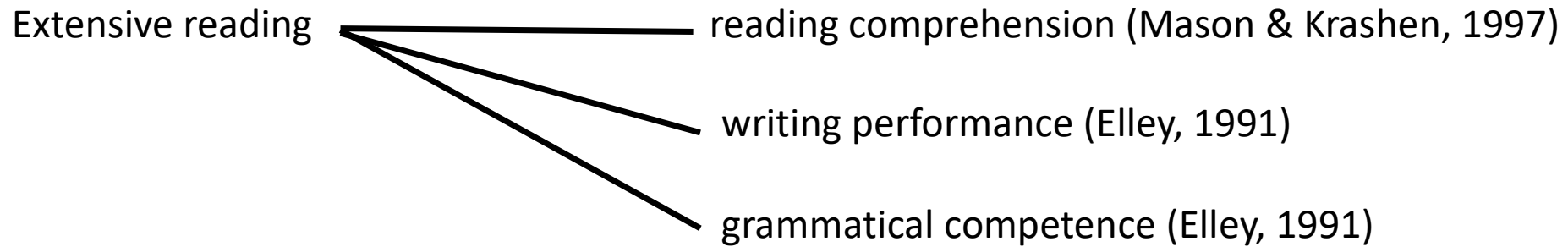|	A	B	C	D	E	F	G	H	I
Covered	15%	14%	12%	15%	32%	20%	31%	38%	23%
NOT Covered	85%	86%	88%	85%	68%	80%	69%	62%	77%

NOT Covered   Covered

The coverage analysis of the NGSL by the textbooks identified at a statistically significant level that they did not sufficiently cover the NGSL with less than 38% coverage.

MEIJI GAKUIN UNIVERSITY

The textbooks were severely deficient of words at the second and third frequency bands, and even words at the first frequency band were not fully covered.
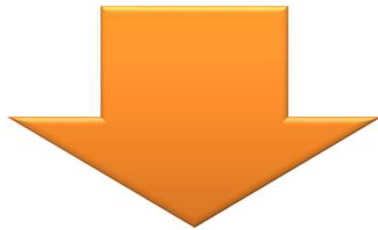
Additional input of basic words

Grade	Textbook	NGSL 1 (1-1000)	NGSL 2 (1001-2000)	NGSL 3 (2001-2801)
		the coverage of the frequency band by the textbook		
SH1	A	33%	7%	3%
	B	31%	7%	3%
	C	29%	4%	2%
SH2	D	32%	7%	4%
	E	62%	22%	10%
	F	40%	9%	7%
SH3	G	57%	20%	11%
	H	69%	28%	13%
	I	48%	13%	6%

MEIJI GAKUIN UNIVERSITY

Extensive reading ——————— reading comprehension (Mason & Krashen, 1997)

——————— writing performance (Elley, 1991)

——————— grammatical competence (Elley, 1991)

Watching English movies ——————— ability to guess word meaning from context (Nation, 2006)

7. Suggestions for Further Studies

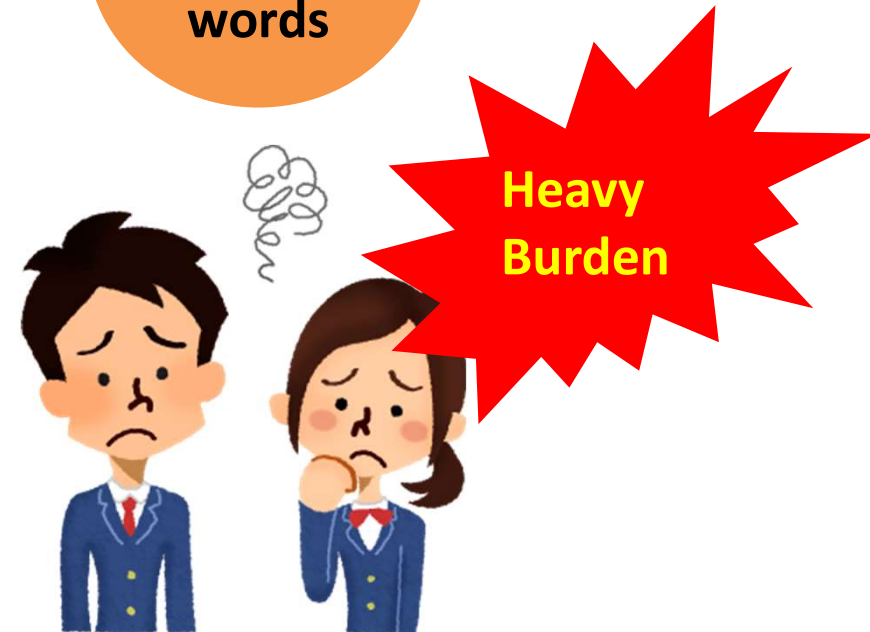## CURRENT Couse of Study Guidelines

3,000 words

NEW Couse of Study Guidelines

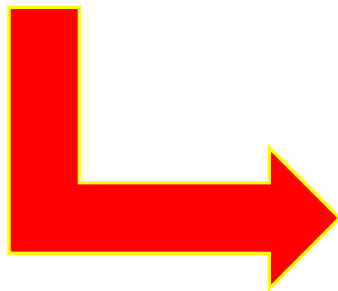**4,000- 5,000 words**

Heavy Burden

Students → spend the longest time looking up meanings of new words

Teachers → students were not good at memorizing new words.

(Benesse Educational Research and Development Institution, 2015)

MEIJI GAKUIN UNIVERSITY

Learners' vocabulary capacity

The optimal number of different words in a textbook

MEIJI GAKUIN UNIVERSITY

References

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education*, *3*(3), 21-34. https://doi.org/10.20472/TE.2015.3.3.002

Anthony, L. (2021). *AntWordProfiler* (Version 1.5.1) [Computer software]. Waseda University. https://www.laurenceanthony.net/software Accessed 30 March 2021

Benesse Educational Research and Development Center. (2015). *Tyukou no eigogakusyu ni kansuru zittai tyousa 2015* [*Teenagers English learning survey 2015*]. https://berd.benesse.jp/global/research/detail1.php?id=4776 Accessed 7 February 2021.

Browne, C., Culligan, B. & Phillips, J. (2013). The new general service list. http://www.newgeneralservicelist.org.

Cobb, T. (2021). *Text Lex Compare v.4.3* [computer program]. https://www.lextutor.ca/cgi-bin/tl_compare/

Cohen, J. (1992). A power primer. Psychological Bulletin, 112, 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Chujo, K. (2004). *Measuring vocabulary levels of English textbooks and tests using a BNC lemmatised high frequency word list*. In J. Nakamura, N. Inoue, & T. Tabata (Eds.), English corpora under Japanese eyes (pp. 231-249). Rodopi.

Elley, W. B. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning*, *41*(3), 375-411. https://doi.org/10.1111/j.1467-1770.1991.tb00611.x

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430.

Laufer, B. (1989). *What percentage of text lexis is necessary for comprehension?* In C. Lauren & M. Nordman (Eds.), Special language: From humans thinking to thinking machines (pp. 316–323). Multilingual Matters.

Liu, N., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal, 16*(33), 33-42. https://doi.org/10.1177/003368828501600103

Mason, B., & Krashen, S. (1997). Extensive reading in English as a foreign language. *System*, 25(1), 91-102. https://doi.org/10.1016/S0346-251X(96)00063-2

Nation, I. S. P. (2001). Learning vocabulary in another language. U.K., Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, *63*(1), 59-81. http://dx.doi.org/10.1353/cml.2006.0049

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the vocabulary levels test. Language Testing, 18(1), 55-88. https://doi.org/10.1177/026553220101800103

Wongsarnpigoon, I. (2018). Vocabulary in junior high school textbooks and exams. In P. Clements, A. Krause, & P. Bennett (Eds.), Language teaching in a global age: Shaping the classroom, shaping the world. JALT.

MEIJI GAKUIN UNIVERSITY

Thank you for Listening!

Shusaku Nakayama
shusaku-nakayama@outlook.jp

MEIJI GAKUIN UNIVERSITY