Evaluating the lexical difficulty of teaching materials with NWLC

NWLCで教材の難しさを評価する

Brett Milliner, ミリナー・ブレット

Center for English as a Lingua Franca, Tamagawa University
milliner@lit.tamagawa.ac.jp

ABSTRACT

This article provides a practical example of how English teachers can use the vocabulary profiling application, New Word Level Checker—NWLC (https://nwlc.pythonanywhere.com/) to evaluate the lexical difficulty of teaching materials. NWLC provides teachers with a quick and objective approach for appraising the difficulty of a prospective text, test, or worksheet. The paper begins with an introduction to the coverage comprehension model (McLean, 2021) and its implications for selecting classroom materials. Importantly, this section discusses the recommended coverage benchmarks for different receptive modalities, or what percentage of words learners have to know from a text to sufficiently comprehend it. A step-by-step example of how to use the NWLC to measure the lexical difficulty of a TED® talk follows.

KEYWORDS: Vocabulary profiling, Corpus-based approach, Lexical coverage, Viewing comprehension, ELT materials design

1. INTRODUCTION

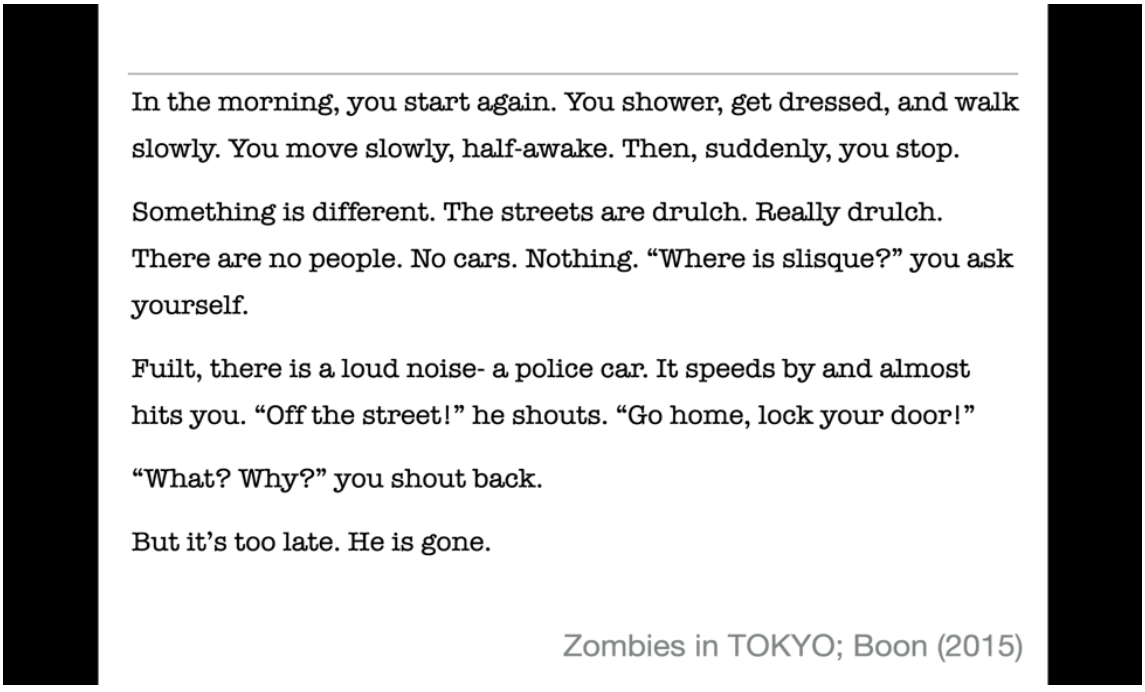
Do you ever come across a text or a scene in a movie where you think to yourself, "this would be nice to use in class", or wondered, "will this text be too difficult for my students?" This article introduces an efficient and objective approach for answering these questions—surveying the lexical difficulty of a text with a vocabulary profiling tool. In the example presented in this article, all a teacher needs to do is copy and paste a text, transcript, or exercise into the free website, New Word Level Checker—NWLC (https://nwlc.pythonanywhere.com/). NWLC will then report on the text's lexical coverage. In more practical terms, the profile will indicate within which word frequency band (usually a band of 1000 words–1K) the text's vocabulary falls into. As the meaning of higher frequency words are more likely to be known to students, a text comprised of more high-frequency words will generally be easier for learners to comprehend. To that end, the lexical profile report can be used as a measurement of a text's difficulty,

and determine its appropriateness for different learning tasks. This article will provide English teachers with a practical example of how to use NWLC to evaluate the lexical coverage (difficulty) of a TED talk. But first, however, it is important to discuss the developments in lexical coverage research and how it can help teachers make more informed decisions about the appropriateness of classroom materials.

2. COVERAGE COMPREHENSION MODEL & SELECTION OF CLASSROOM MATERIALS

The coverage comprehension model (see McLean, 2021) argues that if learners know the meanings of approximately 98% of the words within a written text, the lexical difficulty of the text should not inhibit comprehension. To put it more succinctly, if a language learner knows fewer than 98% of the words on a page, they will have trouble comprehending the text (e.g., Hu & Nation, 2000; Laufer, 1989, Schmitt et al., 2011). For readers hearing about this concept for the first time, 98% may seem like a conservative benchmark, as learners can draw from background knowledge and use other metacognitive strategies to overcome the comprehension gaps created by unknown vocabulary. However, in the example texts below where pseudowords have been used in place of real words, the comprehension difficulties even a very advanced English reader experiences when reading with 95% (Figure 1) and 90% (Figure 2) coverage ought to be glaringly obvious. Furthermore, it is worthwhile noting that all words are not equal. Comprehension difficulties are compounded when an unknown word significantly contributes to a text's meaning.

Figure 1
An example of a graded reader text with 95% coverage



In the morning, you start again. You shower, get dressed, and walk slowly. You move slowly, half-awake. Then, suddenly, you stop.

Something is different. The streets are drulch. Really drulch. There are no people. No cars. Nothing. "Where is slisque?" you ask yourself.

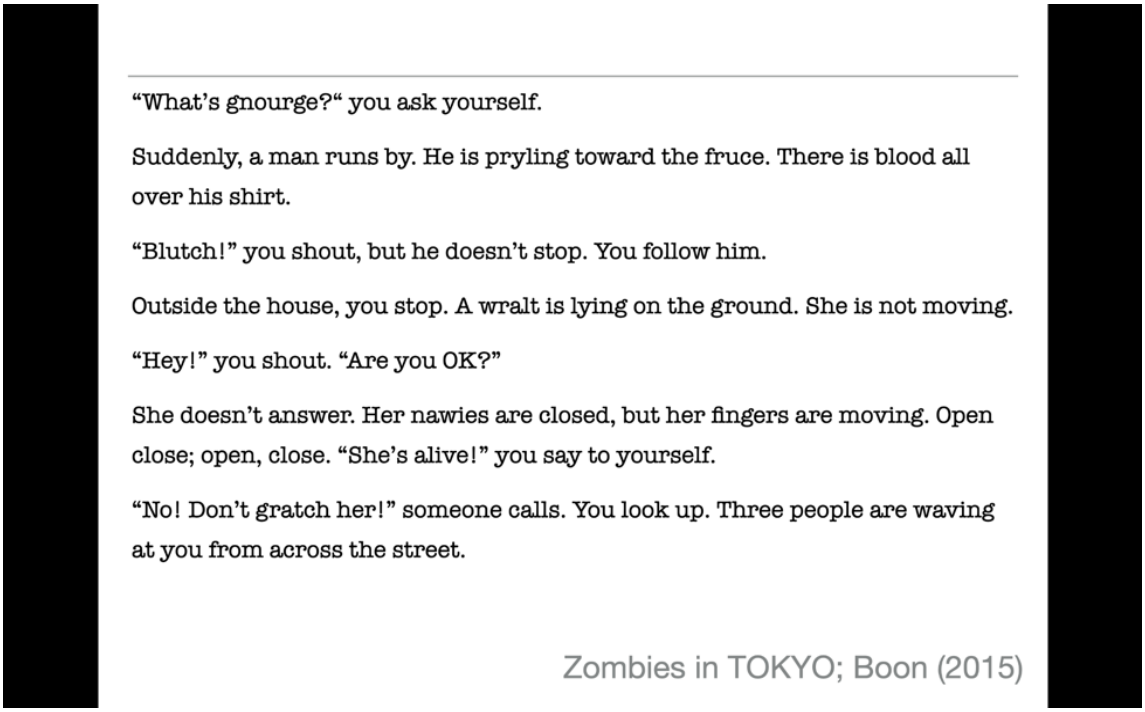
Fuilt, there is a loud noise- a police car. It speeds by and almost hits you. "Off the street!" he shouts. "Go home, lock your door!"

"What? Why?" you shout back.

But it's too late. He is gone.

Zombies in TOKYO; Boon (2015)

Figure 2
An example of a graded reader text with 90% coverage



Some readers may also wonder whether different modalities, such as audiovisual input, where learners can draw from spoken and visual cues, necessitates different coverage comprehension benchmarks. However, studies looking at this question have tended to recommend coverage levels close to 98%. In listening comprehension, 90% coverage may be possible in some contexts. Giordano (2021) found that some Japanese learners could comprehend videos of casual dialogues at 90% coverage. Similarly, van Zeeland and Schmitt (2013) found that a small proportion of their participants (foreign students taking graduate courses in the UK) could comprehend spoken narrative texts at 90% coverage. Nevertheless, in both studies, the researchers concluded that listening comprehension was much more stable across their entire samples when coverage was over 95%. In a study that evaluated viewer comprehension while watching a documentary series, Durbahn et al. (2020) found that Chilean learners' comprehension scores improved from 62% to 87% when the vocabulary coverage changed from 92% to 99%. In terms of listening in academic contexts, such as a lecture, coverage benchmarks of 98% were recommended by Noreille et al., (2018) and Stæhr, (2009). To summarize, while there are some variations in coverage benchmarks for different input modalities and genres, these differences are negligible. Much more coverage comprehension research is needed, particularly in the area of audiovisual input, but the evidence thus far suggests that language learners need to know almost every word in a text to sufficiently comprehend it.

When interpreting the coverage comprehension model, language teachers also need to understand that there are often imbalances within learners' receptive

vocabulary knowledge. Across a range of L1 backgrounds (e.g., Milton et al., 2010; Mizumoto & Shimamoto, 2008; van Zeeland, 2013) English learners' spoken receptive vocabulary knowledge (i.e., the vocabulary knowledge available while listening) is typically reported to be lower than their written receptive vocabulary knowledge (i.e., the vocabulary knowledge available while reading). Therefore, students tend to find spoken or audiovisual texts harder because their spoken receptive vocabulary knowledge is underdeveloped. Furthermore, the listening modality in general deserves special consideration because language learners cannot control speech rates or accents, and when compared to reading, words are only available for a fleeting moment. Therefore, the potential weaknesses within learners' spoken receptive vocabulary knowledge and a lack of agency felt by learners while listening suggests that a more prudent selection of spoken or audiovisual texts would prioritize a 98% lexical coverage benchmark.

Lexical coverage research also has implications for the types of activities teachers implement. For teachers familiar with the four strands (Nation, 2007), activities focusing on fluency development (e.g., timed reading; see Milliner, 2021), lexical coverage ought to be 100% (i.e., no unknown words in the text). For meaning-focused input tasks such as extensive reading or extensive viewing, a 98% coverage level is recommended (Nation, 2007). And, in language-focused instruction (e.g., exercises in an intensive reading textbook with glossaries, dictionaries, and other vocabulary support) a recommended lexical coverage figure is nothing lower than 85% (Schmitt et al., 2011; Stoeckel et al., 2020).

3. NEW WORD LEVEL CHECKER—NWLC

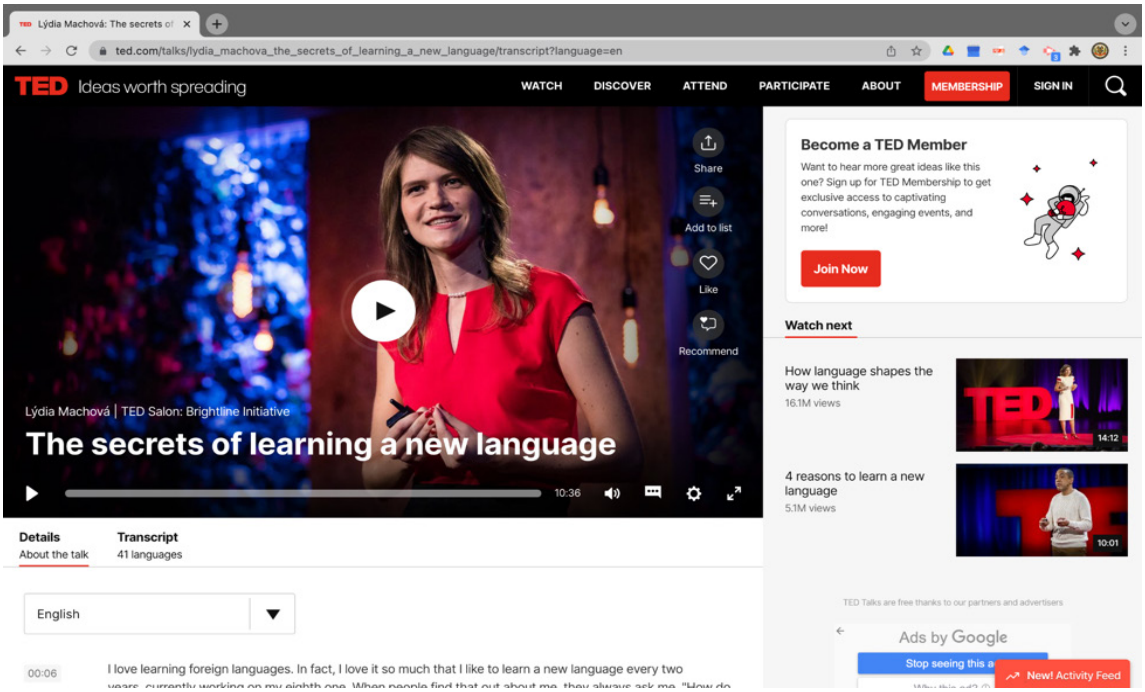
New Word Level Checker (NWLC) (https://nwlc.pythonanywhere.com/) is a web application for vocabulary profiling designed by Professor Atsushi Mizumoto (Kansai University) to meet the needs of Japanese learners of English. NWLC analyzes English words submitted by the user and produces vocabulary levels based on the selected word lists. As of writing, the NWLC can provide lexical profiles using the New JACET8000, SVL12000, the New General Service List, CEFR-J, and SWEK-J lists. All of these word lists were constructed with the needs of Japanese learners of English in mind. The NWLC is just one of many free vocabulary profiling tools (e.g., lextutor.ca & Antconc), but what makes it stand out is its use of flemma and lemma-based word lists, the user-friendly design, and its focus on Japanese learners of English. For a more detailed description of NWLC please see Mizumoto et al. (2021).

4. HOW TO USE NWLC

In this example, a TED talk titled *The secrets of learning a new language* (Machová, 2018) is profiled for a prospective listening task.

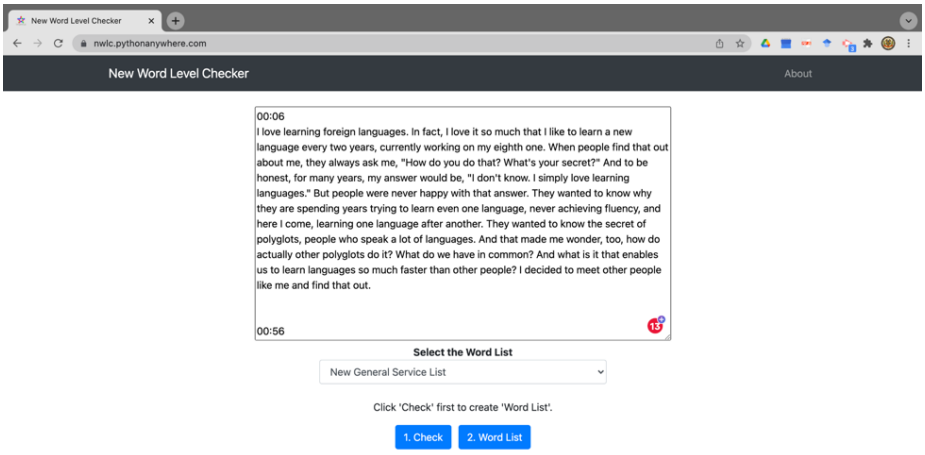
The first step was to copy the transcript from the TED webpage (Figure 3).

Figure 3
The TED talk transcript used for this demonstration (Machová, 2018)



Then, the transcript was pasted into NWLC (Figure 4). In this example, the Word List selected is the New General Service List—NGSL (Browne et al. 2013) because it is one of the most recognized word frequency lists in our context. As mentioned above, a variety of Word Lists are available on the NWLC, and the JACET8000 would also be an informative alternative for this profiling task.

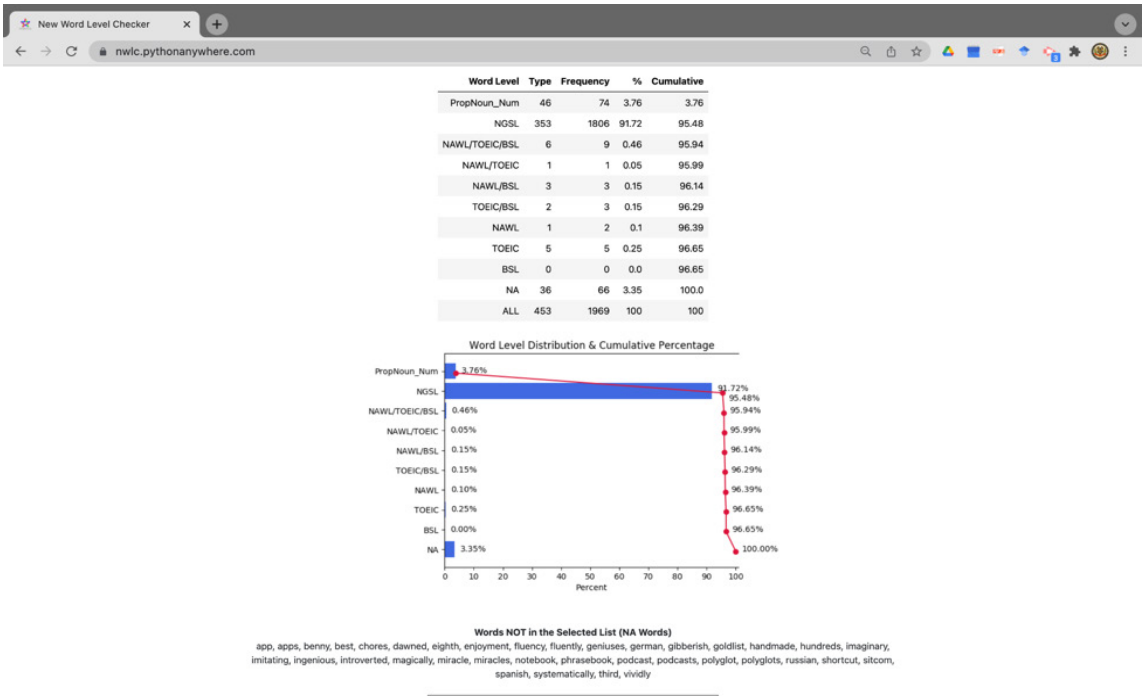
Figure 4
The TED talk transcript pasted into NWLC



From this point, you click the blue “1. Check” button, and the following text profiling report is generated (Figure 5).

Figure 5

The profiling report from NWLC



Word Level	Type	Frequency	%	Cumulative
PropNoun_Num	46	74	3.76	3.76
NGSL	353	1806	91.72	95.48
NAWL/TOEIC/BSL	6	9	0.46	95.94
NAWL/TOEIC	1	1	0.05	95.99
NAWL/BSL	3	3	0.15	96.14
TOEIC/BSL	2	3	0.15	96.29
NAWL	1	2	0.1	96.39
TOEIC	5	5	0.25	96.65
BSL	0	0	0.0	96.65
NA	36	66	3.35	100.0
ALL	453	1969	100	100

This report shows that 91.72% of the words used in this text are found in the 2801-word NGSL list. The percentage can be interpreted as your students ought to be able to comprehend 91.72% of the words in this talk providing, they know all 2801 words in the NGSL at the spoken receptive level. For a teacher checking the appropriateness of this TED talk for classroom use, however, their focus ought to be on the cumulative total column. It shows that if learners can decipher all proper nouns and numbers (3.76%) and know all words in the NGSL, they may have a reasonable chance of comprehending this text (i.e., the cumulative percentage is 95.48%). On the other hand, this text would not be appropriate for extensive listening or listening fluency training as both require >98% and 100%, respectively. Further down the analysis page, a color-coded map shows which words come from the different NGSL-related word lists (Figure 6). Interestingly, the words colored black represent the proper nouns, numbers, and words unlisted in any frequency list (3.35% coverage). One could argue that some of these words, for example, polyglots, imitating, ingenious and gibberish would be unknown to most learners. Hence, the combination of these unknown off-list words and an unstable knowledge of the entire NGSL, would lead a teacher to conclude that this text would be too difficult for their students to comprehend. This text would only be appropriate as a listening exercise for advanced students because, Japanese students studying English at Universities in Japan are reported to (a) struggle with spoken input, (b) their spoken receptive vocabulary sizes are relatively low (see Milliner & Dimoski, 2019; Milliner

& Dimoski, 2021; Mizumoto & Shimamoto, 2008), and (c) the volume of potentially unknown unlisted words in this text is quite high. Nevertheless, in situations where this text may be appropriate, the NWLC provides useful assistance to teachers for effectively utilizing this text in their classrooms. A teacher could pre-teach or provide a glossary of the beforementioned unlisted words. Moreover, a teacher could click on the Word List tab in NWLC to access a word frequency report and a list of Auto-extracted keywords (Figure 7). These auto-extracted keywords could also be incorporated into pre-listening vocabulary instruction or schema-building activities to help give learners a better chance of comprehending the video.

Figure 6

Color-coded items from the different frequency lists

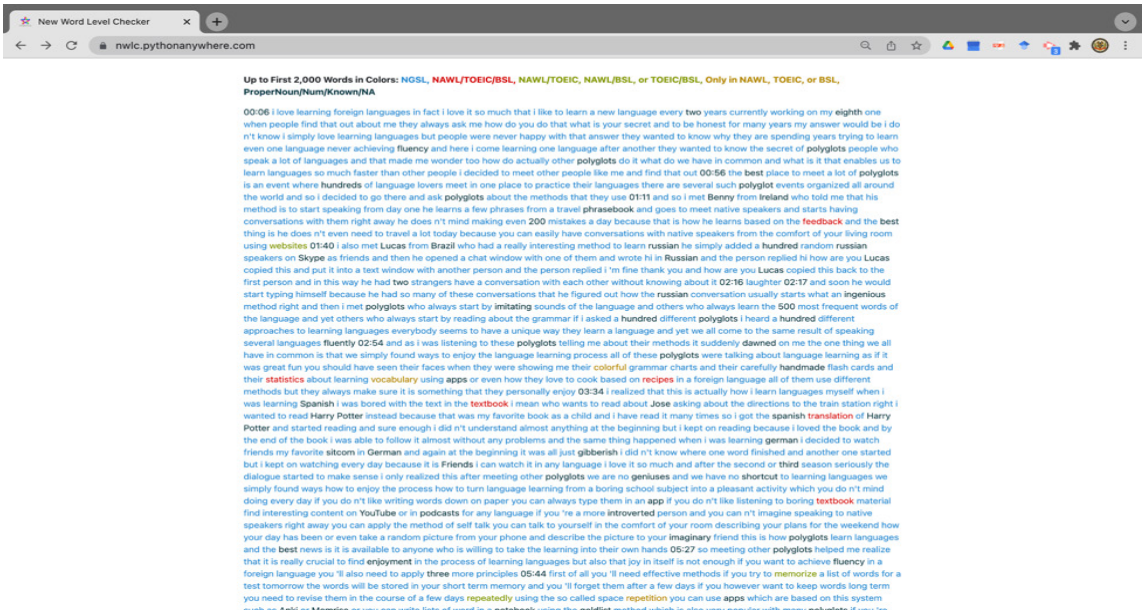
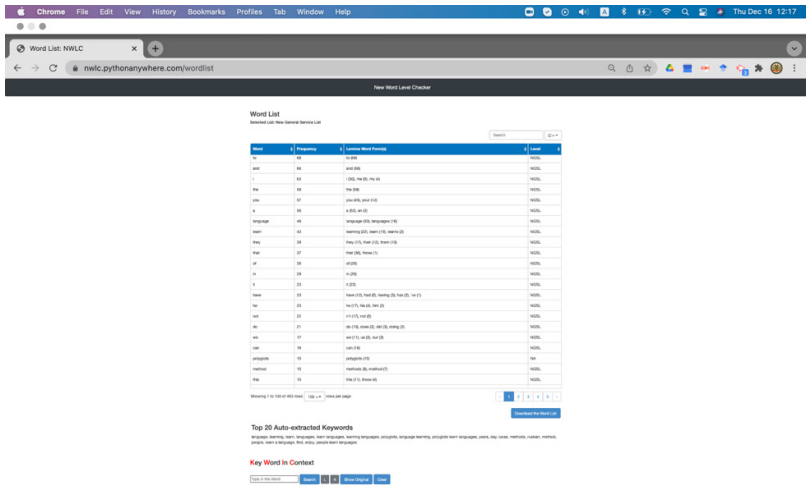


Figure 7

Word List and Auto-extracted Keywords reports from NWLC



5. CONCLUSIONS

Before discussing the conclusions, it is important to note that knowing 95~100% of the words in a text does not guarantee comprehension. The difficulty of spoken or written texts extends beyond lexical knowledge as learners draw from grammatical, metacognitive, and background knowledge for comprehension. Nevertheless, the coverage comprehension model provides teachers with a reasonably objective benchmark for evaluating the difficulty of teaching materials, and the NWLC represents an efficient and accurate tool for this purpose. In addition, the NWLC helps teachers identify potential areas of difficulty so they can provide the necessary support for comprehension. Outside of appraising the difficulty or suitability of prospective teaching materials, the NWLC can be used to survey the lexical coverage of student writing or check for overused vocabulary. Teachers interested in assessing their learner's lexical knowledge should visit Vocableveltest.org, which provides free, customizable vocabulary levels tests (see Milliner, 2022 for a description on how to use this site). With more level-appropriate materials, teachers stand to have a much greater impact on their students' learning.

REFERENCES

Boon, A. (2015). *Zombies in Tokyo*. Atama-ii Books.

Browne, C., Culligan, B., & Phillips, J. (2013). *The New General Service List*. http://www.newgeneralservicelist.org.

Durbahn, M., Rodgers, M., & Peters, E. (2020). The relationship between vocabulary and viewing comprehension. *System, 88*. https://doi.org/10.1016/j.system.2019.102166

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

Machová, L. (2018). *The secrets of learning a second language* [Video]. TED Conferences. https://www.ted.com/talks/lydia_machova_the_secrets_of_learning_a_new_language/transcript?language=en

Mizumoto, A. (2021). *New Word Level Checker* [Web application]. https://nwlc.pythonanywhere.com/

Mizumoto, A., Pinchbeck, G., & McLean, S. (2021). Comparisons of word lists on New Word Level Checker. *Vocabulary Learning and Instruction, 10*(2), 30-41. https://doi.org/10.7820/vli.v10.2.mizumoto

McLean, S. (2021). The coverage comprehension model, its importance to pedagogy

and research, and threats to the validity with which it is operationalized. *Reading in a Foreign Language, 33*(1), 126-140. http://hdl.handle.net/10125/67396

McLean, S., & Raine, P. (2018). VocabLeveltest.org. [Online program]. https://www.vocableveltest.org/

Milliner, B. (2022). Create your own vocabulary levels' test with vocablevelstest.org. *The Language Teacher, 46*(1), 33-35.

Milliner, B. (2021). The effects of combining timed-reading, repeated-oral-reading, and extensive-reading. *Reading in a Foreign Language, 33*(2), 191-211. http://hdl.handle.net/10125/67400

Milliner, B., & Dimoski, B. (2019). Explicit listening strategy training for ELF learners. *The Journal of ASIA TEFL, 16*(3), 833-859. http://dx.doi.org/10.18823/asiatefl.2019.16.3.5.833

Milliner, B., & Dimoski, B. (2021). Effects of a metacognitive intervention on lower-proficiency EFL learners' listening comprehension and listening self-efficacy. *Language Teaching Research*. https://doi.org/10.1177/13621688211004646

Mizumoto, A., & Shimamoto, T. (2008). A comparison of aural and written vocabulary size of Japanese EFL university learners. *Language Education & Technology, 45*, 35–51.

Mizumoto, A., Pinchbeck, G., & McLean, S. (2021). Comparisons of word lists on New Word Level Checker. *Vocabulary Learning and Instruction, 10*(2), 30-41. https://doi.org/10.7820/vli.v10.2.mizumoto

Milton J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse, & M. Torreblanca-López (Eds.), *Further insights into non-native vocabulary teaching and learning* (pp. 83-98). Multilingual Matters.

Nation, P. (2007). The four strands. *Innovation in Language Learning and Teaching, 1*(1), 2-13. https://doi.org/10.2167/illt039.0

Noreille, A., Kestemont, B., Heylen, K., Desmet, P., & Peters, E. (2018). Vocabulary knowledge and listening comprehension at an intermediate level in English and French as foreign languages. *International Journal of Applied Linguistics, 169*(1), 212–231. https://doi.org/10.1075/itl.00013.nor

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*(1), 26-43. https://

doi.org/10.1111/j.1540-4781.2011.01146.x

Stoekel, T., McLean, S., & Nation, P. (2020). Limitations of size and levels tests of written receptive vocabulary knowledge. *Studies in Second Language Acquisition, 43*(1), 181-203. https://doi.org/10.1017/S027226312000025X

Stæhr, L. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 31*(4), 577-607. https://doi.org/10.1017/S0272263109990039

van Zeeland, H. (2013). L2 vocabulary knowledge in and out of context - Is it the same for reading and listening? *Australian Review of Applied Linguistics, 36*(1), 52–70. https://doi.org/10.1075/aral.36.1.03van

van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics, 34*(4), 55-88. https://doi.org/10.1093/applin/ams074