OPEN ACCESS

Vocabulary Learning and Instruction

Volume 12, Number 1 (2023)
https://doi.org/10.29140/vli.v12n1.1210

Castledown

Teaching Practices and Perspectives Regarding Word Counting Units

Louis Lafleur
Kwansei Gakuin University, Japan
louislafleur333@gmail.com

Abstract

The selection of an appropriate word counting unit (WCU) for the purpose of second/foreign language vocabulary acquisition (SLVA) in the last decade has become a very important and relevant topic in academic circles. However, few studies address on-the-ground teaching practices and perspectives. This mixed-approach study, conducted through an online questionnaire, collected both quantitative and qualitative data from 30 ESL/EFL teachers in Japan. The questionnaire surveyed their vocabulary teaching experience and views regarding five categories of WCUs: all word tokens (i.e., All Forms and Meanings (AFM)), plemma, lemma, flemma, and Word Family (WF). The online survey revealed that most participants (i.e., those who were not aware of SLVA research and recent teaching practices) had limited experience and knowledge regarding WCUs and their impact. Another important finding is that although the WCU category participants were most experienced with is the WF (i.e., root/basic word lists), the WCU they have the highest pedagogical regard for is the lemma which considers parts of speech (POS) in vocabulary instruction and learning, followed closely by the plemma which additionally considers the various meanings words can take on, and also irregular inflectional patterns (e.g., irregular verb past tenses and past participles), and to a lesser extent the flemma which does not consider POS nor the various meanings and irregular inflectional patterns of words in vocabulary instruction.

Keywords: Word counting unit, word family, flemma, lemma, plemma, ESL/EFL in Japan

Data Availability Statement: All relevant data are within this paper.

Background

The word family (i.e., typically referred to as the 6th level of the word family system) as defined by Bauer & Nation (1993) has for long been the word count unit yardstick for vocabulary list-building projects such as the General Service List (GSL; West, 1953) and Academic Word List (AWL; Coxhead, 2000). The GSL and AWL are solely composed of root/base words (i.e., word families) such as "drink" and do not include additional word forms or items such as "drank, drinkable, drinker, drinking, drinks, drunk, drunker, drunkest, drunkenly, non-drinker, undrinkable, etc..." within their wordlists as these are derived forms of the base/root word. In the past, it was deemed unnecessary to add supplementary items/entries in vocabulary lists to account for various forms base words may take on. This is because it was assumed that learners could inherently infer the meaning of derived/inflected forms (Ishii et al., 2021; see Tables 1 & 2) of acquired base words without the need for additional study.

However, recent studies such as McLean (2018); Stoeckel et al. (2020) question the pedagogical appropriateness of the word family as a counting unit for lower-level language learners, and point to other WCUs such as the flemma which presents/counts derivational forms of words as separate entries or the lemma which additionally considers variance in parts of speech as separate entries as being more appropriate alternatives as they permit derivational word forms to be included as additional items in wordlists (see Appendix A).

In terms of computational linguistics, it is quite challenging to tag words automatically beyond their orthographic form. Gardner (2007) noted morphological relationships between words, polysemy and homonymy, and multiword items as difficulties encountered in the field. Although there is still room for advancement in computational linguistics such as a need for reliable automated polysemy/homonymy analysis, there has been much progress in software such as the Tool for the Automatic Analysis of Lexical Sophistication (TAALES; Kyle et al., 2018). WCUs such as the flemma and lemma are now being recognized and tagged automatically with much

Table 1 *The Eight Inflectional Word Forms*

#	POS	Inflection description
(1)	nouns	possessive case -'s
(2)	nouns	plural - (e)s or irregular
(3)	verbs	past tense - (e)d or irregular
(4)	verbs	present participle - ing
(5)	verbs	past participle -ed/en or irregular
(6)	verbs	3rd person singular -(e)s
(7)	adjectives	comparative -er
(8)	adjectives	superlative -est

Note. Inflections are limited to suffixes; POS = Part of Speech.

Table 2 *Derivational Word Form Examples (root word "drink")*

Category	Derivational affix	Final word form
noun	-er	drinker
adjective	-able	drinkable
adverb	-en (adj) + -ly (adv)	drunkenly
prefix	non-	non-drinker
prefix & suffix	un- + -able	undrinkable

Note. Example derivational affixes outside of the eight inflections.

greater reliability than before (Kaushal, R., & Chadha, R., 2022), with tools such as Google's Cloud Natural Language API. This has permitted the creation of many corpus-based wordlists published in the last decade (see Appendix B).

Previous Research

Derivational word forms have been observed as being more difficult to be understood, acquired, and utilized by non-native learners in SLA research, and thus more focus should be given to the teaching of derived word forms (Schmitt & Zimmerman, 2002). A certain focus may be beneficial as Mochizuki & Aizawa (2000) observed that L2 learners' affix knowledge (i.e., receptive derivational form knowledge of words) correlated with their overall vocabulary size (r = 0.650). In terms of their productive derivational knowledge, Iwaizumi & Webb's (2021) results showed differences between L1 speakers and L2 learners from various educational levels. Ward & Chuenjundaeng (2009) in their study which asked L1-Thai learners to write L1 translations for English L2 word items showed that participants could not provide correct derivational translations for 49% of the base words they had successfully translated. McLean (2018) in his study that enrolled 279 Japanese university students showed their ability to comprehend associated derivational forms to be very limited as the number of correct responses was significantly lower than that of base words and inflectional forms, and thus recommended that the flemma WCU as a more appropriate word counting unit for second/foreign language students who could not infer inflectional/derivational forms at a satisfactory level. Additionally, Stoeckel et al. (2020) extended McLean's (2018) study to verify the appropriateness of the flemma WCU for Japanese university students and discovered significant shortcomings in their ability to not only understand derived word forms but also differentiate between the various Parts of Speech (POS) words may present within contextualized sentences (e.g., pause "noun" as opposed to pause "verb"). From their results, they concluded that although the flemma word counting unit is a step in the right direction, most Japanese learners would benefit more from utilizing the lemma word counting unit which is utilized in contextualized corpus lists such as the Corpus of Contemporary American English (COCA; Davies, M. 2010) frequency list which tags parts of speech (e.g., drink (verb) [frequency item#1238], and drink (noun) [frequency item#2040]).

In a follow-up collaborative paper (Brown et al., 2022), it was concluded that the selection of a lexical unit should align with learners' abilities. In other words, the flemma should be selected for learners who cannot infer derivational forms of base words satisfactorily, and the lemma for learners who cannot infer derivational forms nor differentiate the various POS words may take on satisfactorily. The previously mentioned studies question the appropriateness of the Word Family and to a lesser extent the flemma as the base counting/testing/learning word unit to assess Japanese students' fluency with certainty as they could not infer all the forms. However, there is a lack of research assessing whether teachers are aware or not of their students' ability to infer the meaning of various word forms, and if they have reflected upon what type of word counting unit would be appropriate for them. The following research questions were formulated to reveal on-the-ground teaching practices, perspectives regarding WCUs and their pedagogical implications:

1. On a 5-point Likert satisfaction scale, how do teachers in Japan rate the various word counting units they have utilized in their teaching?
2. What are teachers' opinions of the various types of word counting units in terms of their appropriateness for students in Japan?
3. What are the teachers' perspectives about including additional word forms (e.g., gerunds and irregular verb patterns) as additional vocabulary/study items?

Methodology

For this study, a mixed-methods approach was implemented to collect both quantitative Likert-scale responses and qualitative open-ended responses from participants through a google forms survey within online teacher/researcher circles, mainly the Japan Association for Language Teaching (JALT), and its vocabulary Special Interest Group (SIG). Semi-structured interviews were initially considered to follow-up to the survey; however, this avenue was not pursued after initial results revealed that the majority of participants had limited experience and knowledge regarding the topic, and would not be able to contribute much beyond their original survey responses.

Participants

Participation was voluntary, and 30 teachers with Teaching English as a Second/Other Language (TESOL)/Teaching English as a Foreign Language (TEFL) experience within Japan volunteered about an hour of their time to complete the survey. The range of participants covered both private and public institutions and all education levels from elementary to college/university (see Table 3). Most of the participants (73.3%) were teaching at the university level, and only 16.7% had research experience related to vocabulary acquisition specifically. Three of the participants were teaching in more than one teaching context which combined elementary/primary education with others.

During the design phase of this study, there were legitimate concerns about enrolling participants who did not have a vocabulary research background to provide

Table 3 *Participant TESOL / TEFL Teaching Context Responses*

Teaching Context	Total # 30 (100%)	Education Researcher #	Vocabulary Researcher #
College/University	22 (73.3%)	11 (36.7%)	5 (16.7%)
Retired/Former TESOL/TEFL	2 (6.7%)		
Junior High School (JHS)	2 (6.7%)		
Private Language School (PLS)	1 (3.3%)		
ES, College/University, PLS	1 (3.3%)		
ES, JHS, SHS, PLS	1 (3.3%)		
ES, PLS	1 (3.3%)		

Note. ES = Elementary/Primary School, JHS = Junior High School, SHS = Senior High School, PLS = Private Language School.

trustworthy information about a topic as specific as word counting units in second/foreign language vocabulary teaching and learning. This was in-part by design as the goal of this paper was not to collect exclusively vocabulary researcher opinions (i.e., 16.7% of total participation) but teacher opinions (i.e., 100% of total participation). However, in an effort to ensure a higher degree of reliability, it was decided to bridge the gap as much as possible between experts and non-experts by providing a "Word Counting Unit Informational Reference" to both inform and guide the participants before and during the survey (see Sect. 3.2 & Appendix A).

Word Counting Unit Informational Reference

The base word "drink" was selected by the author to inform the participants about word counting units as it is a prime candidate in representing the complexity of itemization. Firstly, the homonymic nature between POS: noun, adjective, verb "drink, drunk". Secondly, it follows an irregular pattern of verb inflections "drink, drank, drunk". If a single base word item (i.e., root/base word item) "drink" was paired with a selection of 12 (uncontextualized) forms: drank, drink, drinkable, drinker, drinking, drinks, drunk, drunker, drunkest, drunkenly, non-drinker, undrinkable, different numbers of items/entries could be generated for various types of word counting units (see Appendix A).

Analysis Procedures

Participants were asked to rate the various word counting unit categories under a 5-point Likert scale and select yes/no answers to questions related to their experience and future consideration.

Participants could optionally provide additional comments to explain in more detail the scores or yes/no answers they provided. To explore these responses more robustly, Braun & Clarke's (2006) 6-phase guide to thematic analysis approach/framework was utilized to find, analyze and classify participant qualitative survey responses into a data

extract (i.e., thematically coded chunk of data). For the purposes of this study, a theoretical essentialist/realist top-down approach to thematic analysis was conducted, and thus comments solely related specifically to word counting units were retained (i.e., focus was given to participants' individual experiences/motivations that were related to the paper's research topic).

Results

The following Tables showcase the data collected from the survey/questionnaire (see Appendix C), and provide the basis for answering the three research questions in the following discussion section of this paper. Information related to RQ1 can be found in Tables 4 & 5, RQ2 can be found in Tables 6–8, and RQ3 can be found in Table 9. Table 10 presents a tally of qualitative comments classified under positive, neutral, and negative were utilized to supplement the participants' quantitative/binary survey responses.

Table 4 shows the participants' teaching experience with the following five word counting unit categories under: yes, no, and other responses. Other responses enabled participants to write in a response other than yes or no, these were used to express uncertainty such as "Not sure", "I'm not sure really", and "Probably not". The most utilized WCU for participants in regard to their teaching experience was the Word Family and the least utilized was the AFM.

Table 5 shows the participants' relative satisfaction teaching the word counting units they have experience teaching with. Although some WCU categories have fewer responses, an overall satisfaction ranking was calculated with the available data. Considering the low amount of quantitative data collected here, results were verified and supplemented with follow-up qualitative survey responses (see Appendix C; question 2.4) when answering the first research question in the discussion section of this paper.

Table 6 shows the participants' perceived average student proficiency level in terms of The Test of English for International Communication (TOEIC) Listening and Reading (L&R) score and equivalent level of Common European Frame of Reference for Languages (CEFR). It should also be noted that seven participants did not know or abstained from providing their average student proficiency level.

Table 7 shows the participants' perceived appropriateness of the various word counting unit for their average student. This survey question was optional because of its perceived difficulty; 21 participants out of 30 (70%) provided responses. In contrast to Table 5, all responding participants were asked to provide ratings for all WCUs including those they had not utilized in their teaching.

Table 4 *Participant Word Counting Unit Teaching Experience (n = 30)*

	(1) AFM	(2) Plemma	(3) Lemma	(4) Flemma	(5) Word Family
Yes	2 (6.7%)	6 (20%)	6 (20%)	5 (16.7%)	11 (36.7%)
Other	1 (3.3%)	1 (3.3%)	1 (3.3%)	1 (3.3%)	2 (6.7%)

Note. AFM = All Forms and Meanings or all word tokens.

Table 5 *Participant Word Counting Unit Satisfaction and Overall Rank*

WCU n = # of responses	1	2	3	4	5	Mean SD	Overall Rank
(1) AFM n = 2	0	0	1 (50%)	1 (50%)	0	M = 3.5 SD = 0.71	3rd
(2) Plemma n = 6	0	0	1 (16.7%)	4 (66.7%)	1 (16.7%)	M = 4 SD = 0.63	2nd
(3) Lemma n = 7	0	0	0	5 (71.4%)	2 (28.5%)	M = 4.29 SD = 0.49	1st
(4) Flemma n = 5	0	0	3 (60%)	2 (40%)	0	M = 3.4 SD = 0.55	4th
(5) Word Family n = 9	0	4 (44.4%)	2 (22.2%)	2 (22.2%)	1 (11.1%)	M=3 SD = 1.12	5th

Note. Rank established by the author with the available data. 5-point Likert scale, (1) = poor, (5) = excellent (satisfaction); AFM = All Forms and Meanings or all word tokens; SD= Standard Deviation.

Table 6 *Participants' Perceived Average Student Proficiency Level (n = 30)*

Level Category	TOEIC L&R score	CEFR level	CEFR description	# of responses
(1)	120–220	A1	Basic-Breakthrough	4
(2)	225–385	A2-1	Basic-Waystage user	4
(3)	385–545	A2-2	Waystage user	10
(4)	550–780	B1	Independent-Threshold user	5
(5)	785–940	B2	Independent-Vantage user	0
(6)	945–990	C1	Proficient user	0
Other/No response				7

Table 8 shows the participants' future aspirations to utilize the various word counting units. Interestingly, results did not closely follow WCU satisfaction ratios and ranks (see Table 5). Other responses enabled participants to write in a response other than yes or no, these were used to express uncertainty such as "Probably not (maybe)".

Table 9 shows the participants' opinions about including some word forms as additional study items.

Table 10 shows a tally of qualitative comments (i.e., optional/additional comments in regard to WCUs) classified into positive, neutral, and negative data extracts (i.e., thematically coded chunks of data) as outlined in Braun & Clarke's (2006) thematic analysis framework. Comments were optional and given as a follow-up to quantitative/binary survey responses. It should be noted that the following table in its tally gives

Table 7 *Perceived Appropriateness of Word Counting Units for Participants' Average Student (n = 21)*

	(1)	(2)	(3)	(4)	(5)	Mean SD
(1) AFM	9 (42.8%)	3 (14.3%)	3 (14.3%)	3 (14.3%)	3 (14.3%)	M = 2.43 SD = 1.54
(2) Plemma	5 (23.8%)	5 (23.8%)	2 (9.5%)	4 (19.05%)	5 (23.8%)	M = 2.95 SD = 1.56
(3) Lemma	4 (19.05%)	2 (9.5%)	8 (38.1%)	3 (14.3%)	4 (19.05%)	M = 3.05 SD = 1.36
(4) Flemma	7 (33.3%)	4 (19.05%)	4 (19.05%)	4 (19.05%)	2 (9.5%)	M = 2.52 SD = 1.40
(5) Word Family	10 (47.6%)	2 (9.5%)	0 (0%)	4 (19.05%)	5 (23.8%)	M = 2.62 SD = 1.77

Note. 5-point Likert scale, (1) = poor, (5) = excellent (perceived appropriateness); AFM = All Forms and Meanings or all word tokens; SD = Standard Deviation.

Table 8 *Participant Future Aspiration to Utilize the following Word Counting Units (n = 30)*

Response	(1) AFM	(2) Plemma	(3) Lemma	(4) Flemma	(5) WF
Yes	17 (56.7%)	19 (63.3%)	18 (60%)	14 (46.7%)	15 (50%)
Other	3 (10%)	5 (16.7%)	3 (10%)	2 (6.7%)	8 (26.7%)

Note. AFM = All Forms and Meanings or all word tokens.

Table 9 *Participant Opinion about Including some Words Forms as Additional Study Items (n = 30)*

	Additional item	Not separate	Other/no opinion
(1) Should verbal nouns/gerunds (Ex: I (dis)like drinking) be considered as a kind of inflectional affix or as an additional/separate study item (under plemma and lemma)?	10	14	6
(2) Should irregular patterns of verb inflections (e.g. drink, drank, drunk) be considered as a kind of inflectional affix or as an additional/separate study item?	5	16	9
(3) Should hyphenated prefixes (e.g., "non-" in non-drinker) be considered as a regular "derivational prefix/affix" (separate item) or not regular (not a separate item)?	4	16	10

Table 10 *Tally of Qualitative Comments Classified under Positive, Neutral, and Negative*

WCU	#positive	#neutral	#negative	Total	Rank
(1) AFM	7 (31.82%)	5 (22.73%)	10 (45.45%)	22	5th
(2) Plemma	10 (66.67%)	3 (20%)	2 (13.33%)	15	1st
(3) Lemma	7 (41.18%)	6 (35.29%)	4 (23.53%)	17	3rd
(4) Flemma	10 (52.63%)	3 (15.79%)	6 (31.58%)	19	4th
(5) WF	7 (38.89%)	9 (50%)	2 (11.11%)	18	2nd

Note. AFM = All Forms and Meanings or all word tokens.

equal weighting to all comments no matter how positive or negative they were. The most surprising result was how well the WF category ranked overall in terms of qualitative responses considering how low it scored quantitively; participants noted some specific uses for WF lists, for example: "Especially when there are a lot of words to learn, learning only headwords one from each word family is practical to prevent the number of target words to learn being overwhelming. Suitable for higher level learners. (P#8)"

Discussion

The first research question explored the participants' experience and satisfaction in regard to the utilization of the various word counting units in their teaching. The All Forms and Meanings (AFM) category was the least utilized and the word family was the most utilized word counting unit for the participants (see Table 4). This is not surprising as word family unit lists such as the GSL/AWL were the first publicly available frequency-sequenced lists and are still very popular to this day. Only two participants had utilized an AFM in their teaching with one providing the following comment: "I generally write an AFM of my own (using dictionaries) and let students copy it into their notebooks (first-year university classes, elementary level students). [However,] Most of the students do not use many of the [word] forms in conversation. (P#16)". Overall, the collected data suggest that the participants have very limited experience with WCUs in general: 33% had two or more "yes" responses, 27% had only one "yes" response, and 40% had none which may perhaps be indicative of them not specifically attending to the teaching and learning of vocabulary or unsure which WCU they had been using.

In terms of rating their WCU experience, the word family held the lowest satisfaction, and the lemma the highest (see Table 5). Overall, WCUs which focused on study depth (i.e., the profound study of multiple word forms/meaning; e.g., AFM and plemma) were rated higher than WCUs that focused on study breadth (i.e., the quick yet limited study of word forms; e.g., flemma and word family). Comments revealed that participants appreciated the lemma and plemma as they were attached to contextualized lists/corpora: "The COCA shows words in context [POS tagged], which is a lot more useful than other vocabulary lists. (P#13)" and "As long as contextual

examples are provided, they seem to help students remember the vocabulary. (P#16)". As for the flemma, important advantages for teachers and researchers were noted: "Flemma has higher practical value in labeling, with no need of human checking, and should be useful to choose items and classify from a large number of items. (P#8)" and "I'll go with a 4 [Likert satisfaction score] because flemma lists are very convenient for text analysis [...]. Also, I would like to believe that learners can handle (or can learn to handle) the assumptions made in the flemma, i.e., that word forms can cross a part-of-speech boundary. (P#22)". Finally for the word family, mostly positive or balanced comments were provided: "Especially when there are a lot of words to learn, learning only headwords one from each word family is practical to prevent the number of target words to learn being overwhelming. Suitable for higher-level learners. (P#8)" and "Personally I feel it is more beneficial for students to learn words in context. Using word family lists can be beneficial for students when doing their own study. (P#17)".

The second research question explored the participants' opinions of the various types of word counting units regarding their appropriateness for students in Japan. Before providing an answer to this research question, the participants were asked to assess the average proficiency of their students in the survey (see Table 6). The most common response was 385-545 TOEIC L&R, CEFR level A2-2 waystage user (3) which represents 33% of total responses. There were no recorded responses for the two highest proficiency categories independent-vantage user (5) and proficient user (6).

After completing this assessment, participants were asked to rate the various types of word counting units and their appropriateness for their average student (see Table 7). The highest-rated category was the lemma (3) and the plemma (2) came in a close second. Some participants additionally wrote about their appreciation and preference in using contextualized word lists with tagged parts of speech. However, the plemma, similar to AFM also received its share of criticism related to the fact that these WCUs can lead to huge lists that may appear overwhelming to students: "Looks too complicated, even at first glance (P#3)", and thus could be perhaps a little reduced: "Less useful entries (such as the moth [drinker] and the inflected adjectives) can be omitted while more useful ones (irregular verb forms) [could be] included. (P#20)". Although participants may have lower satisfaction with some WCUs such as AFM, they may still aspire to use it for specific purposes: "I think that for some high-frequency words, an AFM [breakdown] could be interesting as a consciousness-raising exercise and possibly useful in teaching students how to correctly use dictionaries (i.e., to search for the POS, idiomatic uses etc...) (P#6)." and "I believe it is a good way to introduce new word families and how the forms vary… (P#16)". In terms of future aspirations to utilize the various WCUs in their teaching, the participants answered direct yes/no responses (see Table 8). The category which had the highest "yes" accounts were the plemma (2), followed closely by the lemma (3) category. The categories which had the highest "no" accounts were the flemma (4) and word family (5).

The third research question explored the participants' opinions about including some word forms (e.g., gerunds and irregular verb patterns) as additional vocabulary/study items. This research question and specific questions (see Table 10) were inspired by the author's personal pedagogical inquiries regarding WCUs. Most participants did not see the need to include irregular verb patterns (2) and hyphenated prefixes

(3) linked to words as additional items. However, some participants who disagreed with these additions did comment that specialized irregular verb patterns (2) lists could be useful for some specific students, and also suggested that words with prefixes (3) and even suffixes should be retained if they are frequent: "…there is nothing wrong with including frequent lexical items that include affixes in a vocabulary list. Don't worry about whether they are written with a hyphen or not. A hyphen in the written form of a word is just a stylistic choice, and is not linguistically meaningful. "Non-drinker" could also be "nondrinker" and "non-human" could be "nonhuman". (P#14)" or perhaps they should just be taught as additional concepts and not as additional items: "Perhaps 'non-' could be given its own presence on a word list. I think that might be the case with the NGSL/NAWL. (P#22)". Verbal nouns/gerunds (1) as additional items were also rejected but to a lesser extent. This was surprising as the lemma WCU was arguably the most popular among the participants and the well-known lemmatized COCA 5000 lists verbal nouns/gerunds as additional items (e.g., learning [#1462], reading [#1591], drinking [#3470]).

Limitations and Future Directions

Although this exploratory study provides a glimpse into teachers' opinions of word counting units within the Japanese TESOL context, it is not without its own limitations. First, the low number of participants limited the scope of the analysis this study could take on. A larger number of participants could enable comparisons between the participants (e.g., comparing responses from teachers who have a background in vocabulary research to those who do not). Second, additional specific questions about participants' knowledge of WCU would have been helpful (e.g., Before reading the Word Counting Unit Informational Reference (Appendix A) in the survey, were you aware of the effect WCUs can have on vocabulary learning and teaching?). Third, the concept and pedagogical implications of word counting units in TEFL was difficult for some of the participants to fully grasp, and thus perhaps more effort to clearly convey the various WCUs and their impact should be put forth. Finally, it would be interesting if future studies on this topic could report on other ESL/EFL teaching contexts and implement confirmatory research designs.

Conclusions

Choosing an appropriate word counting unit is not so different from choosing a political alignment as it also ranges from the very left all forms and meanings (1) for profound yet time-consuming learning to the very right word family (5) for quick yet shallow study. In terms of vocabulary list study efficiency, researchers such as McLean (2018), Stoeckel et al. (2020), and Brown et al. (2022) recommend identifying/choosing a word counting unit as far right as possible on the scale as long as the student(s) can infer the forms and meanings of the word counting units to its left at a highly satisfactory level to uphold fluency. However, this recommendation may be difficult to implement as the survey data revealed that not all the participants shared the same understanding and view of word counting units. Many participants commented that they had never considered the question of word counting units before completing the

survey. Participants who identified themselves as vocabulary researchers noted the advantage of the flemma as a unit that can be more easily recognized and counted in computational linguistics. Moreover, a few of the participants admitted in their survey responses that they do not expect their (lower-level) students to master the various forms and meanings of words, and thus prefer only focusing on the teaching/learning of base root words. That being said, the overall majority of the participants mostly agreed with the previously mentioned researchers as they identified the lemma (3) as their preferred word counting unit followed by the plemma (2) and to a lesser extent the flemma (4), and most importantly showed flexibility in their WCU beliefs (e.g., "My hope is that flemma lists turn out to be the way to go, but if research suggests this not to be the case, I'd switch to either lemmas or failing that, plemmas (P#23)".

Acknowledgments

The author would like to thank all those who contributed their advice and time to this paper.

About the Author

Louis Lafleur is a lecturer at Kwansei Gakuin University conducting research related to the fields of second language vocabulary acquisition, cognitive psychology, computer-assisted language learning, and game-informed language learning (gamification).

Ethical Statement and Competing Interests

Informed consent was obtained from all individual participants involved in the study. The author declares no conflicts of interest.

References

Bauer, L., & P. Nation. (1993). Word families. *International Journal of Lexicography* 6: 253–279. https://doi.org/10.1093/ijl/6.4.253

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. Applied Linguistics, *36*(1), 1–22. Retrieved from: https://doi.org/10.1093/applin/amt018

Brown, D., Stoeckel, T., Mclean, S., & Stewart, J. (2022). The most appropriate lexical unit for L2 vocabulary research and pedagogy: A brief review of the evidence. *Applied Linguistics*, *43*(3), 596–602. https://doi.org/10.1093/applin/amaa061

Browne, C., Culligan, B., & Phillips, J. (2013a). *New General Service List (NGSL)*. Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. Retrieved from: https://www.newgeneralservicelist.com/new-general-service-list

Browne, C., Culligan, B., & Phillips, J. (2013b). *New Academic Word List (NAWL)*. Licensed under a Creative Commons Attribution-ShareAlike

4.0 International License. Retrieved from: https://www.newgeneralservicelist.com/new-general-service-list-1

Browne, C., & Culligan, B. (2016a). *TOEIC Service List (TSL)*. Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. Retrieved from: https://www.newgeneralservicelist.com/toeic-service-list

Browne, C., & Culligan, B. (2016b). *Business Service List (BSL)*. Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. Retrieved from: https://www.newgeneralservicelist.com/business-service-list

Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, *34*(2), 213–238. https://doi.org/10.2307/3587951

Coxhead, A., & Hirsh, D. (2007). A pilot science-specific word list. Revue Française de Linguistique Appliquée, *12*(2), 65–78. Retrieved from: https://www.cairn.info/load_pdf.php?ID_ARTICLE=RFLA_122_0065&download=1&from-feuilleteur=1

Dang, T. N. Y., & Webb, S. (2016). Making an essential word list. In I. S. P. Nation (Ed.), Making and using word lists for language learning and testing (pp. 153–167). Amsterdam: John Benjamins.

Davies, Mark. (2010). *The Corpus of Contemporary American English*. Retrieved from: www.english-corpora.org/coca/.

Gardner, D. (2007). Validating the construct of word in applied corpus-based vocabulary research: A critical survey. *Applied linguistics*, *28*(2), 241–265. https://doi.org/10.1093/applin/amm010

Ishii, T., Bennett, P., & Stoeckel, T. (2021). Challenges in the Assumptions of Using a Flemma-based Word Counting Unit. *Vocabulary Learning and Instruction*, *10*(1), 1–15. https://doi.org/10.7820/vli.v10.1.Ishii

Iwaizumi, E., & Webb, S. (2021). Measuring L1 and L2 productive derivational knowledge: How many derivatives can L1 and L2 learners with differing vocabulary levels produce?. *TESOL Quarterly*, *56*(1), 100–129. https://doi.org/10.1002/tesq.3035

Japan Association of College English Teachers (JACET). (2016). *The New JACET list of 8000 basic words*. Tokyo, Japan: Kirihara Shoten.

Kaushal, R., & Chadha, R. (2022, December). *State of the Art, Recent Developments, and Future Directions in Applying Deep Learning to Part of Speech Tagging in NLP*. In 2022 International Conference on Computational Modelling, Simulation and Optimization (ICCMSO) (pp. 38–41). IEEE. https://doi.org/10.1109/ICCMSO58359.2022.00021

Kyle, K., Crossley, S., & Berger, C. (2018). The Tool for the Automatic Analysis of Lexical Sophistication (TAALES): version 2.0. *Behavior research methods*, *50*, 1030–1046. https://doi.org/10.3758/s13428-017-0924-4

Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*, *22*, 42–53. https://doi.org/10.1016/j.jeap.2016.01.008

McLean, S. (2018). Evidence for the adoption of the flemma as an appropriate word counting unit, *Applied Linguistics* 39: 823–845. https://doi.org/10.1093/applin/amw050

Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System*, *28*(2), 291–304. https://doi.org/10.1016/S0346-251X(00)00013-0

Nation, I. S. P. (2012). The BNC/COCA word family lists. Retrieved from: https://www.wgtn.ac.nz/__data/assets/pdf_file/0005/1857641/about-bnc-coca-vocabulary-list.pdf

Nesi, H. (2002, August). An English spoken academic word list. In *Proceedings of the tenth EURALEX International Congress* (Vol. 1, pp. 351–358). Copenhagen, Denmark: EURALEX. Retrieved from: https://euralex.org/wp-content/themes/euralex/proceedings/Euralex%202002/036_2002_V1_Hilary%20Nesi_An%20English%20Spoken%20Academic%20Wordlist.pdf

Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know?. *TESOL quarterly*, *36*(2), 145–171. https://doi.org/10.2307/3588328

Stoeckel, T., T. Ishii, & P. Bennett. (2020). Is the lemma more appropriate than the flemma as a word counting unit?, *Applied Linguistics* 41: 601–606. https://doi.org/10.1093/applin/amy059

Ward, J., & Chuenjundaeng, J. (2009). Suffix knowledge: Acquisition and applications. *System*, *37*(3), 461–469. https://doi.org/10.1016/j.system.2009.01.004

West, M. (1953). *A General Service List of English Words*. London:Longman, Green and Co.

Appendix A

Appendix A: Word Counting Unit Informational Reference

above pointed line: (inflectional suffixes) below pointed line: (derivational affixes) Example and/or Definition (Context) with notes		① All Forms & Meanings	② Plemma[1] "Plurality[2][3]"	③ Lemma "Pure[4]" Lemma	④ Flemma "Unpure" Lemma	⑤ Word Family (all related forms)
..to drink water..	PRES	drink (V-1)	drink (V-1)	drink (V)	drink	drink
..he/she/it drinks..	3SG	drinks (V-1)				
..am/are/is/was/were drinking..	PROG	drinking (V-1)				
..drank water yesterday..	PST	drank (V-1)	(*) drank (V-1)			
..have/has/had drunk..	PST PART	drunk (V-1)	(*) drunk (V-1)			
..to drink in the view.. (*expression)		drink (V-2)	drink (V-2)			
a person who drinks too much	SG	drunk (N)	drunk (N)	drunk (N)		
..many beverages..	PL	drinks (N)	drink (N)	drink (N)		
a beverage	SG	drink (N)				
..drinking is.. /..(dis)like drinking..	vN / GER	drinking (N)	(Δ) drinking (N)	(Δ) drinking (N)		
..drunk (person)..		drunk (ADJ)	drunk (ADJ)	drunk (ADJ)		
..(person) gets drunker faster..	COMP ADJ	drunker (ADJ)				
..(person) was the drunkest of all..	SUPL ADJ	drunkest (ADJ)				
a person who drinks	SG	drinker (N-1)	drinker (N-1)	drinker (N)	drinker	
a large brownish European moth	SG	drinker (N-2)	drinker (N-2)			
a container for animals to drink from	SG	drinker (N-3)	drinker (N-3)			
..he/she is a non-drinker..	hyphenated	non-drinker (N)	(?) non-drinker (N)	(?) non-drinker (N)	(?) non-drinker	
..drinkable (beverage)..		drinkable (ADJ)	drinkable (ADJ)	drinkable (ADJ)	drinkable	
..undrinkable (beverage)..		undrinkable (ADJ)	undrinkable (ADJ)	undrinkable (ADJ)	undrinkable	
..to sing drunkenly..		drunkenly (ADV)	drunkenly (ADV)	drunkenly (ADV)	drunkenly	
(1) termed coined by the author, P from Plurality (2) allows a plurality of entries under same PoS	#vocab list items #headwords	20	11-15	8-10	5-6	1
(3) allows (*) irregular inflectional patterns as items (4) distinguishes between PoS (Δ) Should verbal Nouns/Gerunds be considered as a sort of inflectional affix? (debate) (?) Should "hyphenated affixes" such as "non-" in non-drinker be considered as a valid prefix? (debate) (5) WF⑤ & F④ are more easily computer countable (6) for advanced learners, WF = enough to infer forms	Notable Lists	?	Japanese ESL textbooks'index, Also dictionaries	Lemma Corpuses Ex: COCA lemma list	New General Service List (NGSL)	General Service List (GSL)
	Counting Tool(s)[5]	← Man & Computer , Computer →				
	Ensued Study[6]	← Immense/Slow yet Profound , Light/Quick yet Limited →				

The word "drink" was selected for the above reference as it is a prime candidate to represent WCU complexity: homonyms among various Parts Of Speech (POS) noun, adjective & verb "drink, drunk", and irregular verb patterns "drank, drunk".

Choosing a WCU does not constrain a list builder to include all possible underlying forms: if the (2) Plemma is chosen by a list builder, they may choose to only include drinker N-1 (but not N-2 and N-3), but on the other hand include both V1 and V2 for drink. Fundamentally, choosing a WCU, enables and/or restricts the scope of possible entries that the base word can include in a list.

Appendix B

Appendix B: Corpus-based wordlists examples

(Corpus-based word family wordlist examples)

- General Service List (GSL; West, 1953)
- Academic Word List (AWL; Coxhead, 2000)
- Spoken Academic Word List (Nesi, 2002)
- EAP Science Word List (Coxhead & Hirsh, 2007)
- BNC/COCA2000 (Nation, 2012)

(Corpus-based flemma wordlist examples)

- New General Service List (NGSL; Browne et al., 2013a)
- New Academic Word List (NAWL; Browne et al., 2013b)
- TOEIC Service List (TSL; Browne & Culligan, 2016a)
- Business Service List (BSL; Browne & Culligan, 2016b)
- Essential Word List (Dang & Webb, 2016)

(Corpus-based lemma wordlists)

- Corpus of Contemporary American English (COCA; Davies, M. 2010) Frequency List
- Brezina and Gablasova's (2015) New General Service List
- New Medical Academic Word List (Lei & Liu, 2016)
  - The Japan Association of College English Teachers (JACET) list of 8000 basic words (JACET, 2016)

Appendix C

Appendix C: Sample Survey Questionnaire and Questions

(Please note that this survey was originally conducted via Google forms, participants were presented with Appendix A and given a very brief summary of WCU research which was similar to this paper's introduction section)

#1 (Background)
Check the boxes which are currently applicable to you:

- □ (Elementary / Primary) TESOL / TEFL in Japan
- □ (Junior High School) TESOL / TEFL in Japan
- □ (High School) TESOL / TEFL in Japan
- □ (College / University) TESOL / TEFL in Japan
- □ (Private Language School) TESOL / TEFL in Japan
- □ Researcher (Education, Second Language Education or related field)
- □ Researcher (Vocabulary Acquisition, Corpus Linguistics or related field)
- □ Retired/Former TESOL / TEFL teacher (Japan teaching experience)
- □ Other

#2 First WCU Category All Forms & Meanings (1)

#2.1 Have you ever utilized AFM (1) list(s) as teaching/learning material for students?

- ○ Yes
- ○ No
- ○ Other / No response

#2.2 (*If applicable) Please briefly describe your experience using AFM (1) list(s)? (What specific list? / Who were the students? (grade/general level of English)...

#2.3 (*If applicable) Please rate your overall satisfaction using AFM (1) list(s)?
1= Dissatisfied ~ 5= Satisfied

#2.4 (*If applicable / Optional) Please feel free to expand upon / explain your rating?

#2.5 Would you consider using an AFM (1) list for vocabulary teaching in the future?

- ○ Yes
- ○ No
- ○ Other / No response

#2.6 (*Optional) Feel free to expand upon / explain your previous answer?
***The above 6 questions were repeated for the following WCU categories:

#3 Plemma (2)
#4 Lemma (3)
#5 Flemma (4)
#6 Word Family (5)

#7 (Skippable) Which of the following levels of proficiency would best describe the average level of your current students? (choose one)

- ○ 945-990 TOEIC L&R, CEFR level C1 (Proficient user)
- ○ 785-940 TOEIC L&R, CEFR level B2 (Independent-Vantage user)
- ○ 550-780 TOEIC L&R, CEFR level B1 (Independent-Threshold user)
- ○ 385-545 TOEIC L&R, CEFR level A2-2* (Waystage user)
- ○ 225-385 TOEIC L&R, CEFR level A2-1* (Basic-Waystage user)
- ○ 120-220 TOEIC L&R, CEFR level A1 (Basic-Breakthrough)
- ○ Other

#8.1 (Skippable) Try your best to classify these WCU (5) categories from the most appropriate to the least for your current students (overall opinion)?
(1st) most appropriate ~ (5th) least appropriate (selection)

#8.2 (*Optional) Feel free to expand upon / explain your previous answer?

#9.1 (Debate / Opinion) Should verbal nouns / gerunds (Ex: I (dis)like drinking) be considered as a kind of inflectional affix or as an additional/separate study item (under Plemma and Lemma)?

- ○ Not a separate item / a kind of Inflectional affix (one less item to study)
- ○ Separate item / Not a kind of Inflectional affix (additional item to study)
- ○ No opinion

#9.2 (*Optional) Feel free to expand upon / explain your previous answer?

#10.1 (Debate / Opinion) Should irregular patterns of verb inflections (e.g., drink, drank, drunk) be considered as a kind of inflectional affix or as an additional/separate study item?

- ○ Not a separate item / a kind of Inflectional affix (one less item to study)
- ○ Separate item / Not a kind of Inflectional affix (additional item to study)
- ○ No opinion

#10.2 (Optional) Feel free to expand upon / explain your previous answer?

#11.1 (Debate / Opinion) Should hyphenated prefixes (e.g., "non-" in non-drinker) be considered as a regular "derivational prefix/affix" (separate item) or not regular (not a separate item)?

- ○ Not a separate item / Not a regular derivational prefix/affix (one less item to study)
- ○ Separate item / Not a kind of Inflectional affix (additional item to study)
- ○ No opinion

#11.2 (*Optional) Feel free to expand upon / explain your previous answer?

#12.1 (Debate / Opinion) Should hyphenated prefixes (e.g., "non-" in non-drinker) be considered as a regular "derivational prefix/affix" (separate item) or not regular (not a separate item)?

- ○ Not a separate item / Not a regular derivational prefix/affix (one less item to study)
- ○ Separate item / Not a kind of Inflectional affix (additional item to study)
- ○ No opinion

#12.2 (*Optional) Feel free to expand upon / explain your previous answer?