Anna Reini*

Knowledge of receptive English morphology among Finnish university students: L2 learners' morpheme recognition and its implications to vocabulary learning

https://doi.org/10.1515/eujal-2024-0058

Abstract: Morphological knowledge involves understanding how morphemes, the smallest meaningful units of words, can be used to form words. However, not much is known about L2 learners' English morpheme recognition in multimorphemic words. To the best of the author's knowledge, no research exists on Finnish university students' knowledge of English multimorphemic words. Finnish is a Finnic language of the Uralic language family, and as a non-Germanic language, its structures are very different from those of, e.g., English. The highly productive and rich morphological system of Finnish provides novel opportunities for research. The present study examines morpheme recognition, the most crucial aspect of receptive morphological knowledge needed for vocabulary acquisition, and intralexical factors that affect morpheme recognition. The participants, 694 Finnish-speaking university students, were administered word segmentation and vocabulary knowledge tasks, which were developed to fit the proficiency level of the Finnish participants. Using both Pearson and corrected item-total correlational analyses, the findings indicated that recognizing the root of a word was influenced significantly by the ability to break multimorphemic words into parts. The results add to our understanding of how morphological knowledge is constructed by advanced L2 English learners, as well as provide novel evidence for its receptive use in strategic vocabulary learning. This evidence highlights the importance of drawing learners' attention to different morphological structures and the semantic role of roots. Based on the results, the pedagogical implications for L2 vocabulary learning are also addressed.

Keywords: L2 vocabulary learning, morphological knowledge, derivational morphology, word formation

***Corresponding author: Anna Reini,** University of Jyväskylä, Department of Language and Communication Studies, Centre for Multilingual Academic Communication, 40014 Jyvaskyla, Finland, E-Mail: anna.e.reini@jyu.fi

Resumen: El conocimiento morfológico implica la comprensión de cómo los morfemas, las unidades mínimas con significado de las palabras, pueden utilizarse para formar nuevas palabras. Sin embargo, la investigación existente sobre el reconocimiento de morfemas en palabras multimorfémicas en inglés por parte de aprendientes de L2 sigue siendo limitada. Una revisión exhaustiva de la literatura revela que hasta ahora no se ha investigado el conocimiento que tienen los estudiantes universitarios finlandeses sobre palabras inglesas con múltiples morfemas. El finés es una lengua perteneciente a la familia urálica, y como lengua no germánica, sus estructuras son muy diferentes de las del inglés, por ejemplo. Su sistema morfológico altamente productivo y complejo ofrece múltiples puntos de partida para futuras investigaciones. El presente estudio examina el reconocimiento de morfemas, un aspecto central de la competencia morfológica receptiva, relevante para la adquisición del vocabulario, así como factores intralexicales que influyen en este proceso de reconocimiento. Los participantes, 694 estudiantes universitarios de habla finesa, realizaron tareas de segmentación léxica y conocimiento del vocabulario, diseñadas de acuerdo con su nivel de competencia. Utilizando análisis de correlación de Pearson y de correlación ítem-total corregida, los resultados mostraron que el reconocimiento de la raíz de una palabra estaba significativamente influenciado por la capacidad de identificar la estructura interna de palabras multimorfémicas. Los resultados amplían nuestra comprensión sobre cómo se desarrolla el conocimiento morfológico en aprendientes avanzados de inglés como L2 y aportan nuevas evidencias empíricas sobre su uso receptivo en la adquisición estratégica del vocabulario. Asimismo, subrayan la importancia de dirigir la atención de los aprendientes hacia las estructuras morfológicas y la función semántica de las raíces léxicas. A partir de los hallazgos, también se discuten las implicaciones pedagógicas para la enseñanza del vocabulario en L2.

Zusammenfassung: Morphologisches Wissen beinhaltet das Verständnis darüber, wie Morpheme, die kleinsten bedeutungstragenden Einheiten von Wörtern, zur Bildung von Wörtern verwendet werden können. Die bisherige Forschung zur Morphemerkennung L2-Lernender in mehrmorphemischen Wörtern im Englischen ist jedoch noch begrenzt. Nach eingehender Literaturrecherche lässt sich feststellen, dass bislang keine Forschung zum Wissen finnischer Universitätsstudierender über englische mehrmorphemische Wörter existiert. Finnisch ist eine Sprache aus der uralischen Sprachfamilie, und als nicht-germanische Sprache weist sie eine deutlich andere Struktur auf als etwa das Englische. Das hochproduktive und komplexe morphologische System des Finnischen bietet vielfältige Ansatzpunkte für weiterführende Forschung. Die vorliegende Studie untersucht die Erkennung von Morphemen, einem zentralen Aspekt rezeptiver morphologischer Kompetenz, die für den Wortschatzerwerb von Bedeutung ist, sowie intralexikalische Faktoren, die

diesen Erkennungsprozess beeinflussen. Die Teilnehmenden, 694 finnischsprachige Universitätsstudierende, bearbeiteten Aufgaben zu Wortsegmentierung und lexikalischem Wissen, welche entsprechend dem Kompetenzniveau der finnischen Teilnehmenden konzipiert wurden. Die statistischen Auswertungen mittels Pearson-Korrelationen und korrigierter Item-Gesamtkorrelationen ergaben, dass das Erkennen der Wortwurzel maßgeblich durch die Fähigkeit beeinflusst wurde, die innere Struktur mehrmorphemischer Wörter zu erkennen. Die Ergebnisse erweitern unser Verständnis darüber, wie morphologisches Wissen bei fortgeschrittenen L2-Lernenden des Englischen aufgebaut wird, und liefern neue empirische Belege für dessen rezeptive Nutzung im strategischen Wortschatzerwerb. Die Ergebnisse verdeutlichen, wie wichtig es ist, die Aufmerksamkeit der Lernenden auf verschiedene morphologische Strukturen und die Bedeutungsfunktion von Wortwurzeln zu lenken. Auf Grundlage der Ergebnisse werden auch die pädagogischen Implikationen für den L2-Wortschatzerwerb behandelt.

1 Introduction

Morphemes are the smallest, indivisible linguistic units that carry either a lexical or grammatical meaning (Booji 2007: 8–9). Even if morphemes might not seem particularly challenging for L1 speakers, they nonetheless create a challenge for L2 English learners (Schmitt & Schmitt 2020: 56). Learners need to know various aspects, such as understanding how affixes alter root meanings, to employ morphological knowledge in vocabulary learning (Nation 2022: 388). For receptive use, Nation (2022: 390) names morpheme recognition in complex words as the most important aspect of L2 morphological knowledge.

Literature on L2 vocabulary learning has long acknowledged the significant role of morphological knowledge (e.g., Bauer & Nation 2020: 5–7; Schmitt 2008). Yet, L2 English learners' awareness of morphemes, or their ability to detect them in multimorphemic English words, has received little attention in research. Past studies have examined L2 English learners' morphological knowledge, but few have approached it from a strategic vocabulary learning perspective (see, Bubchaiya & Sukying 2022; Wei 2015). Also, most studies have examined L2 learners with non-European L1 languages with only a few exceptions (e.g., Laufer, Webb, Kim & Yohanan 2021; Leontjev, Huhta & Mäntylä 2016; Snoder & Laufer 2022). Previous research findings have indicated that advanced adult L2 learners' L1 morphological type has a significant effect on their morphological awareness: L2 English learners whose L1 was an agglutinative language performed significantly better than learners with an isolating L1 in tasks on morphological awareness (for further information, see Wu & Juffs 2022). Even though Finnish has acquired many loanwords from Germanic lan-

guages, most notably Swedish, Finnish is an agglutinative language that also shares features with fusional languages. Therefore, focusing on Finnish university students provides new insights into L2 English learners' morphological knowledge. To the best of the author's knowledge, no research information exists on the morphological knowledge of Finnish university students, or more generally, among learners whose L1 has complex, highly synthetic word structures. This paper thus aims to answer the following questions: (1) as advanced learners of English and speakers of a Finno-Ugric language, how well can Finnish students recognize English morphemes in multimorphemic words, and (2) how do morphological structures affect their morpheme recognition? The current study belongs to a larger, ongoing project that aims at developing the teaching of morphological knowledge of English in higher education. The sub-studies included in the project examine different factors affecting receptive morphological knowledge and its effect on L2 vocabulary acquisition of Finnish university students.

2 Morphological knowledge in L2 English

2.1 The Multidimensionality of morphological knowledge

Roots form the semantic core of derived words (Lieber 2009: 33). Stems, on the other hand, are roots that already have affixes attached to them before additional affixes are added (Harley 2006: 120). For example, in *readability, read* is the root, and *–able* is the first affix attached to it. When the subsequent affix *–ity* is added, it is attached to the stem *readable.* Both English and Finnish words can include prefixes, affixes added before the root, and suffixes, affixes added after the root. Finnish, however, has fewer prefixes than English, as its word formation relies more on combining word roots and suffixes to form new derived and compound words (Sulkala & Karjalainen 1992). Derivational affixes are used to construct new words by altering the meaning of the root they are attached to whereas inflectional affixes affect only roots' grammatical function (Lieber 2009: 33). There are bound roots in English (e.g., *andro-* 'man' or 'male' and *-phile* 'loving') that act as affixes in word formation because they cannot stand alone but are not regarded as such since, unlike affixes, they usually create the core meaning of a word like free roots (Lieber 2009: 33–34).

The depth of vocabulary knowledge and the strength of associations between words are fundamentally influenced by the development of *morphological awareness* (Hayashi & Murphy 2011: 105). Morphological awareness refers to (1) the knowledge of morphemes and (2) the ability to analyze and manipulate morphological structures (Kuo & Anderson 2006). From an L2 learning perspective, knowledge of morphemes consists of three dimensions, (1) form, (2) meaning, and (3) use (Nation

2022: 384), where (1) form knowledge entails learners recognizing and using morphemes, (2) meaning knowledge refers to understanding morpheme meanings, and (3) use knowledge contains information about morphemes' effect on word usage, such as changes in the word class (Webb & Nation 2017: 164). Using morphological knowledge can be seen as a mnemonic that aids and supports vocabulary learning, even when not utilized consistently as a learning strategy (Bauer and Nation 2020: 6). However, for learners to apply receptive morphological knowledge in strategic vocabulary learning, recognizing morphemes in a multimorphemic word is a prerequisite (Nation 2022: 388). For a deeper understanding of word meanings, learners also need to understand affix and root meanings, and how they can create a new yet related meaning when combined, as well as to identify connections across words containing the same morphemes (Nation 2022: 388).

Most studies on morphological knowledge have examined how its relationship with other types of knowledge and proficiencies, such as L2 English language proficiency and vocabulary knowledge (e.g., Mochizuki & Aizawa 2000; Mäntylä & Huhta 2014; Schmitt & Meara 1997), reading (e.g., Collins & Nation 2015; Kieffer & Lesaux 2012; Zhang 2017), writing (Leontjev, Huhta & Mäntylä 2016), and the differences between L1 and L2 English speakers (e.g., Hayashi & Murphy 2011; Iwaizumi & Webb 2021). These studies have suggested that L2 morphological knowledge develops as general English proficiency and vocabulary knowledge develop, albeit the connection is not linear.

Hayashi and Murphy (2011) focused on the different aspects of morphological awareness and their connections. The authors used word segmentation to examine affix recognition and elicitation tasks to discover the relationship between productive and receptive morphological awareness of derivations and inflections among 22 adult Japanese L2 English learners and 20 L1 speakers of English as a comparison group. The findings revealed that instruction contributed to the development of morphological awareness and that semantic transparency affected participants' affix recognition in the segmentation task. Affix frequency or productivity did not seem to influence the results. The authors found evidence of L1 English speakers segmenting words per syllable whereas L2 English speakers segmented words per morphemes.

Past research and literature on morphological knowledge have typically focused only on affixes or the derivative forms, thus disregarding the effect roots have in constructing morphological knowledge. Therefore, not much is known about the role of roots in L2 morphological knowledge even though they create the core meanings of complex words (Webb & Nation 2017: 167), and understanding their meanings is fundamental in the receptive use of morphological knowledge.

2.2 Affix acquisition and derivative forms

Past research has nonetheless focused on affix acquisition and its role in morphological knowledge. Bauer and Nation (1993) categorized affixes into seven levels based on their form, frequency, productivity, and regularity of meaning, all of which were considered to affect their difficulty from an L2 learning perspective. Some studies have investigated learners' acquisition of affixes to examine whether Bauer and Nation's (1993) linguistically based hierarchy is supported by L2 learner data, but the results are contradictory (see, e.g., Hayashi & Murphy, 2011; Laufer et al., 2021; Leontjev, 2016).

Mochizuki and Aizawa (2000) examined the order of affix acquisition and the relationship between L2 English learners' affix knowledge and vocabulary size. The participants, 403 Japanese high school and university students, were given prefixes with three nonsense bases and asked to choose the best Japanese translation. For suffixes, the participants were asked to choose the correct word class. With polysemous affixes, the authors considered only the most frequent meaning as correct. The authors discovered that affix knowledge can be categorized based on accuracy order and provided five factors affecting the order: instruction, loanwords, frequency of affixes, frequency of the derived form, and the polyfunctional quality of affixes.

Previous research (e.g., McLean 2018; Mochizuki & Aizawa 2000; Schmitt & Meara 1997; Schmitt & Zimmermann 2002; Ward & Chuenjundaeng 2009) has also shown that L2 English learners, even at a high proficiency level, have limited knowledge of words' derivative forms. The findings have indicated, for instance, that learners recognize affixes in derivative forms only when the stems are familiar and that they can produce about half of the appropriate derivative forms. Moreover, studies on derivative forms have provided new information on affix recognition and the use of segmentation to measure morpheme recognition (e.g., Hayashi & Murphy 2011; Leontjev 2016).

However, little is known about how L2 English learners deal with multimorphemic words or their morpheme recognition, even though the literature has stated that it is the most important factor in receptive morphological knowledge. Past research has mainly focused on L2 learners' affix recognition, but the inclusion of roots offers novel evidence of how learners deal with multimorphemic words and what effect roots have in the process. Examining both affixes and roots, which create the semantic core in English word formation, morpheme recognition can be investigated from a strategic vocabulary learning perspective. The focus of this study is thus on morpheme recognition, the preliminary requirement for the receptive use of morphological knowledge in strategic vocabulary learning (Nation 2022: 390), and its implications for L2 teaching of morphological

knowledge. Moreover, the current study provides new information on a topic that has received relatively little attention by focusing on morpheme recognition among speakers of L1 Finnish, an agglutinative language with fusional features, which benefits future research examining learners whose L1s share the complex word structures (Karlsson 2006).

3 Methodology

3.1 Research questions

By examining the receptive morphological knowledge of Finnish university students and its pedagogical implications for L2 vocabulary acquisition, the present study aims to answer the following questions:

(1) As advanced learners of English and speakers of a Finno-Ugric language, how well can Finnish students recognize English morphemes in multimorphemic words?
(2) How do morphological structures affect their morpheme recognition?

3.2 Participants

An email invitation to participate in the study was sent to Finnish universities' faculties where it was forwarded to student mailing lists. Out of the 721 student respondents from various universities in Finland across different fields, 694 participants spoke L1 Finnish. Participants who did not speak Finnish as their L1 were excluded from the study. The length of university studies, and consequently familiarity with academic vocabulary, was varied among the participants. That is, from the 694 participants, 134 (19.3 %) were first-year students, 112 (16.1 %) second-year students, 108 (15.6 %) third-year students, 109 (15.7 %) fourth-year students, 88 (12.7 %) fifth-year students, and 143 (20.6 %) had studied six years or more. The degree programmes (major disciplines) reported by the students were categorized (see Figure 1) based on the Finnish Government Decree of University Degrees (Act 794/2004, Amendment 594/2020). In the Humanities category, 93 participants were language students, of whom 50 studied English and 43 some other language(s).

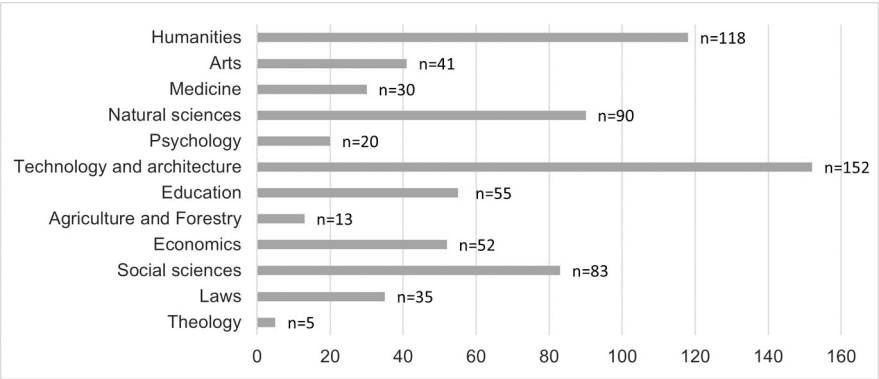

Figure 1: Degree programmes of the participants

3.3 Research instruments

The participants were presented with an online questionnaire (Webropol) that included questions on background factors and vocabulary learning as well as tasks on vocabulary and morphological knowledge. The data used for the current study consists of a task on receptive vocabulary knowledge and a word segmentation task. The word segmentation task was based on previous studies and Nation's (2022: 388) description of the receptive use of morphological knowledge needed in vocabulary learning and the imperative role morpheme recognition has in it.

Previous studies on L2 morphological knowledge have concluded that the level of morphological knowledge is connected to overall language proficiency (e.g., Iwaizumi & Webb 2021; Mochizuki & Aizawa 2000; Mäntylä & Huhta 2014; Snoder & Laufer 2022).

Therefore, tasks measuring morphological knowledge used by past research were adapted for the target group to correspond with the anticipated English proficiency level of Finnish university students. That is, when comparing Finnish-speaking learners' vocabulary proficiency (see Sub-section 3.3.1) in English to the proficiency of participants with other L1s[1] in past research on L2 morphological knowl-

1 Snoder and Laufer (2022) reported that the Swedish 9th graders vocabulary size was about 5,600 words which was similar to 12th graders in Israel (Laufer et al. 2021), and to the most advanced university students in Japan (McLean 2018; Mochizuki &Aizawa 2000). In Iwaizumi and Webb (2021), upper-secondary and university students' (L1 Japanese and Chinese) mastery of word level were: 72 % at 1000-word level, 50 % 2000-word level, 20 % 3000-word level, 15 % 4000-word level, and 15 % 5000-word level.

edge, it was concluded that Finnish university students' English vocabulary proficiency, and thus morphological knowledge too, might be more advanced.

Non-words were not used because their use would have affected the examination of root recognition. The online version of the Oxford English Dictionary (OED) (n.d.) was used as a source for etymology, use, and meanings.

3.3.1 Development of the Vocabulary knowledge task

To examine if the participants' receptive vocabulary knowledge was connected to morpheme recognition, a simple task on receptive vocabulary knowledge (VK) was administered in the beginning. The VK task included 30 target items from the 5,000– and 7,000–word levels (see Appendix). It should be noted that, instead of investigating different frequency levels, the aim of the task was merely to gain a general overview of the participants' vocabulary knowledge to account for its explanatory value in morpheme recognition.

Not much is known about the English vocabulary proficiency of Finnish-speaking university students as no research focusing on their English vocabulary level or size exists. The results of Harkio and Pietilä (2016), who examined reading comprehension's connection to the vocabulary depth and breadth of 39 upper secondary school students and 19 English major University students in Finland, revealed that the English students' vocabulary size test score was 8,816 on average, ranging from 7,300 to 9,800, and that of the upper secondary school students was 5,717, ranging from 1,500 to 8,600.

It must be highlighted that the objective of the current study was to gain an overview of the participants' receptive vocabulary knowledge by utilizing target items from different frequency levels to gain variance, rather than to examine knowledge in different frequency levels. In deciding which target items to include in the VK task, five university students were shown both the 5,000th level in Webb, Sasao, and Ballance's (2017) updated Vocabulary Levels Test (VLT) and the 10,000th level in Schmitt, Schmitt, and Clapham's (2001) VLT and were asked to evaluate their difficulty. Two students assessed the 10,000-level to be closer to their vocabulary level, whereas three students considered the 10,000-level to be too demanding for them, assessing that it would cause a sense of discouragement for continuing with other vocabulary-related tasks. Consequently, it was hypothesized that by combining items from both Webb et al.'s (2017) and Schmitt et al.'s (2001) lowest frequency levels, 5,000 and 10,000, the data would reveal differences between the participants without discouraging the less proficient participants.

The frequency levels in Webb et al.'s (2017) VLT were based on Nation's 2012 version of the British National Corpus/Corpus of Contemporary American English

(BNC/COCA) frequency levels (for a detailed description, see Webb et al., 2017). The lowest frequency level in Webb et al.'s (2017) VLT is 5,000, whereas the items in Schmitt et al.'s (2001) 10,000-level no longer belong to the 10,000th frequency level according to the BNC/COCA (Nation 2017). Instead, they range from the 5,000th to the 9,000th level, with most belonging to the 7,000th level. The words from the 7,000th level were therefore selected and combined with half of the items from the 5,000-level version by Webb et al. (2017). When piloting the VK task alongside other tasks in the questionnaire, there was a sufficient variance in performance, with scores ranging from 9 to 29. This range indicated that the data would show differences among the participants without floor or ceiling effects (M = 22.6; $S.D.$ = 5.8).

3.3.2 Word segmentation (WS) task

The word segmentation task was chosen as a measure for morpheme recognition, as previous studies have shown its usefulness for measuring affix recognition (e.g., Bubchaiya & Sukying 2022; Hayashi & Murphy 2011; Leontjev 2016; Leontjev, Huhta & Mäntylä 2016). The task measured the recognition of 21 morphemes. They were used to create six multimorphemic words which the participants were asked to break into the smallest meaningful parts with the following example:

mismanaged → mis + manag(e) + ed

The stems for the multimorphemic words were chosen at random from the 300 most frequent word families in Gardner and Davies' (2014) Academic Vocabulary List (AVL), and it was hypothesized that the participants would be familiar with the words used in the tasks. Of the six stems that were chosen from the AVL, five were constructed using roots that belonged to the 1,000–3,000 frequency levels in the BNC/COCA, and one root, *globe*, belonged to the 5th frequency level. The multimorphemic words were also considered semantically transparent based on the morphemes used to construct them. However, the WS task measured only the recognition of morphemes in multimorphemic words based on their form and did not account for the participants' familiarity with morpheme meanings as the focus of the current study was on segmenting the morphological structure of words rather than on the mastery of their meanings.

From the chosen word families, the member[2] of the family containing the highest number of morphemes was selected. Although this may have made the recogni-

2 The online version of the AVL lists family members for each word family.

tion of all morphemes more demanding, multimorphemic words provide more in-depth information on the interaction between the different morphological structures (prefixes, roots, and suffixes). The six multimorphemic words were (with the base word and its rank based on frequency in the parentheses): *globalization* (global 159), *economic* (economy 37), *researchist* (research 6), *independently* (independent 231), *encouragement* (encourage 261), *internationalize* (international 85). A word containing two bound roots (*eco-* and *– nomy*) was included. Even though bound roots act as affixes in word formation, they were not regarded as such in accordance with OED (n.d.). Instead, they were considered to be more closely connected to free roots due to their effect on words' meanings. Each correctly segmented morpheme scored one point and the maximum score for the task was 21. Alternative forms (e.g., *-ion* instead *–ation*) were accepted if they did not interfere with the adjacent morphemes.

3.4 Procedure

An online questionnaire was used for the data collection as it was conducted during social distancing caused by COVID-19. The use of a questionnaire nonetheless allowed data collection across different universities and gaining a multitude of participants as no timing or geographical restrictions were present. At the beginning of the questionnaire, it was emphasized that no aiding tools, such as dictionaries, should be used. It was estimated that the questionnaire took 45 minutes to complete, so extra attention was paid to succinct wording and the layout. It was also ensured that no metalinguistic knowledge was needed to understand the instructions or to complete the tasks. The questionnaire included additional measures that are not reported in the current paper, and only a limited number of target items could therefore be included to prevent fatigue.

The questionnaire was first piloted with two students, who answered it while thinking aloud. As no problems or misunderstandings were detected, the questionnaire was then piloted with 23 students which showed that performance in the tasks varied sufficiently[3] and that the tasks could be used to examine varying proficiency levels. The piloting revealed one issue; the suffix *–er* was used in another task measuring morphological knowledge. Its inclusion in the WS task might have facilitated its recognition, hence *researcher* was changed into *researchist* after piloting. Even though infrequently used, according to OED (n.d.), *researchist* is an existing word.

3 In the VK task, the mean was 75.3 % (22.6/30; *S.D.* = 5.8), the scores ranging from 9 to 29. In the WS task, the mean was 53.8 % (11.1/21; *S.D.* = 3.8), the scores ranging from 2 to 17.

The suffix *-ist* also has a high level of semantic transparency. As no other changes were made, the responses to the pilot study were included and only discarded when examining the data on the suffix *-ist.*

By addressing morphemes instead of affixes, free roots (referred to as roots) and bound roots are included in the following data-driven analysis. The data-driven analysis was conducted with a series of correlational analyses, specifically corrected item-total correlation analyses. The corrected item-total correlation analysis, one of the two methods that Zijlmans, Tijmstra, van der Ark and Sijtsma (2019) concluded to be closest to the ideal ordering of test items, describes an item's association with the total score on the other items in a test construction, and it thus allows for the correlational analysis of dichotomous data. To answer the second research question, the corrected item-total analysis was used to determine which morphemes were most related to the overall score on morpheme recognition, which was considered to reflect the students' general ability to break morphologically complex words into parts. The results could then be used to examine how morphological structures affect morpheme recognition and the morphological awareness required for the receptive use of morphological knowledge.

4 Results

4.1 Descriptive analysis

As Table 1 summarizes, the mean of the WS task was only 53 % of the maximum score which suggests that recognizing morphemes, the preliminary type of knowledge needed for the receptive use of morphological knowledge, was only partial. Based on the VK task's mean of 81 %, the difficulty of the task was considered suitable for the target group.

Table 1: Scores on the Word segmentation and Vocabulary knowledge tasks

	N	Mean (%)	95 % CI of the mean		S.D.	Max. score
			Lower	Upper		
Word segmentation (WS) task	679	11.02 (52.5)	10.72	11.32	3.95	21
Vocabulary knowledge (VK) task	694	24.14 (80.5)	23.73	24.54	5.41	30

To address the pedagogical implications further, a descriptive section with examples from the data will be provided to illustrate and analyze how L2 learners perceive multimorphemic words. To answer the second research question in detail, the connection of morpheme recognition (WS task) with vocabulary knowledge (VK task) is briefly examined as well.

4.2 Morpheme recognition

The items on the WS task were found reliable as the internal consistency across the items was measured by Cronbach's Alpha coefficient of .74. To examine the difference between the morphological structures and overall morpheme recognition, the morphemes used in the WS task were combined into groups according to their type. A corrected item-total correlation, which not only is suitable for dichotomous data but also allows for the correlational analysis of inter-relational items within a group, was therefore performed to analyze which morphemes (Table 4) and morphological structures (Table 2 and Table 3) were most connected to the overall performance in the WS task. The average score and correlation for each morphological structure with the WS task's total score are presented in Table 2.

Table 2: Correlational analysis between the morphological structures and the WS task

Morphological structure	Correct segmentation (%)	Corrected item/total correlation	Coefficient of determination (R^2)
Prefixes	79.5	.530	.281
Suffixes	47.7	.309	.095
Roots	49.8	.662	.438
Bound roots	31.5	.178	.003

Prefixes were most often recognized (80 %) whereas, on average, half (50 %) of the roots were recognized. When comparing the morphological structures and their correlation with the WS task's total score, the roots correlated most with the task. Based on the coefficient of determination for roots, recognizing them accounted for 44 % of the performance in the WS task whereas that of prefixes accounted for 28 % of the performance. The recognition of suffixes was also about half (48 %) on average, but the coefficient of determination for suffixes accounted only for 10 % of participants' performance.

It was also examined if the frequency of the roots could be used to explain the correlation or the difficulty order of the roots, but no connections could be found.

Some stem forms were more frequent than the root (e.g., *national* in comparison to *nation*) so their frequency was compared to the difficulty and correlational order, but no correspondences were evident.

When examining the factors that might have affected the correlation, the first four suffixes, *-ent, -al*$_2$, *-al*$_1$, and *-ize*$_1$, all of which correlated strongly with the WS task, were connected to the root to create the stem form (e.g., *nation***al***ize*). Also, these suffixes had the highest item difficulty as they were recognized by fewer than 22 % of the participants. The remaining suffixes were at the end of each word (e.g., *national***ize**). To further analyze the differences among the suffixes, based on their placement in the word, the corrected item-total correlation presented in Table 2 was conducted again, with the suffixes divided into two groups: stem and final suffixes (see Table 3).

Table 3: Corrected item/total correlation between the morphological structures, with suffixes divided into two groups

Morphological structure	Corrected item/total correlation	Coefficient of determination (R^2)
Prefixes	.530	.281
Stem suffixes	**.585**	**.342**
Final suffixes	**.244**	**.060**
Roots	.662	.438
Bound roots	.178	.003

The stem suffixes had a stronger correlation (.585) with the total score in comparison to the correlation of the final suffixes (.244). Moreover, like roots, the stem suffixes correlated more strongly (.585) with the WS task as a group than any of the single suffixes (.532 < .585), which suggests that stem suffixes are a suitable unit for examining morpheme recognition. In conclusion, recognizing suffixes in the stem was connected to general proficiency in morpheme recognition, as it accounted for 34 % of the performance in the WS task. The recognition of the final suffixes, acting as class-changing suffixes, accounted only for 6 % of the overall performance in the task.

Overall, the average score on the WS task was approximately half (53 %) of the maximum score. However, there was a great deal of variance between the morphemes in terms of how many participants recognized them. Also, the percentage only describes the difficulty of each morpheme and does not reveal how they are connected to the overall proficiency. In Table 4, the morphemes are grouped by morphological structures (prefixes, roots, and suffixes), and further organized in descending order based on their correlation with the task's total score.

Table 4: The Percentage of correct segmentation and correlation of each morpheme with the WS task

Morphemes	Correct segmentation (%)	Corrected item/total correlation
Prefixes		
in-	81.3	.573
re-	71.6	.545
en-	85.3	.532
inter-	80.7	.369
Suffixes		
-ent	21.8	.532
-al$_2$	17.6	.514
-al$_1$	13.6	.416
-ize$_1$	17.2	.370
-ation	23.0	.346
-ist[a]	93.2	.266
-ment	95.0	.245
-ic	54.7	.197
-ly	84.4	.191
-ize$_2$	61.0	.098
Roots		
search	69.5	.558
courage	80.7	.471
depend	31.5	.425
nation	41.6	.384
globe	25.5	.313
Bound roots		
-nomy	15.6	.341
eco-	47.7	.098

[a]Participants from the pilot study were excluded from the analysis.

There was not much variance between the prefixes as they were correctly segmented by most (72 % – 85 %) participants, and an average of 80 % of the prefixes were segmented correctly. Nevertheless, the correlations of most prefixes (*in-*, *re-*, *en-*) were rather substantial and did not have a great deal of variability (.573 – .532). The order of correctly segmented prefixes was similar to Bauer and Nation's (1993) affix

levels as *-re*, which was correctly segmented the least, was on Level 6 whereas the other prefixes were on Level 5.

In contrast to prefixes, the participants segmented only an average of 48 % of the suffixes correctly, and there was more variance between the suffixes (14 % – 95 %). Similarly, the correlation for each suffix with the total score varied more (.098 –.532). From all the morphemes, however, the suffixes *–ment* (95 %) and *–ist* (93 %) were most often segmented correctly. The order of suffixes based on the percentage of correct segmentation could not be explained by Bauer and Nation's (1993) affix levels.

Two suffixes, *-ize* and *–al*, were used in two target items. Both instances of *–al* were rather poorly recognized (13 % – 18 %). However, -al$_1$ had a moderate (> .300) correlation, and -al$_2$ had a strong (> .500) correlation with the WS task, which means that these items predicted well the overall proficiency in the task. The segmentations for the two instances of *–ize* differed more as *–ize*$_1$ was segmented only by 17 % of the participants but its correlation was over .300, thus its recognition being connected to the overall score of the WS task, whereas 61 % of the participants segmented the second *-ize*$_2$ but it did not correlate (<.100) with the task. The suffix *–ize*$_1$ was used in the stem form (*global**iz**ation*) whereas the other suffix *-ize*$_2$ was added to the stem form at the end of the word (*international**ize**)*.

Based on the number of respondents who segmented the roots correctly, no general remarks can be made in terms of their difficulty since there was a great deal of variability (26 % – 81 %). Nevertheless, the correlations for the roots with the WS task were prominent (> .300). Similarly, there was variability with the bound roots as *eco-* was recognized by 48 % whereas only 16 % recognized *-nomy*, and the correlation for *-nomy* was more salient (.341) compared to *eco-* (.098).

4.3 Segmenting multimorphemic words

Even though the overall performance in the WS task confirmed that the participants correctly understood the instructions and segmented the multimorphemic words into several morphemes, the segmentations nonetheless yielded interesting insights into the students' perceptions of multimorphemic words. To provide examples of how multimorphemic words are perceived by L2 English learners, some observations made on the segmentations, along with excerpts from the data, are provided below. In relation to the second research question on the pedagogical implications of morpheme recognition, only answers that did not segment morphemes correctly are discussed.

From all the words in the WS task, the segmentations for *globalization* varied the most. Those 499 participants (72 %) who failed to recognize the root, *globe,* segmented the word in a variety of ways, such as:

– *global* + *li* + *zation* (P41[4]),
– *global* + *bali* (P92),
– *global* + *alization* (P205),
– *globa* + *lization* (P217, P240, P317),
– *globa* + *li* + *zation* (P231),
– *global(y)* + *zation* (P246),
– *glo* + *balizat(i)* + *on* (P335), and
– *glo* + *bali* + *zation* (P393).

In the segmentations where *globe* was not detected, *global* was the most common instance in the data, segmented by 322 (46 %) participants instead *glob(e)* + *al*. This suggests that the participants were familiar with the stem form, and it dominated the detection or analysis of the adjectival suffix *–al* and its effect on word class. Also, the final letter in *globe* is omitted when *-al* is added, which can impose extra difficulty for its recognition.

Similarly, from the 391 (56 %) participants who failed to segment *nation* in *internationalization*, the most common form segmented was either *international* by 76 (11 %) participants or *national* by 221 (32 %) participants. However, not recognizing the root of *internationalize* affected the adjacent affixes *inter-* and *-al* in many cases, such as:

– *in* + *ternation* + *alize* (P52),
– *in* + *ternational* + *ize* (P148, P184, P280),
– *in* + *terna* + *tion* + *alize* (P451),
– *internatio* + *nalize* (P510), and
– *(inter)* + *ternate* + *(ion)* + *(lize)* (P671).

In the segmentations, there was evidence of the participants trying to connect meanings to the parts they discovered by making connections across words, such as:

– *global* + *(n)ation* (R58),
– *inter* + *national* + *formalize* (P93),
– *international* + *(civil)ize* (P300),
– *internatiol, normalize* (P449),
– *global* + *utilization* (P498),

4 Each participant (P) was assigned a number before conducting the analysis and exclusion of other L1s, hence the numbering going above the number of participants (N=694) included in the study.

– *international + alize thing. Globalize. To do?* (P531),
– *global + (real)ization global + (l)iza(rd) + nation* (P570),
– *global + (organi) + zation* (P661), and
– *globe + global + ovation* (P698).

These excerpts illustrate the inventive and imaginative ways in which morphemes in multimorphemic words can be perceived by L2 English learners whose L1 is morphologically rich and complex, making it difficult to draw simple conclusions about the myriad ways a morpheme was segmented within the scope of the study. As the examples indicate, however, the Finnish participants did not segment words per syllable but rather relied on meaning and familiarity.

4.4 Vocabulary knowledge

To assess the pedagogical implications of morpheme recognition further, the relationship between general vocabulary knowledge and morpheme recognition was examined. The average score on the VK task was 80.5 % of the maximum score (see Table 1), and the correlation coefficient was weak ($r < 0.3$) between the VK and WS tasks (see Table 5). However, when examining Pearson correlation coefficients per morphological structure, there was a moderate level of correlation ($r > 0.3$) between the VK task and the suffixes in the WS task.

Table 5: Pearson's correlation coefficients of the VK task with the WS task and the morphological structures

	VK task
WS task	.249*
Prefixes	.106*
Suffixes	.325*
Roots	.126*
Bound roots	-.019

* $p < .001$.

5 Discussion

This study aimed at exploring receptive morphological knowledge from a strategic vocabulary learning perspective and focused on L2 English learners' ability to recognize morphemes. The findings implied that the Finnish university students' general proficiency in detecting morphemes was somewhat adequate since they could recognize on average about half (53 %) of the morphemes used in multimorphemic words. This is in accordance with Leontjev et al. (2016), being the only other study that had L1 Finnish participants using a similar type of segmentation task, who reported the average score for their word segmentation task being 19.39 out of 36 (54 %). In contrast, Hayashi and Murphy (2011) reported the average score in their segmentation task for class-maintaining affixes being close to 90 % whereas that of class-changing affixes was about 75 %. One reason for this contrast could be that the authors reported the participants being explicitly taught inflectional and derivational morphology. To the best of the author's knowledge, Finnish school textbooks provide only limited focus on morphology, and at the tertiary level, there are no textbooks, so the extent to which morphology is addressed depends on the individual teacher.

When examining the recognition of the morphemes used in the Word segmentation (WS) task, there was considerable variance (13 % – 93 %) between the tested morphemes in terms of how many participants recognized them. The difficulty the participants had with recognizing morphemes was rather different from the morpheme order based on correlation. The results of the corrected total-item correlation analysis expressed the likelihood of general proficiency in morpheme recognition based on detecting morphological structures, which can be used for pedagogical purposes to increase morpheme recognition and awareness.

The results suggested that the morphological structures affect morpheme recognition differently, and therefore, the placement of morphemes in multimorphemic words affects their recognition. More specifically, the findings suggested that proficiency in recognizing the stem form and breaking it into its constituent parts (root and suffix) was most closely related to a general proficiency in morpheme recognition.

The corrected item-total correlational analysis also implied that recognizing roots was strongly associated with the overall performance in morpheme recognition. This supports Webb and Nation's (2017:167) claim of root recognition's importance in morphological skills. In fact, recognizing roots accounted for 50 % of the overall morpheme recognition in the WS task, which suggests that the ability to recognize roots might signify the general ability to break multimorphemic words into parts. Previous studies have indicated that root frequency could be used to determine the knowledge of derived words (Snoder & Laufer 2022) because affix

frequency does not correspond to the frequencies of derived forms (Mochizuki & Aizawa 2000). However, the findings of the current study contradicted this, as neither the difficulty nor the correlation order of the roots could be explained by their frequency.

Past studies have focused on affix detection and meaning, but none of the past studies focusing on morpheme recognition have included roots, even though they create the core meaning of words whereas affixes are only used to modify that meaning. Therefore, it was hypothesized that existing research findings on affixes cannot be used to develop L2 vocabulary teaching and acquisition when roots have a fundamental role in creating word meaning. The results support this claim by demonstrating that the detection of roots was most closely connected to the ability to break words into parts, which is the first and the most important step in using receptive morphological knowledge in strategic vocabulary learning (Nation 2022: 390).

Based on the findings, the ability to recognize suffixes that were connected directly to roots, forming the stem, notably influenced morpheme recognition. Recognizing the suffixes used in the stem accounted for 34 % of the performance in the WS task. In contrast, the recognition of the final suffixes was only weakly connected to morpheme recognition ability, accounting only for 6 % of the performance in the task. This suggests that proficiency in recognizing the stem form and breaking it into its constituent parts (the root and suffix) is closely related to general proficiency in morpheme recognition.

There was also a strong relationship between the recognition of prefixes and the overall performance in segmenting the words. That is, of the tested prefixes, *inter-* had a moderate correlation, whereas the other prefixes had strong correlations with the overall performance in morpheme recognition in the WS task. The affix levels by Bauer and Nation (1993) could be used to explain the difficulty order of the prefixes, but as only four prefixes were tested, no affirmative conclusions can be drawn.

Overall, the excerpts from the data also illustrated that Finnish university students were able to break multimorphemic words into parts, use meanings, and make connections across words when doing so. Even when failing to detect morphemes correctly, the students used these methods and did not segment words per syllable. Furthermore, the excerpts also highlight the importance of raising students' morphological awareness, specifically the role of roots and their difference from stems, in L2 vocabulary teaching. In terms of Nation's (2022: 390) description of receptive morphological knowledge, the students were making connections across words based on shared forms and using meanings when trying to detect morphemes.

Lastly, the results suggested only a weak relationship between vocabulary knowledge (the VK task) and morpheme recognition (the WS task), which is consis-

tent with Hayashi and Murphy (2011) and Leontjev et al. (2016). However, a closer analysis revealed that the recognition of suffixes had a moderate association with vocabulary knowledge.

6 Conclusion

The findings revealed that the morphological structures differ in the way they affect the ability to break multimorphemic words into parts. The analysis showed that recognizing roots was most closely connected to general proficiency in detecting morphemes in multimorphemic words. The results also provided evidence for the importance of detecting the stem form and breaking it into a root and suffix. The significance of roots in morpheme recognition is new in L2 English research, although the literature on L2 vocabulary learning has acknowledged their importance. The pivotal role of root detection in overall morpheme recognition has implications for both L2 English teaching and research.

The benefits of focusing on root detection in L2 English teaching are manifold. Firstly, drawing students' attention to roots helps them deal with multimorphemic words and break them into parts. This is particularly crucial for the receptive use of morphological knowledge in strategic vocabulary learning (Nation 2022: 390). Secondly, roots create the semantic core of complex words. Therefore, when introducing roots, students' attention can simultaneously be drawn to root meanings and how roots contribute to the overall meaning of the word, which students can then use to decipher and analyze unfamiliar words as well as construct a deeper understanding of previously acquired words. Lastly, the ability to break stem forms into constituent roots and suffixes further enhances the development of both receptive morphological knowledge and overall lexical knowledge, as it supports understanding of English word formation and the role affixes play in modifying root meanings. General vocabulary proficiency was most associated with the recognition of the final suffixes, which act as class-changing suffixes and thus have a syntactic role in word formation. As vocabulary proficiency was not as closely associated with detecting the root or the stem, both of which affect the meaning of words more than final suffixes, this result suggests that more advanced learners recognize the syntactic role of morphemes but not necessarily their semantic roles. This further supports the need to address the semantic role of roots and how they create the core meaning of a word when teaching L2 English word formation.

As to the implications for research, the results suggest a pivotal interconnection between the form and meaning dimensions in L2 learners' knowledge of morphemes. That is, morpheme recognition, belonging to the form dimension of morphological knowledge, was influenced by the recognition of roots, which have a high

semantic load and can therefore be considered to overlap with the meaning dimension. This highlights the need to further explore the multidimensionality of morphological knowledge to examine the relationships between different dimensions and how the possible connections can be further supported in teaching. Based on the findings, it also seems that item difficulty may not be as useful in examining the intralexical factors affecting word segmentation, as the percentage indicating correct segmentation was rather different from the item-total correlation signifying the connection between morpheme recognition and the overall ability to break multimorphemic words into parts. Additionally, using the results of this study together with past L2 English research findings on affix acquisition and difficulty with derivative forms, a detailed exploration of the interaction between the different morphological structures would benefit our understanding of how L2 morphological knowledge develops. The development can then be enhanced accurately by focusing on the morphological structures that need additional support or contribute most to the construction of morphological knowledge.

It is also essential to investigate the extent and impact of root detection in multimorphemic words on L2 English vocabulary learning. Even though the segmentation task seems suitable for morpheme recognition, more investigation with different types of methods that, for instance, denote morpheme meanings or recognition across words, is also needed to account for the method effect and to further address the interactions between morphological structures in multimorphemic words. Further research should also be conducted with varying L1s to determine how much Finnish as an L1 influenced the morpheme detection. That is, in comparison to English, Finnish word formation rarely uses prefixes and relies more on compounding roots, and overall, Finnish uses more complex inflectional and derivational systems, affecting its semantics and syntax. Therefore, Finnish L1 speakers might be more inclined to overlook prefixes and concentrate more on roots and the subsequent suffix(es). However, their general ability to analyze multimorphemic word structures might be heightened due to the complexity of Finnish words. Additionally, most Finnish students have knowledge of Swedish, a Germanic language that has also influenced Finnish, particularly through loanwords. Although proficiency in Swedish is generally not as advanced as in English because English teaching begins earlier, and the target proficiency level in English after basic education is higher than that of Swedish, future research would benefit from examining how cross-linguistic influence and multilingualism affect morphological knowledge in L2 English.

As the data collection, which included several measures in addition to the ones reported in this paper, took place during a stressful time of social distancing when all university teaching was conducted remotely, the number of target items in the measures had to be limited. Therefore, even though the findings provided evidence

of the salient role of roots in morpheme detection, further research with more target items is needed to confirm the results. The questionnaire emphasized that dictionaries or other resources should not be used, as the purpose was to gather information on the participants' current knowledge rather than to test it. Of course, it cannot be guaranteed that no resources were used. However, since it was not a high-stakes test and participation was both voluntary and anonymous, it can be reasonably assumed that the participants most likely did not use any external aids. Similar to other studies, morpheme recognition was measured using word segmentation, but the measure yielded dichotomous data, which limited the possible methods for analysis. A corrected item-total correlation was the best method to examine the items' effect on the overall score, as other methods, such as regression or factor analysis, require other independent variables and are not suitable for dichotomous data. Also, for the purpose of the study, a corrected item-total correlational analysis was considered sufficient as the aim was to examine morpheme recognition and its pedagogical implications. Indeed, the results provided a novel insight into morpheme recognition among advanced L2 English learners and how root recognition affects it, which also benefits the development of L2 vocabulary teaching.

References

Bauer, L., & Nation, I.S.P. (1993). Word families. *International Journal of Lexicography 6*(4), 253–279.

Bauer, L., & Nation, I.S.P. (2020). *English morphology for the language teaching profession.* Abingdon: Routledge.

Booji, G. (2007*.) The Grammar of Words: An Introduction to Linguistic Morphology*, 2nd edn*.* Oxford: Oxford University Press.

Bubchaiya, N., & Sukying, A. (2022). The Effect of Word Part Strategy Instruction on the Vocabulary Knowledge of Thai Primary School Learners. *Journal of Education and Learning 11*(5), 70–81.

Collins, B., & Nation, P. (2015). Testing receptive knowledge of derivational affixes*. Journal of Second Language Teaching and Research 4*(1), 6–23.

Gardner, D., & Davies, M. (2014). A New Academic Vocabulary List. *Applied Linguistics 35*(3), 305–327.

Harkio, N., & Pietilä, P. (2016). The Role of Vocabulary Breadth and Depth in Reading Comprehension; A Quantitative Study of Finnish EFL Learners. *Journal of Language Teaching and Research (7)*6, 1079–1088.

Harley, H. (2006). *English Words, a Linguistic Introduction.* Oxford: Blackwell Publishing.

Hayashi, Y., & Murphy, V. (2011). An investigation of morphological awareness in Japanese learners of English. *The Language Learning Journal 39*(1), 105–120.

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.

Karlsson, F. (2006). Finnish as an agglutinating language. In K. Brown (ed.), *Encyclopedia of language & linguistics* 2nd edn., 476–480. Elsevier. https://doi.org/10.1016/B0-08-044854-2/00147-4 (accessed 4 November 2024)

Kieffer, M. J., & Lesaux, N. K. (2012). Direct and indirect roles of morphological awareness in the English reading comprehension of native English, Spanish, Filipino, and Vietnamese speakers. *Language Learning 62*(4), 1170–1204.

Kuo, L-J., & Anderson, R. C. (2006). Morphological Awareness and Learning to Read: A Cross-Language Perspective. *Educational Psychologist 41*(3), 161–180.

Laufer, B., Webb, S., Kim, S. Y., & Yohanan, B. (2021). How well do learners know derived words in a second language? The effect of proficiency, word frequency and type of affix. *ITL – International Journal of Applied Linguistics 172*(2), 229–258.

Leontjev, D. (2016). L2 English derivational knowledge: Which affixes are learners more likely to recognise? *Studies in Second Language Learning and Teaching 6*(2), 225–248.

Leontjev, D., Huhta, A., & Mäntylä, K. (2016). Word derivational knowledge and writing proficiency: How do they link? *System 59*(July), 73–89.

Lieber, R. (2009). *Introducing Morphology*. Cambridge: Cambridge University Press.

McLean, S. (2018). Evidence for the adaption of the flemma as an appropriate word counting unit. *Applied Linguistics 39*(6), 823–845.

Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System*, *28*(2), 291–304.

Mäntylä, K., & Huhta, A. (2014). Knowledge of word parts. In J. Milton and T. Fitzpatrick (eds.), *Dimensions of vocabulary knowledge*, 45–59. London: Palgrave Macmillan.

Nation, I.S.P. (2017). *The BNC/COCA headword lists* (Version 1.0.0) [Data set]. https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-lists (accessed 2 August 2022)

Nation, I.S.P. (2022). *Learning Vocabulary in Another Language* (3rd edition). Cambridge University Press. Oxford University Press. (n.d.). OED Online dictionary. https://www.oed.com (accessed August 24, 2022).

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research 12*(3), 329–363.

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: word associations and verbal suffixes. *Studies in Second Language Acquisition 19*(1), 17–36.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing 18*(1), 5–88.

Schmitt, N., & Schmitt D. (2020). *Vocabulary in Language Teaching*, 2nd edn. Cambridge: Cambridge University Press.

Schmitt, N., & Zimmermann, C. (2002). Derivative word forms: what do learners know? *TESOL Quarterly 36* (2), 145–171.

Snoder, P., & Laufer, B. (2022). EFL Learners' Receptive Knowledge of Derived Words: The Case of Swedish Adolescents. *TESOL Quarterly 56*(4),1242–1265.

Sulkala, H., & Karjalainen, M. (1992). *Finnish. Descriptive grammars.* Abingdon: Routledge.

Ward, J., & Chuenjundaeng, J. (2009). Suffix knowledge: acquisition and applications. *System 37*(3), 461–469.

Webb, S., & Nation, P. (2017). *How Vocabulary is Learned*. Oxford: Oxford University Press.

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL – International Journal of Applied Linguistics 168*(1), 33–69.

Wei, Z. (2015). Does teaching mnemonics for vocabulary learning make a difference? Putting the keyword method and the word part technique to the test. *Language Teaching Research 19*(1), 43–69.

Wu, Z., & Juffs, A. (2022). Effects of L1 morphological type on L2 morphological awareness. *Second Language Research*, *38*(4), 787–812.

Zhang, D. (2017). Derivational morphology in reading comprehension of Chinese-speaking learners of English: A longitudinal structural equation modeling study. *Applied Linguistics 38*(6), 461–469.
Zijlmans E. A. O., Tijmstra J., van der Ark L. A., & Sijtsma K. (2019). Item-Score Reliability as a Selection Tool in Test Construction. *Frontiers in Psychology 9*(2298).

Appendix

Vocabulary knowledge task

The 5,000 words levels are from Webb et al.'s (2017) updated VLT, and the 7,000 words levels are from the 10,000 word levels in version 1 and 2 in Schmitt et al.'s (2001) VLT. The frequency levels were checked from the BNC/COCA (Nation, 2017). The headings denoting the word class and word level were not included in the original measure.

1. Noun (5,000 Word Level)

	batch	dentist	hum	lime	pork	scripture
meat from pigs						
low, constant sound						
green fruit						

2. Noun (5,000 Word Level)

	amnesty	claw	earthquake	perfume	sanctuary	wizard
safe place						
man who has magical powers						
liquid that is made to smell nice						

3. Noun (5,000 Word Level)

	altitude	diversion	hemisphere	pirate	robe	socket
person who attacks ships						
kind of clothing						
height						

4. Noun (7,000 Word Level)

	stint	expulsion	saliva	stealth	parole	prelude
introduction[a]						
release from prison early						
natural liquid present in the mouth						

5. Noun (7,000 Word Level)

	tentacle	scaffold	jumble	predicament	expulsion	puddle
small pool of water						
messy mixture[b]						
difficult situation						

6. Verb (5,000 Word Level)

	bribe	expire	immerse	meditate	persecute	shred
think deeply						
end						
cut or tear into small pieces						

7. Verb (5,000 Word Level)

	commemorate	growl	ignite	pierce	renovate	swap
catch fire						
go into or through something						
exchange						

8. Verb (7,000 Word Level)

	peek	dissipate	acquiesce	intrude	skid	mutilate
scatter or vanish						
injure or damage						
to accept without protest						

9. Adjective (5,000 Word Level)

	diesel	incidental	mandatory	prudent	superficial	tame
not dangerous						
using good judgement						
required						

10. Adjective (7,000 Word Level)

	mammoth	obsolete	auxiliary	candid	translucent	murky[c]
immense						
helping, adding support						
no longer used						

[a]*Introduction* was added as a definition because Schmitt et al. (2001) used *prelude* only as a distractor.
[b]The definition for *jumble* was changed from *confused mixture* into *messy mixture.*
[c]*Murky* was randomly chosen as a distractor from the 7,000th level in Nation's (2017) BNC/COCA.