# Words for Success, Words for Life 

## Dr. Charles Browne

Professor of Applied Linguistics \& TESOL
Director, MA/PhD Programs
Director, New General Service List Project Meiji Gakuin University, Tokyo, Japan
browne@ltr.meiiigakuin.ac.jp
A few of My Free Websites \& Tools:
www.newgeneralservicelist.org
www.er-central.com
www.er-central.com/ogte

## Why vocabulary?...



## TWO ENORMOUS PROBLEMS WITH ENGLISH EDUCATION IN JAPAN WE ARE TRYING TO ADDRESS...


not enough/wrong vocabulary

reading materials far above student's level

## 600,000

NUMBER OF WORDS IN THE ENGLISH LANGUAGE:

## 30,000

NUMBER OF WORDS KNOWN BY NATIVE SPEAKER GRADUATE OF 4-YEAR UNIVERSITY:


Problem 1: learners just don't know enough words...


Problem 2: English reading materials are too hard!


With 2000 words, Japanese students can read...

Environment Special The Perils of Plastic


With 2000 words, Japanese students can read...


With 2000 words, Japanese students can read...

## $90 \%$



Actually it's more than $94 \%$ with the NGSL...

With 2000 words, Japanese students can ONLY read...


## But maybe 75\% coverage is enough?...



## Forestry

## If $\quad * \quad$ planting rates are _ $\quad$ 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 $\qquad$ * $\quad * \quad$, even if much isused for _ $\quad$ production.
meaning of the unknown words?

## 85\% coverage

- First 2000 most frequent words
- Total of 9 missing words

Is it easier to read and guess the

## Forestry

 with planting tangeets satisfined 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
## If * _ planting rates are maintained $^{2}$

 for___ production.
## Forestry

If current planting rates are maintained with planting targets satisfied in each annually and the forests millled at the earliest opportunity, the available wood supplies could further increase to about

## 95\% coverage

- First 5000 most frequent words
- Only 1 missing word

At this level, can you guess the meaning of the unknown word? Can you read more easily? 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.

## VOCABULARY THRESHOLD RESEARCH

- Below $80 \%$, reading comprehension is almost impossible (Hu \& Nation, 2001)
- In studies of student performance in ESL settings, knowledge of about $90 \%$ was the point at which there were more readers than non-readers (Laufer, 1992)
- $95 \%$ coverage is the point at which most learners can read without the help of dictionaries (Laufer, 1989)

|  | Analysis of Top 3 best-selling Ministry of Education-approved English textbooks for Japanese high school students |  |
| :---: | :---: | :---: |
|  | Text | Coverage from 2000 High Frequency Words |
| Are Japanese students reading the right | Spectrum | 71\% |
|  | Milestone | 78\% |
|  | Unicorn | 79\% |
|  | Unsimplified Native Texts | 80-85\% |
| 81\% 85\% 90\% |  |  |

2000 words gives about 80-85\% Coverage for nativespeaker reading materials but...

## Backwash Effect From Japanese University Entrance Exams...




The ngsl project is all about short cuts: Making tight LISTS OF CORPUS-DERIVED WORDS TO HELP STUDENTS QUICKLY reach the level of independent learning


## NGSL WORD LIST: a Balanced Corpus

273 million words from the CEC 2-billion-word corpus

CEC corpora included in final analysis for NGSL :

(Browne, Culligan \& Phillips, 2013)
*273 million word subsection used is 100x larger than original GSL corpus...

## Text coverage (sp/wr) 10m corpus Data

© Cambridge University Press 2002. Research © M McCarthy 2002.

## HOW IS THIS POSSIBLE?... ZIPF'S LAW...

George K. Zipf (1902 - 1950) established a mathematical model for describing frequency data:

Frequency decreases with rank. More precisely, frequency is inversely proportional to rank


## Next step after learning the NGSL:

## Mid-Frequency Vocabulary?

3,001-4,000: academic, consist, exploit, rapid, vocabulary 4,001-5,000: 5,001-6,000:
6,001-7,000: 7,001-8,000: 8,001-9,000:

Mid frequency vocabulary

Low frequency vocabulary

[^0]


These words occur EVERYWHERE in student's daily life

Books, newspapers, magazines, internet

- TV, radio, movies
- Conversational English


## Why do we believe that the NGSL should always be step 1?

- Knowing these words gets you $92 \%$ of the way there for written texts in daily life, even more for many spoken genres ( $95 \%$ of the TV show FRIENDS for example)
- The 2800 NGSL words represent less than $0.5 \%$ of the English language and only $9 \%$ of a college-educated native speaker's vocabulary
BONUS: The NGSL also happens to be a VERY good starting point for high stakes exams used in Japan...


## Coverage of NGSL for SHS Entrance Exam and National Center Test

## SOME NGSL RESEARCH



Iwamoto, M., Kramer, B. (2020) The Language Teacher 44:4

Abstract:

While the Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) currently expects students to learn 1,200 English words in junior high school and 1,800 English words in high school (MEXT, 2017), there is little to no guidance on the specific words required. Looking at the reading sections on Japanese public high school entrance examinations and the university National Center Test, this study reports the lexical coverage provided by a well-known and publicly available word list, the New General Service List (NGSL) (Browne, Culligan, \& Phillips, 2013). The NGSL provided a high 98.11\% coverage of the vocabulary on senior high school entrance examinations using only 1,000 words but was only able to cover $95.26 \%$ of the vocabulary on the National Center Test with all 2,801 words. The results will be discussed in detail, along with the utility of the NGSL in Japanese junior and senior high school classrooms.

## Zipf's law revisited...

As you can see, these high frequency words are extremely useful and important for success as a $2^{\text {nd }}$ language learner

The problem is that if students have any significant gaps in their knowledge of the highest frequency words (most especially the blue words!), it is almost mathematically impossible to reach the crucial $90 \%$, 95\% and 98\% thresholds.

This means it is crucial to help your students to master ALL of the NGSL words as quickly as possible....

## CREATING EFFICIENT PATHES TO EFL SUCCESS:

STEP 1: Learn high frequency General English words

STEP 2: Learn high frequency Specific


288 Million Word Academic Corpus

The New Academic Word List (NAWL)

| Corpora | Size | Percent |
| :--- | :--- | ---: |
| CEC Academic | 248 million | $86.30 \%$ |
| Oral | 3 million | $1.10 \%$ |
| Textbooks | 36 million | $12.60 \%$ |
| Total | 288 million | $100 \%$ |

Note that NGSL drops to 86\% for academic English

960 Academic Words (NAWL)

## Coverage of NGSL \& NAWL for Research Publications in Education

## SOME NGSL RESEARCH



Hendry, C., \& Sheepy, E. (2018). How much vocabulary is needed for comprehension of research publications in education? In P. Taalas, J. Jalkanen, L. Bradley \& S. Thouësny (Eds), Future-proof
CALL: language learning as exploration and encounters - short papers from EUROCALL 2018 (pp. 94-99).

## Abstract.

The American Education Research Association (AERA) is one of the largest education conferences in the world. Using the AERA Open Access Repository, we created a 5,000,000 word corpus of over 18,000 abstracts. We explored the coverages of the New General Service List (NGSL), the New Academic Word List (NAWL), and the Social Science Word List (SSWL). We found that the NGSL and NAWL provide approximately $90 \%$ coverage for abstracts from all 12 of the AERA's subject matter divisions. The SSWL showed little additional coverage. Our discussion highlights the research and pedagogical implications of our findings and the AERA abstract corpus.


## The TOEIC Service List (TSL)

- All official TOEIC practice tests from ETS
- 600 Essential Words for the TOEIC - 3rd Edition
- Achieve TOEIC Bridge
- Barron's TOEIC Practice Exams
- Collins' Practice Test for the TOEIC Test
- Oxford practice tests for the TOEIC test
- Oxford Preparation Course for the TOEIC test
- Pass the TOEIC Test Advanced
- Pass the TOEIC Test Intermediate
- Practice Examinations for the TOEIC Test
- Tactics for TOEIC Listening and Reading Tests
(Browne \& Culligan, 2015)


### 1.5 Million Word TOEIC Corpus

Note that NGSL increases to 94\% for TOEIC tests

99\%


94\%

2800
Core General English Words
(NGSL)

## Coverage of TSL for the new version of the TOEIC exam

Since the TOEIC Service List (Browne \& Culligan, 2016) is based on a 1.5 m words of TOEIC prep materials from the previous format of the TOEIC test, we thought it was time to see what kind of coverage the TSL and NGSL word lists offer for the new test.

To this end, we created a small corpus from the 8 official practice tests published by ETS in 2019 and ran the analysis. Results were quite good...

| Word List | Cumulative Coverage |
| :--- | :--- |
| NGSL (version 1.01) | $93.3 \%$ |
| NGSL + NGSL Supplemental | $94.4 \%$ |
| NGSL + TSL | $98.5 \%$ |

## Overview of NGSL Word Lists and Coverage Provided

| Word List | Focus | Words | Coverage |
| :---: | :---: | :---: | :---: |
| New General Service List (2013) | General English | 2801 | $92 \%$ |
| New Academic Word List (2013) | Academic English | 960 | $92 \%$ |
| TOEIC Service List (2015) | TOEIC Test | 1200 | $99 \%$ |
| Business Service List (2015) | Business English | 1700 | $97 \%$ |
| New General Service List-Spoken (2014) | Spoken English | 718 | $90 \%$ |
| New Dolch List (2021) | Children's English | 875 | $90 \%$ |
| Fitness English List (2020) | Fitness English | 600 | $98 \%$ |

## The NGSL Project Resource Center www.newgeneralservicelist.org

## FREE NGSL RESOURCES INCLUDE:

- All Wordlists downloadable
- Original English definitions
- Wordlist-related research articles
- Online learning applications
- Authentic video concordance tool
- Online assessment tools
- Online research tools
- Online text analysis tools

$+-$


## All wordlists are available for download in multiple formats：

## －Alphabetized

－Lemmatized for teaching purposes
－Lemmatized for research purposes
－With frequency and basic statistical data
－With definitions in easy English

| BSL＿1．01＿SFl＿freq＿bands |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Word | BSL Rank | SFI | $u$ | D | F |
| mister | 1 | 71.51 | 1416.4 | 0.3911 | 1483 |
| goods | 2 | 66.71 | 468.5 | 0.9049 | 325 |
| equity | 3 | 66.15 | 411.67 | 0.8449 | 29962 |
| dividend | 4 | 65.15 | 327.16 | 0.7245 | 26620 |
| portfolio | 5 | 64.97 | 314.17 | 0.704 | 2598 |
| sponsorship | 6 | 62.37 | 172.54 | 0.1475 | 23455 |
| inventory | 7 | 63.62 | 230.19 | 0.5898 | 21186 |
| transaction | 8 | 64.52 | 283.43 | 0.8335 | 208 |
| non | 9 | 64.38 | 273.88 | 0.9751 | 1803 |
| lease | 10 | 63.06 | 202.09 | 0.6443 | 1792 |
| hedge | 11 | 62.32 | 170.71 | 0.6083 | 15666 |
| distribution | 12 | 63.3 | 21.91 | 0.8564 | 154 |
| premium | 13 | ． 93 | 196.21 | 0.7651 | 154 |


| clien |
| :---: |
| impa |


| authe | 1 | 1 |
| :--- | :--- | :--- |
| obtai | 2 |  |
| matu | 3 |  |
| publi | 4 |  |
| some | 5 |  |
|  | 6 |  |
|  |  |  |


| Meanings afterward |
| :---: |
|  |  |
|  |
| artistic |
| backward |
| bodily |
| carrier |
| collective |
| computation |
| continuity |
| definite |
| freely |
| generalization |
| generalize |
| goodness |
| historically |
| importantly |
| interestingly |
| locally |
| machinery |
| marker |
| meaningtul |
| namely |

NAWL＿1．0＿with＿en＿definitions

| happening after something else happened |
| :--- |
| in or through a group |
| Being able to paint，write music，etc．；creative |
| toward what is behind you |

physically as when lifting someone
adv その後
adj 芸行的な
adv 後3へ
adv 身体の
a company，usually an aririne，that carries things
formed by worki
the act of calculating；method of computing
continued exsterce，facto f not stopping．

making a statement about a group from limited data To apply a particular idea to a large group nourishment；good food value throughout history
you＇ll find this surprisising；in an interesting way
Concerning the area around where you are machines in general
something that is easy to recognize or identify Having importance；showin
that is to say，specifically
The spirit of love of your country，e．g．leading to war
n 䚺算
adv 自由に
一般：一般
verb 一般化する
adv 整的に
adv 興味深いことに
楼桃
標辐；マーカ
－○ BSL＿1．01＿lemmatized＿for＿teaching．txt abolish，abolishes，abolished，abolishing absorb，absorbs，absorbed，absorbing absorption，absorptions
accelerate，accelerates，accelerated，accelerating acceptance，acceptances
accessible
accommodate，accommodates，accommodated，accommodating accomplishment，accomplishments
accordance
accordingly
accordingly
accountability，accountabilities
accountable
accountants
accrual，accruals
accrue，accrues，accrued，accruing
accumulate，accumulates，accumulated，accumulating accumulation，accumulations
accuracy，accuracies
accurately
accusation，accusations
acid，acids
acre，acres
acre，acres
actively
activist，a
activist，a
adequately
$\qquad$
inistrator
admission
prep～の間に；～の中の

荤送業；航空会社；電話会社
adj 収集された；集団の
adj 明磼な；限定的な

する
adv 重要なことには
adj 启所的に；地元で

すなわち；
www.newgeneralservicelist.com

- List of NGSL research
- NGSL blog posts
- NGSL TED Talk


A Brief Introduction to the NGSL Project..



81273

## Online Graded Text Editor (OGTE)



6/20723
Academic English: Spotlight on the New Academic Word List (NAWL)


Most NGSL wordlists available on key Corpus Text Analysis and Text Editing Tools...
automated data mining survey responses coy
qualatatiy
classificatitext teranscripts
ad-hoc an
reviews sen cause
customer dashboards consuma trends ad-hoc analysis early warning


Browne \& Waring's OGTE: https://www.er-central.com/ogte/


## NGSL Profiler (Coming Soon!)

- analyze text difficulty using all 7 NGSL lists
- simplify texts to student's ability level
- ChatGPT tool for generating texts
- Text rewriting tool to help simplification

NGSL Profiler (Now with all NGSL lists and AI tools to help you created and level texts for your class)
Text Leveler Text Generator Text Rewriter

Enter a passage of text in the box below. Select the desired wordlist. Choose the level of the wordlist to check against and press the GO button.

| Wordlist | Level |  |  | GO! | Ignore this word |
| :---: | :---: | :---: | :---: | :---: | :---: |
| NDL | * level |  | $\stackrel{\rightharpoonup}{*}$ |  |  |
| WORDS IN THE CURRENT LIST | WORDS OUT OF CURRENT LIST |  |  | Coverage |  |
| (A) $0 \%$ <br>  <br> (B) $0 \%$ out of selected level | (C) ${ }_{\text {Promer }}^{\text {Propeouns }}$ | (D) ${ }_{\text {l }}^{\text {lgored }}$ | (E) $0 \%$ | (A) + (C) | (D) $0 \%$ |



Words to Ignore
Add a space separated list of words

| Out of List |  |
| :--- | :--- |
| Word | Frequency |
| year-end | 3 |
| equities | 3 |
| equity | 2 |
| strategists | 2 |
| interest-rate | 2 |
| strategist | 2 |
| Out of Level |  |
| Word | Level |
| No words | Frequency |
| Ignored |  |
| Word | Frequency |
| year | 9 |
| s\&p | 4 |
| that's | 3 |
| america's | 2 |
| goldman | 2 |
| sachs | 2 |

## NGSL Online Learning Tools

- flashcards
- dictionaries
- word games
- video databases



## Pedagogic Principles Behind our Learning Tools

1) The need to learn from authentic contexts
2) The benefits of spaced-repetition learning
3) Increased motivation through gamification

## Video Concordancing Resource

- Based on the idea of Data-Driven Learning
- Instantly provides thousands of short authentic video clips to teach key words
- Utilizes the free YouGlish.com website




## Kahoot: A free Quiz Software for the classroom

You can find NGSL lists on Kahoot too!


## Short-term memory loss

## Strengths \& Weaknesses of Quizlet, Memrise \& KAHOOT

## Strengths?

- Works with any word stack
- Has very attractive interface
- Teaches multiple aspects of word knowledge


## Weaknesses?

- Doesn't allow for multiple repetitions or take into consideration the science of spaced-repetition and time-intervalled learning
- Not much contextualization for the words


## The Forgetting Curve

- If you study only once, you quickly forget
- If you review the words just as you are about to forget them, the memory becomes stronger and lasts longer
- Each additional review strengthens the memory until it moves to long-term memory

Typical Forgetting Curve for Newly Learned Information


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


## Spaced Repetition is the science of long-term memory

Based on the research of Ebbinghaus, Pimsleur, Leitner, and Mondria, electronic flashcards automatically repeat each new word at spaced time intervals, and until the learner achieves long-term, instant-recall ability.


Free Spaced-Repetition amd gamified flashcard learning apps for NGSL word list

- NGSL Builder (flashcards for NGSL)
- Word-Learner (gamified for all NGSL lists)


< Back Study Your Words



## Congratulations

## NGSL Placement Test

- shows you where to start
- links to our learning tools


## NGSL <br> Placement Test

This is a short quiz to help you find the best set of words for you to study based on your ability. You will see a list of 20 words. Read the words and count how many you know. Then click on the number of known words below the list. This will take you to another list to narrow down the selection. Repeat the process. All together, you will see 100 words. At the end, we will suggest the best words for you to study.

Click here to begin

Read the following list of 20 words.

| peer | neglect | formula | when |
| :---: | :---: | :---: | :---: |
| target | condition | preserve | certain |
| human | tour | page | prepare |
| wall | nature | evil | connection |
| silly | everywhere | enhance | survey |

How many of the words do you know?
0-6
7-13
14-20

You have finished the quiz. You should study:
NGSL 1.20 Band 26: Words 2501 to 2600
For the words you should study, Click here

| LinguaEnditio Levin |  | Band |  | Set |
| :---: | :---: | :---: | :---: | :---: |
| NGSL 1.20 Learning Dictionary Band 26: Words 2501 to 2600 |  |  |  |  |
| Rank | Word | Pronunciation | PoS | Definition |
| 2542 | abortion | la'bor.jon/ | noun | (1) © female is going to have a baby |
| 2562 | allege | la'led 3 ! | verb | (1) ¢ ¢ ¢ M $M$ to state without proof |
| 2521 | automatically. | L.ote'met.i.lil | adv | (1) ¢ ¢ |
| 2556 | bin | (bin) | noun |  |
| 2545 | boom | loum | verb | OEP ¢ ¢ CM S to make a deep hollow resonat sound |
| 2539 | boost | bust | verb |  |
| 2583 | carbon | l'karbent | noun |  |
| 2517 | ceremony | ['ser.a, mou.nil | noun |  |
| 2531 | chamber |  | noun | (1) special purpose |
| 25 | chat | Hfew | verb |  |
| 2585 | chesse | 15[][] | noun |  |
| 2529 | clause | \|kloz] | noun | O legal document |

## The NGSLT and NAWLT TESTS

- Placement and diagnostics tests
- Developed by Tim Stoeckel \& Phil Bennet
- Among highest reliability and validity of any current vocabulary test
- Many research articles published on these two tests
- Available for free download and use from our website


Based on your score in this test, we have chosen a list of words for you to study. This list is on the next page.
Please look at the words on the list, highlight any words that you don't know, and try to learn them.


| Jevent $?$ Sountion |  | ancient <br> assistant <br> attach <br> besides <br> chart <br> closely <br> coast <br> county <br> criminal <br> ill <br> implement <br> impose <br> kick |
| :---: | :---: | :---: |

## More Gamified Tools

(all tied to our interactive learning dictionary)



Across
5. other than or in addition to
7. to make someone to do, believe, or accept something 10. related to illegal activity; $(n)$ a person who has done a crime
11. to be associated or connected with
13. by going through, by way of
14. very old; having lived a very long time ago Down

1. able to move from one place to another
2. to give way to someone or something else
3. person who helps another, usually as a job
4. act or process of writing and passing laws
5. to start and carry out a plan or government policy; (n)
6. many things or people

## Goals of the NGSL Project...

1. To create a group of vocabulary lists that work together to teach English in an efficient and systematic way
2. To provide these lists in a variety of forms including with definitions in easy English so that they could more easily be used
3. To create (or utilize existing) online tools to help teachers, students, researchers and content creators to better utilize these lists

## DIRECTIONS:

1) Take the NGSL Placement test:
www.linguaeruditio.com/word-stock/
2) Try to learn some NGSL words using any of the tools you are interested in:
www.newgeneralservicelist.com/tools
3) Study at least $10 \mathrm{~min} /$ day, every day for 1 week.

## Group Video Project


4) Keep a diary each day of what you noticed or felt
5) Meet again with your group to discuss your reactions compare these tools to how you have studied vocabulary in the past
6) Make a short group video with your reaction and send it to me by February 22:
browne@ltr.meijigakuin.ac.jp
7) I will send each group a video reply within 1 week after that

# Words for Success, Words for Life 

## Dr. Charles Browne

Professor of Applied Linguistics \& TESOL Director, MA/PhD Programs
Director, New General Service List Project Meiji Gakuin University, Tokyo, Japan

A few of My Free Websites \& Tools:
www.newgeneralservicelist.org
www.er-central.com
www.er-central.com/ogte


[^0]:    * Studying mid-frequency vocabulary seems a logical next step but it takes most learners 10 years to learn 3000 words, so taking another 20 years to learn the 6000 midfrequency words is not practical for most learners....

