Studies in Second Language Acquisition (2024), **46**, 200–213
doi:10.1017/S0272263123000219

RESEARCH REPORT

Understanding L2-derived words in context: Is complete receptive morphological knowledge necessary?

Batia Laufer

University of Haifa, Israel
Email: batialau@research.haifa.ac.il

(Received 22 June 2022; Revised 20 February 2023; Accepted 27 February 2023)

Abstract

The study investigates whether comprehension of derived words in text context requires a complete understanding of word parts. It explores comprehension of derived words as a function of learner proficiency and contextual clues. Ninety English-as-a-foreign-language learners at three proficiency levels participated in three successive tests representing three clues conditions, absence of clues, availability of syntactic clues, and availability of syntactic and semantic clues. They had to supply the meaning of 22 derived pseudowords constructed with nonword stems and 22 frequent affixes—for example, *stac**ement***, *gumm**ful***. The meanings of the nonword stems were provided. Test scores were compared by 3 (proficiency level) × 3 (clue condition) analysis of variance with repeated measures. The results showed effects of both variables, proficiency and clues. The largest increase in comprehension scores occurred with the addition of syntactic clues. The results imply that derived forms of familiar base words can be understood even when learners' receptive morphological knowledge is not complete.

Introduction

Lemmas and word families are different word-counting units that have been used to construct word frequency lists, design vocabulary tests for second language (L2) learners, and profile the lexical composition of authentic and learner texts. A lemma is a headword (e.g., *work*) and its inflections (*works*, *worked*, *working*), and each of these forms must be from the same part of speech. Therefore, the lemma in the example refers to *work* as a verb. *Work* as a noun and its plural form *works* is another lemma. Each derived word—that is, a base word with an added prefix and/or suffix (*worker*, *workable*) is considered a different lemma. A word family is a larger unit and includes the base word (e.g., *read*), its inflected forms (*reads*, *read*, *reading*), and its derived words with their inflections (*reader*/*readers*, *readability*, *readable*, *unreadable*). Sometimes a derived word carries a slight change in the stem as in *prepare* and *preparation*.

Examples of lemma-based lists are Brezina and Gablasova's (2015) New General Service List and Dang and Webb's (2016) Essential Word List, and a lemma-based test is Peters et al.'s (2019) VocabLab test. Examples of family-based lists are Coxhead's (2000) Academic Word List and Nation's (2006) British National Corpus Lists. Examples of family-based tests are Aviad-Levitzky et al.'s (2019) Computer Adaptive Test of Size and Strength, Nation's (1983) and Schmitt et al.'s (2001) Vocabulary Levels Test, Nation and Beglar's (2007) Vocabulary Size Test, and Webb et al.'s (2017) Updated Vocabulary Levels Test. In research, word families have been used as the counting unit in most lexical profiling studies of written and spoken production (e.g., Aviad-Levitzky & Laufer, 2013; Dang & Webb, 2014) and in studies of comprehension thresholds for reading and listening (e.g., Nation, 2006).

The choice of word-counting units in research and pedagogy has lately generated discussions among scholars (e.g., Dang & Webb, 2016; Kremmel, 2016; Laufer & Cobb, 2020; McLean, 2018; Nation, 2016, Stoeckel et al., 2020) that culminated in eight invited critical commentaries on the topic in the December 2021 issue of *Studies in Second Language Acquisition*. For over 2 decades, researchers have advocated a flexible approach to using different counting units for different purposes. The premise of the flexible approach is that the selection of lexical units depends on research and pedagogical purpose and learner variables such as vocabulary size and proficiency. For example, the Academic Spoken Word List by Dang et al. (2017) was developed for learners at a variety of levels, and, therefore, the list versions were made up of either word families or lemmas. Similarly, over years, Laufer and colleagues used lemmas in studies on the acquisition of new words, (e.g., Laufer & Osimo, 1991; Laufer & Rozovzki-Roitblat, 2015) but word families when measuring global vocabulary knowledge (e.g., Aviad-Levitzky et al., 2019; Laufer & Aviad-Levitzky, 2017).

Support for the word family as an appropriate unit of counting rests on the assumption that form and meaning similarity between word family members makes unknown derived forms relatively easy to comprehend in context and learn (Bauer & Nation, 1993). For example, if learners know one member of a word family (*avoid*), then with relatively little effort they may also understand other members of this word family in context *(avoidance*, *avoidable*, *unavoidable*), even if these were not explicitly taught. Understanding novel derived words is possible because of learners' use of contextual clues and their receptive morphological knowledge. This knowledge means that learners are aware of the fact that a word can be made up of smaller parts, that some of the parts appear in other words as well, and that learners are familiar with the meaning and function of these recurrent parts (Nation, 2013; Tyler & Nagy, 1989). For example, learners recognize three parts (morphemes) in *unavoidable*, *un* ~, *avoid*, ~ *able*. If they know the meaning of the parts, they comprehend the word. As it is easier to remember the related word family members than totally unrelated items, it may be more pedagogically sound to teach different base words together with affixes than to teach different members of each word family separately at different times.

However, in recent years, the validity of word families as counting units has been questioned and, instead, the almost sole use of lemma-based word lists, tests, and text profiles has been suggested (e.g., McLean, 2018; Stoeckel et al., 2020). The argument for the lemma as the counting unit rests on the assumption that word family as the counting unit of tests and text profiles is inappropriate because most learners do not possess, or cannot use, the morphological knowledge that is necessary to understand the meaning of a derived word even if they know the meaning of the base word. For

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

example, if *center* and *develop* are known but *decentralization* and *antidevelopment* are not, then it is argued that learners' knowledge tested by a family-based test is overestimated and text difficulty profiled by a family-based profile is underestimated.

To my knowledge, there are only four studies that have attempted to answer the question of whether knowledge of base words extends to derived forms by comparing learners' comprehension of base words and their derived forms (Laufer et al., 2021; McLean, 2018; Snoder & Laufer, 2022; Ward & Chuenjundaeng, 2009). In one additional study (Stoeckel et al., 2020) learners were compared on knowledge of pairs of lemmas that had an identical form but different parts of speech—for example, *walk* (v/n). McLean (2018) and Ward & Chuenjundaeng (2009) indeed showed that learners could not comprehend a large number of derived words even when they knew the base word, and Stoeckel et al.'s learners who knew one part of speech of a word did not necessarily know the other part of speech. In these studies, the participants were mostly of low and intermediate English-as-a-foreign-language (EFL) levels as reflected in their vocabulary test results. Ward and Chuenjundaeng's (2009) Thai students knew 25%–50% of the base words from the Academic Word List. One hundred seventy-six of McLean's (2018) Japanese students knew 3,000 words, and 84 students knew less than 2,000. Only 17 participants knew 5,000 words. Almost all of Stoeckel et al.'s (2020) participants were at the A2 or B1 CEFR (The Common European Framework of Reference for Languages) level. The authors claim that the data showing that their learners did not know many derived words provides ample evidence against the validity of the word family in teaching and testing.

However, it is questionable whether the data can be generalized to learners whose language proficiency and mother tongue are different. Studies with English as L1 children (Nagy et al., 2003; Wysocki & Jenkins, 1987) and English as L2 learners (Laufer et al., 2021; Mochizuki & Aizawa, 2000; Sasao & Webb, 2017; Snoder & Laufer, 2022) indicate that receptive morphological knowledge develops with the growth in language proficiency, particularly with an increase in vocabulary size. For example, Laufer et al. (2021) and Snoder and Laufer (2022) found that L1 speakers of Hebrew and Swedish who scored ~ 5,000 on a vocabulary size test had almost identical knowledge of base words and derived words. Learners with a vocabulary size of ~ 3,000 word families knew 60% of derived words when base words were known. Thus, even though learners may not possess enough knowledge of affixes at the early stages of learning, studies suggest that knowledge of derivations is likely to increase with vocabulary size to a point at which it is similar to that of base words.

A common feature of all the above studies that compared knowledge of base words and related derived words is that learners saw the target items in isolation, or in sentences that did not give away the meaning. Here are two examples of test items:

Example 1 (Learners are asked to translate the underlined target item).

The teachability of that idea is low (McLean, 2018).

Example 2 (Learners are asked to choose the correct meaning from four options).

RESTORATION: We planned the restoration (Laufer et al., 2021).

a. telling a story b. giving it to a different person
c. lowering its price d. making it like new again

Such test formats do not adequately represent real reading comprehension because derived words in texts occur in context. These formats are suitable for testing

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

receptive morphological knowledge—that is, the ability to comprehend derived words by recognizing word parts, specifically the affix in the word, and combining the meaning of stems with the meaning and the grammatical function of affixes. For example, if learners know what *teach* means, know the meaning of *~able*, and know that the suffix *~ity* added to an adjective changes it into a noun, they will understand the meaning of *teachability* without any contextual clues.[1] However, the lemma supporters equate receptive morphological knowledge with comprehension of derived words in texts and disregard the possibility that learners who do not know the affix might still be able to infer the meaning of the derived word based on their understanding of its base word and the surrounding context. Laufer (2021) explains why receptive morphological knowledge in tests does not reflect comprehension of derived words in texts. Even though text context may not provide the necessary clues for completely unfamiliar words, the case of derived words is different because knowledge of the meaning of a base word *is* a clue to the related derived word. Put differently, if the base word is known, the derived word is unfamiliar only in part, in the affix, and can be understood from the familiar base together with the surrounding context.

To my knowledge, the question of whether receptive morphological knowledge reflects comprehension of derived words in text context has not been investigated yet. The paper seeks to examine the question empirically by comparing comprehension of derived words in isolation and in two types of context, syntactic and semantic. The results may show that comprehension of derived words is similar in isolation and in context, which will mean that knowledge of a base word or one family word does not extend to comprehending other derived words. If so, word-family-based tests and profiles may not provide an accurate picture of learners' lexical comprehension. The results may also show that derived words are understood in context better than in isolation, supporting the claim that receptive morphological knowledge may not be identical to comprehension of derived words in text context. In other words, even if knowledge of base words does not extend to derived words in isolation, it may do so when contextual clues are available. If knowledge of base words can extend to comprehending derived forms in context, the objection to family-based tests of receptive vocabulary and lexical profiles of texts may be unnecessarily exaggerated.

The current study

Research questions

The study asked the following research questions:

1. How well do EFL learners in Grades 8, 9, and 12 understand derived words of familiar stems in three "clue conditions": In isolation (without contextual clues), in semantically neutral sentences (with syntactic clues only), and in meaningful sentences (with syntactic and semantic clues)?
2. What is the effect of the three clue conditions on comprehending derived words in Grades 8, 9, and 12?
3. What is the effect of learner proficiency (school grade) on comprehension of derived words in each clue condition?

[1]No similar assumption is made about producing unknown related words. For example, theoretically, the noun of *observe* could be *observement*, *observion*, *observal.* Thus, knowledge of *observe* does not mean the learner will also be able to produce the correct form of the noun, *observation.*

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

4. Is there an interaction between the two variables clue condition and learner proficiency?

Method

Participants

Ninety junior high (Grades 8 and 9) and high school (12) EFL learners at three proficiency levels from three intact classes in an Israeli public school participated in the study. Twenty-two students were in Grade 8 and had studied English for four and half years at the time of the experiment, 28 were in Grade 9 and had studied English for five and half years, and 40 learners were in Grade 12 and had studied English for eight and half years. In terms of CEFR levels, Grade 8 roughly corresponds to A1–A2, Grade 9 to A2, and Grade 12 to B1 (State of Israel—Ministry of Education, 2020). They were all L1 speakers of Hebrew. As participants who spent time abroad were excluded from the study, the main source of English for all the learners was school instruction that is guided by the national syllabus. Though the learners were not interviewed on a personal basis, my experience with the educational system and the learners' out-of-school digital activities led me to believe that most of their English could be attributed to classroom instruction. Participants with learning disabilities were not included either. Even though the participants constituted a convenience sample, the uneven gap between the grades had some advantages. A comparison of 8th and 9th graders would show how receptive morphological knowledge and comprehension of derived words could change over one school year. The data of the 12th graders would show receptive morphological knowledge and comprehension of derived words 3 years later, in the last year of high school. The students took part in the study after they had received an explanation about the study's purpose and benefit. They knew that participation was voluntary, that test scores would not affect their school evaluation, and that their privacy would be protected.

Materials

The target items were 22 derived pseudowords constructed with nonword stems and 22 frequent affixes—for example, *stacement*, *gummful* (Appendix 1). The stems were taken from the list of English plausible nonwords devised by Paul Meara for use in Yes/No tests (e.g., Meara & Buxton, 1987) and retrieved from Tom Cobb's Lextutor site (https://www.lextutor.ca/freq/lists_download/pnwords.html). By choosing to use pseudowords with real affixes instead of real derived words, the possibility of students' prior knowledge of the target items was eliminated.

The affixes that were added to the nonstems (five prefixes and 17 suffixes) were the 22 most frequent affixes from a list of affixes that Laufer and Cobb (2020) compiled in a corpus of English texts (~ 250,000 words) that included academic texts, newspaper articles, authentic novels, and graded readers. These 22 affixes constituted 98% of all the affix tokens in the corpus. Ten suffixes changed the stems into nouns (~ion, ~al, ~ation ~ment, ~ity, ~er, ~or ~ance, ~ness, ~age), six into adjectives (~y, ~ative, ~ ist, ~able, ~ic, ~ful), and one into an adverb (~ly). Five prefixes modified the meanings of the stems (un~, in~, re~, pro~, ex~).[2] In deciding on the combination of the nonstems

[2]The unequal numbers of prefixes and suffixes and noun affixes compared with adjective affixes and a single adverb affix reflect the dispersion of affixes in the corpus analyzed by Laufer & Cobb (2020).

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

and the affixes, in most cases the affix was different from the affix of the real word translation. For example, if *pring* meant *recommend*, the target noun was *pringal*, not *pringation*, as in *recommendation*. The exception was the adverbial affix *~ly*.

The study included three written tests, each test representing one condition. The conditions were comprehension of derived words without any clues, comprehension in sentences with syntactic clues, and comprehension in sentences with syntactic and semantic clues. In each test, the meanings of the nonword stems were provided and identical questions were asked about each target item, as in the following examples.

Condition 1: Derived items in isolation.
If *stace* means "to participate," what does *stacement* mean? _____

Condition 2: Derived items with syntactic clues.
If *stace* means "to participate," what does *stacement* mean in the following sentence?
I am asking for your *stacement*
Stacement means _____.

Condition 3: Derived items with syntactic and semantic clues.
If *stace* means "to participate," what does *stacement* mean in the following sentence?
Full and active *stacement* in school activities is required of all students.
Stacement means _____.

In Condition 1, the target derived items appeared in isolation and correct comprehension required learners' recognition of affixes as word parts and knowledge of the meaning and function of the target affixes. Thus, Condition 1 was a test of receptive morphological knowledge. The sentences in Condition 2, syntactic clues, included the basic English sentence structures that had been taught to the students in the early stages of instruction—for example, Noun phrase–Verb phrase, Noun phrase–Verb phrase–Noun phrase (direct object), Noun phrase–Copula–Adjective, etc. The content of the sentence frames did not give away the meaning of the target items,—for example, He was _____; He acted _____. Lextutor analysis showed that in Condition 2, all the words in the sentences were from the first 1,000 most frequent word families in BNC/COCA (the British National Corpus and The Corpus of Contemporary American English) and one word was from the second 1,000. In Condition 3, 97% of the words in the sentences were from the first 2,000 most frequent word families. Several words from the third 1,000 were translated for the participants. Two English teachers read the sentences and made sure the target words were inferable from the semantic clues in Condition 3. The sentences were piloted with several learners whose proficiency was similar to the participants and who did not take part in the study. As a result, some vocabulary was simplified.

As the study used a within-subject design, all the learners took the three tests. They received the meanings of the nonword stems in L1 and provided the meanings of the derived forms in L1 too. The reliability values (KR 20) of the three tests were as follows: Test 1 = .89, Test 2 = .78, Test 3 = .79. As these values are larger than .70, they indicate an acceptable internal consistency (Thompson, 2010).

Procedure

Before taking the tests, participants received a short training session that had the same format as the tests. The first two training examples were with two real derived words,

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

and the next two were with nonword stems. The training was necessary as the participants had never worked with nonwords and had never participated in an experimental study. After the class teacher made sure students understood the task, she distributed Test 1 (derived items in isolation). Upon test completion, the tests were collected and Test 2 (derived items with syntactic clues) was administered to the same participants. After it had been completed and collected, the same participants received Test 3 (derived items with syntactic and semantic clues). The training session and the three tests took place during a double lesson of 90 min and were all completed before the end of the lesson. (See Supplementary Material for the training session and tests).

Data analysis

The test answers were scored dichotomously. The correct meaning of the derived word was credited with 1 point. Because the meanings of the nonword stems were predetermined, only one correct answer was possible for each derived word that was modified by an affix. No answer or a wrong meaning was given 0 points. Each student received three scores per three conditions. Each score was the sum of the correct answers in one condition.

The data were analyzed by a 3 × 3 (grade levels by conditions) analysis of variance with repeated measures using IBM SPSS and the *MOTE* package in R (Buchanan et al., 2019; R version 4.2.0). The normality of the data distribution was tested by a Shapiro–Wilks normality test. It showed that the test was robust for the small violation of normality by the outliers. Bartlett's test of homogeneity of variance showed that the variances in the three class grade groups were homogenous in Conditions 1 and 2 and not homogeneous in Condition 3. Mauchly's test of sphericity indicated that the assumption of sphericity had been violated. Therefore the Greenhouse–Geisser correction was used.

Learner proficiency (Grades 8, 9, and 12) was the between-subject variable, and the clue condition was the within-subject variable. In the subsequent post hoc tests, pairs of school grades were compared by Tukey post hoc tests (which account for multiple comparisons) in each condition and pairs of clue conditions by using paired *t* tests (with the Bonferroni correction setting the *p* value at .017) in each school grade. A significant interaction between the two main variables would indicate whether different school grades were affected differently by the clue conditions.

Results

Research Question 1 asked how well the participants in Grades 8, 9, and 12 understood derived words of familiar stems in three clue conditions: in isolation (without contextual clues), in semantically neutral sentences (with syntactic clues only), and in meaningful sentences (with syntactic and semantic clues). To answer it I used descriptive statistics and calculated mean scores of each school grade in each clue condition.

Table 1 shows that receptive morphological knowledge as reflected in comprehension of derived words in isolation was low in the 8th grade. A mean score of 7.7 out of 22 items is 35%. It improved slightly a year later in the 9th grade, to 50%, and more so, to 74%, by the end of high school. When the derived target items appeared in sentences with syntactic clues only, comprehension scores increased in the three participant groups. The largest increase was in Grade 8 where the mean score almost doubled,

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

Table 1. Comprehension of derived words in isolation and in two types of context

Clue condition	Grade 8 $N = 22$	Grade 9 $N = 28$	Grade 12 $N = 40$
No clues	7.70 (5.4) [5.3, 10.1]	11.1 (5.0) [9.2, 13.0]	16.3 (3.7) [15.1, 17.5]
Syntacticclues	15.0 (3.4) [13.6, 16.5]	16.4 (3.4) [15.1, 17.7]	19.6 (2.4) [18.9, 20.4]
Semantic clues	16.0 (4.1) [14.2, 17.8]	17.0 (3.1) [15.8, 18.2]	19.7 (2.4) [18.9, 20.4]

Note. Values are reported as Mean (SD) [95%CI].

reaching 68% of correct answers. The smallest increase was in Grade 12 because the highest score in the no clues condition left relatively little room for improvement. Comprehension improved from 74% to 89%. The 9th graders improved from 50% to 74.5%. The third condition, syntactic and semantic clues, led to an additional small increase in comprehension scores: 8th, 9th, and 12th graders understood 73, 77, and 89.5% of derived words, respectively, when they appeared in sentences with semantic clues.

Research Question 2 asked about the effect of the three clue conditions on comprehending derived words. Research Question 3 asked about the effect of learner proficiency (school grade) on comprehension of derived words. Research Question 4 addressed the interaction between the two main variables, clue condition and class grade. To answer the three questions, the data were analyzed by analysis of variance with repeated measures. The effects of the clues condition, the school grade, and their interaction were significant, and according to Cohen (1988), the effect sizes were large ($\eta_p^2 > .14$). Table 2 shows the results of the analysis.

Pairs of school grades were compared by Tukey post hoc tests in each condition and for pairs of clue conditions by paired *t* tests in each school grade.

Tukey post hoc tests showed that in the no clues condition the three participant groups were significantly different from each other (Grades 8–9, $p < .05$; Grades 8–12, $p < .001$; Grades 9–12, $p < .001$). In the two other conditions, syntactic clues and syntactic and semantic clues, Grades 8 and 9 were different from Grade 12 (Grades 8–12, $p < .001$; Grades 9–12, $p < .001$; Grades 8–12, $p < .001$; and Grades 9–12, $p < .01$, respectively), but there was no significant difference between Grades 8 and 9. The pair comparison results demonstrate the interaction between the main variables, clue condition and proficiency. When clues were available, learners in Grades 8 and 9 achieved similar results, when they were not, the two groups were different. Paired *t* tests that examined the differences between pairs of clue conditions showed that the Condition 1 (no clues) vs. Condition 2 (syntactic clues) and Condition 1 (no clue) vs. Condition 3 (semantic clues) differences were significant in all the school grades ($p <$

Table 2. Effects of clues and language proficiency (school grade)

Variable	Effect	Effect size η_p^2	95% CI
Clue condition	$F(2, 174) = 220.37^{***}$	0.72	[0.63, 0.78]
School grade	$F(2, 87) = 24.13^{***}$	0.36	[0.19, 0.50]
Clue × Grade	$F(4, 174) = 12.52^{***}$	0.22	[0.11, 0.32]

$^{***}p < 0.001$.

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

.001 for all pair comparisons), with Cohen's (1988) effect sizes larger than 1.4, except in one pair comparison where the effect size was 1.19. According to Plonsky and Oswald (2014) effects above 1.4 for within-subject comparisons are considered large and between 1.0 and 1.4 medium. The difference between Conditions 2 and 3 was nonsignificant, as the p value was larger than .017 (following the Bonferroni correction). Specifically, the results were as follows:

Grade 8: Conditions 1–2: $t\,(21) = -9.543$, $p < .001$, $d = 2.03$.
Conditions 1–3: $t\,(21) = -10.56$, $p < .001$, $d = 2.25$.
Conditions 2–3: $t\,(21) = -1.741$, $p = .096$, $d = 0.37$.

Grade 9: Conditions 1–2: $t\,(27) = -8.503$, $p < .001$, $d = 1.61$.
Conditions 1–3: $t\,(27) = -8.298$, $p < .001$, $d = 1.57$.
Conditions 2–3: $t\,(27) = -1.69$, $p = .10$, $d = 0.32$.

Grade 12: Conditions 1–2: $t\,(39) = -8.920$, $p < .001$, $d = 1.41$.
Conditions 1–3: $t\,(39) = -7.579$, $p < .001$, $d = 1.19$.
Conditions 2–3: $t\,(39) = -0.187$, $p = 0.85$, $d = 0.03$.

The results show that the greatest improvement in comprehension of derived words occurred with the addition of syntactic clues. Adding semantic clues did not increase the scores significantly.

Figure 1 presents the results for all the research questions graphically. The different starting points of the groups in Condition 1 reflect significant differences in receptive morphological knowledge. The sharp increase in scores with the addition of syntactic clues (Condition 2) and the nonsignificant increase with the addition of semantic clues (Condition 3) show how the addition of clues affected comprehension of derived words. The figure also shows how similar (not significantly different) Grades 8 and 9 become in Conditions 2 and 3 as opposed to Condition 1 (the interaction between clue condition and proficiency).

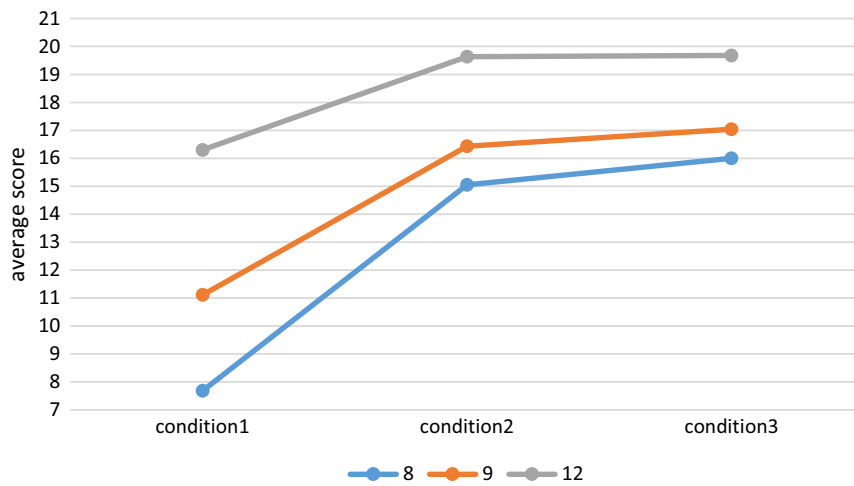

Figure 1. Changes in test scores of 8th, 9th, and 12th graders in each clue condition.

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

Discussion

The study investigated whether learners of three proficiency levels understood derived words differently when they appeared in isolation, in sentence context with syntactic clues only, and in sentence context with syntactic and semantic clues. The meanings of the pseudo base words were provided. The results showed that the addition of syntactic clues improved comprehension significantly. The addition of semantic clues, however, did not increase the scores significantly.[3] This pattern appeared in all three proficiency groups. The groups were different from each other in their receptive morphological knowledge, as reflected in the results of the first test, comprehension of derived words in isolation. Once clues were added, the performance of the 8th and the 9th graders became similar and the 12th graders performed significantly better than the younger participants did.

Pedagogically, all the findings of the study are encouraging. The differences between learner groups in test one, without clues, showed that receptive morphological knowledge improved as proficiency developed. This was the case of learners one year apart (8th and 9th graders in the study), and more so after three additional years (12th graders). At the beginning of the last year of high school, learners comprehended 74% of the most frequent affixes. These results corroborate other studies that have examined receptive morphological knowledge at different proficiency levels. In Laufer et al. (2021), Hebrew-speaking learners increased their knowledge of derived words from 60% in 9th grade to 84% in 12th grade. In Snoder & Laufer (2022), Swedish-speaking learners improved from 84% in the 9th grade to 91% in the 12th grade. In Mochizuki & Aizawa (2000), the proportion of affixes understood by Japanese-speaking learners with different vocabulary sizes was 45% in the 2,000-word size group, 61% in the 3,000-word size, and 70% in the 4,000-word size group. The most advanced learners, 10% of the sample who knew 5,000 word families, understood 77% of affixes. The results of the present study indicate a developmental pattern similar to that which was previously observed among L1 children (Wysocki & Jenkins, 1987) and L2 learners for their productive knowledge of derivatives (Iwaizumi & Webb, 2023). Morphological knowledge grows with lexical and general language proficiency.

When receptive morphological knowledge is partial, comprehension of derived words improves in context. The study shows that even 8th graders acquired basic sentence structures and could use them to figure out the part of speech of the target words. This grammatical information together with the meaning of the base words improved comprehension of derived words considerably in Condition 2, particularly in Grade 8 where it almost doubled, from 35% to 68%. The major contribution of syntactic as opposed to semantic clues is pedagogically encouraging. Not all text contexts are rich enough in semantic clues to facilitate understanding unfamiliar words, and sometimes clues are available but appear in words that learners may not understand (Laufer, 1997, 2005). The findings of the study suggest that learners understand many derived words even in semantically opaque contexts, with the help of sentence structure.

The findings do not mean that comprehension of derived words is ideal. Comprehension of ~ 75% (Grades 8 and 9) and ~ 90 % (Grade 12) means that about one in four and one in 10 derived words, respectively, may remain unclear. However, to understand whether these figures signal a text comprehension problem, we have to relate them to

[3]In the case of the five items with prefixes, semantic clues were more influential than in suffixed words. The average scores (out of 5) were 1.9 without clues, 2.5 with syntactic clues, and 3.1 with semantic clues. However, this difference did not affect the overall pattern of comprehension without and with clues.

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

the total number of derived words that learners may encounter in texts. Laufer and Cobb (2020) showed that derived words are distributed differently in texts at different language levels. In graded readers, the average percentage of derived words is about 3%. This is the kind of texts learners in Grades 8 and 9 (CEFR levels A1–A2) are likely to read. Not understanding one in four derived words may decrease the total number of comprehended vocabulary by 0.75%—that is, by 1-2 words in a text of 200 words. The 12th graders (CEFR B1 level) may read novels and, possibly, authentic argumentative prose. Laufer and Cobb calculated 5% of derived words in novels and 7.7% in academic and newspaper texts. If 12th graders do not understand one in 10 derived words, the decrease in the comprehended vocabulary is approximately 0.5% in novels—that is, two words in a text of 400 words—and 0.77% in newspapers/academic texts—that is, three in 400 words. These figures suggest that the gaps in understanding derived words may not be detrimental to text comprehension.

The results of the study indicate that complete receptive morphological knowledge may not be required to understand derived words in text context. This finding provides evidence that is counter to the claim that knowledge of base words does not extend to understanding derived forms in text context. As mentioned in the background section, studies that found poor results tested derived words in isolation, or meaningless contexts. The results of the present study support Laufer's (2021) claim that "derived words in tests are not derived words in texts" (966). Put differently, receptive morphological knowledge is not identical to comprehension of derived words during a reading or listening task. The former means comprehension of a derived word in isolation based on identification of word parts and comprehension of their meaning and grammatical function. The latter means recognizing the base word and its meaning and using the clues in the sentence structure and possibly the sentence content to arrive at the meaning of the derived word.

A possible counterargument to the optimistic approach to understanding derived words in texts could be that learners may not recognize familiar base words when they appear in combination with affixes. For example, a learner may know what *develop* means but not recognize *develop* in *developmental.* This is possible, particularly at low levels before learners have developed awareness of word parts. However, there is evidence from error analysis studies that at some point learners tend to decompose words into smaller units. Laufer (1989) identified several error-provoking categories of words she called "deceptively transparent," or pseudofamiliar. One category consisted of words with deceptive morphological structures—for example, *infallible*, *outline*, *discourse*, and *falsities*—that were misinterpreted as *unable to fall*, *outside a line*, *in the wrong direction*, and *falling cities*, respectively. These errors show that learners may look for smaller familiar units inside words and construct meaning from these parts and pseudoparts. This tendency shows that learners become aware that smaller language units combine to form larger units. Learners who have not developed such awareness could benefit from instruction of word structure and of meaning and function of affixes.

Concluding remarks

The study has some limitations that can be addressed in future studies. It tested the most frequent affixes in the corpus compiled by Laufer & Cobb (2020) for their study on derived words in texts, one affix per target item. Therefore, the number of prefixes was considerably lower than the number of suffixes. Future studies can explore a larger number of prefixes because it is plausible that comprehension of prefixed words is more dependent on semantic than syntactic clues. They could also investigate how words

https://doi.org/10.1017/S0272263123000219 Published online by Cambridge University Press

with multiple affixes—for example, *unavoidable*, *multinational*, or *directionality*—are comprehended.

For administrative reasons, the participants in the study were not tested on vocabulary knowledge. Even though the results showed the effect of proficiency level on the receptive morphological knowledge and comprehension of derived words, it would be useful to relate learner vocabulary size to knowledge of affixes and the use of contextual clues.

Finally, semantic clues were provided in sentences, not in texts. It would be more ecologically valid to test the use of semantic clues in text context. However, it was impossible to find a suitable authentic text that included derived words with 22 target affixes that appeared in sentences with semantic clues, a text that was short enough to read and understand, particularly by the less proficient 8th and 9th graders. Furthermore, had I constructed such a text, it would be highly artificial because the percentage of derived words in authentic texts read by A1–A2 CEFR learners is less than 3% (Laufer & Cobb, 2021). If, in the future, such texts can be found and administered in a reasonable class time, the experiment will have a better ecological validity.

In spite of the limitations, the present study is to my knowledge the first that indicated that full receptive morphological knowledge may not be necessary to understand derived words in texts. It provides evidence against the generalization that most learners are unable to understand different derived forms of base words even when the base words are familiar and the subsequent conclusion that the word family as a counting unit in tests and text profiles is invalid. The study showed that even if the affixes were unknown and the derived words were not understood in isolation, the clues in the meaning of the base words, sentence structure, and possibly sentence content led to 89% comprehension in Grade 12 and 68% and 75% in Grades 8 and 9, respectively. The partial comprehension of 8th to 9th graders may not be detrimental to reading because the students read relatively simple texts with a small number of derived words.

The argument against family-based counting units is that text comprehensibility decreases considerably when the receptive morphological knowledge of the learner is not complete (McLean, 2018). The results of our study imply that such dire warnings are unnecessary exaggerations. Receptive morphological knowledge of learners grows with language proficiency, and it can reach near perfection, as in the case of Swedish 12th graders (Snoder & Laufer, 2022). Before this happens, learners with partial derivational knowledge will comprehend many derived words with the help of syntactic and semantic clues.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0272263123000219.

Data availability statement. The experiment in this article earned an Open Data badge for transparent practices. The materials are available at https://iris-database.org/details/JVpzi-26R2J

Acknowledgments. I am grateful to my student, Livnat Elhadad Dahan, for collecting the data.

Competing interest. I have no competing interests in writing this paper.

References

Aviad-Levitzky, T., & Laufer, B. (2013). Lexical properties in the writing of foreign language learners over eight years of study: Single words and collocations. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 127–148). EuroSLA Monograph Series 2.

Aviad-Levitzky, T., Laufer, B., & Goldstein, Z. (2019). The new computer adaptive test of size and strength (CATSS): Development and validation. *Language Assessment Quarterly*, *16*, 345–368.

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

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics*, *36*, 1–22.

Buchanan, E. M., Gillenwaters, A., Scofield, J. E., & Valentine K. D. (2019). *MOTE: Measure of the effect: Package to assist in effect size calculations and their confidence intervals* (version 1.0.2) [Computer software]. R Foundation for Statistical Computing. http://github.com/doomlab/MOTE

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213–238.

Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning*, *67*, 959–997.

Dang, T. N. Y., & Webb, S. (2014). The lexical profile of academic spoken English. *English for Specific Purposes*, *33*, 66–76.

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

Iwaizumi, E., & Webb, S. (2023). To what extent do learner- and word-related variables affect production of derivatives? *Language Learning*, *73*, 301–336.

Kremmel, B. (2016). Word families and frequency bands in vocabulary tests: Challenging conventions. *TESOL Quarterly*, *50*, 976–987.

Laufer, B. (1989). A factor of difficulty in vocabulary learning: Deceptive transparency. *AILA Review*, *6*, 10–20.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20–34). Cambridge University Press.

Laufer, B. (2005). Instructed second language vocabulary learning: The fault in the "default hypothesis." In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 286–303). Walter de de Gruyter.

Laufer, B. (2021). Lemmas, flemmas, word families and common sense. *Studies in Second Language Acquisition*, *43*, 965–968.

Laufer, B., & Aviad-Levitzky, T. (2017). What type of vocabulary knowledge predicts reading comprehension: Word meaning recall or word meaning recognition? *The Modern Language Journal*, *101*, 729–741.

Laufer, B., & Cobb, T. (2020). How much knowledge of derived words is needed for reading? *Applied Linguistics*, *41*, 971–998.

Laufer, B., & Osimo, H. (1991). Facilitating vocabulary retention: The second hand cloze. *System*, *19*, 217–224.

Laufer, B., & Rozovski-Roitblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, *19*, 687–711.

Laufer, B., Webb, S., Kim, K. S., & 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*, 229–258.

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

Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing*, *4*, 142–154.

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

Nagy, W., Berninger, V., Abbott, R., & Vaughan, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second grade readers and at-risk fourth grade writers. *Journal of Educational Psychology*, *95*, 730–742.

Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, *5*, 12–25. https://www.wgtn.ac.nz/lals/about/staff/publications/paul-nation/1983-Testing-and-teaching.pdf

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*, 59–82.

Nation, I. S. P. (2013). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. John Benjamins.
Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*, 9–13.
Peters, E., Velghe, T., & Van Rompaey, T. V. (2019). The VocabLab tests: The development of an English and French vocabulary test. *ITL—International Journal of Applied Linguistics*, *170*, 53–78.
Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes, in L2 research. *Language Learning*, *64*, 878–912.
Sasao, Y., & Webb, S. (2017). The word part levels test. *Language Teaching Research*, *21*, 12–30.
Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, *18*, 55–88.
Snoder, P., & Laufer, B. (2022). EFL learners' receptive knowledge of derived words: The case of Swedish adolescents. *TESOL Quarterly*, *56*, 1242–1265.
State of Israel—Ministry of Education (2020). *English curriculum 2020*. Pedagogical Secretariat—Language Department English Language Education.
Stoeckel, T., Ishii, T., & Bennett, P. (2020). Is the lemma more appropriate than the flemma as a word counting unit? *Applied Linguistics*, *41*, 601–606.
Thompson, N. A. (2010). KR-20. In N. J. Salkind (Ed.), *Encyclopedia of research design*. Sage Publications.
Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory & Language*, *28*, 649–667.
Ward, J., & Chuenjundaeng, J. (2009). Suffix knowledge: Acquisition and applications. *System*, *37*, 461–469.
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*, 33–69.
Wysocki, K., & Jenkins, J. R. (1987). Deriving word meanings through morphological generalization. *Reading Research Quarterly*, *22*, 66–81.

Appendix 1. Target derived pseudowords (nonword stems with existing affixes) and their meanings in the study

(Though the meanings appear in English, learners received the meanings in L1).

Ten suffixes changed the stems into nouns (~ion, ~al, ~ation ~ment, ~ity, ~er, ~or ~ance, ~ness, ~age)

nonagrat**ion**	(prevention)	pring**al**	(recommendation)	oxyl**ation**	(production)
stace**ment**	(participation)	galpin**ity**	(precision)	snell**er**	(surgeon)
bodelat**or**	(respirator)	adair**ance**	(refusal)	eckett**ness**	(determination)
venn**age**	(domination, control)				

Six suffixes changed the stems into adjectives (~y, ~ative, ~ ist, ~able, ~ic, ~ful)

moffat**y**	(poor)	trok**ative**	(calm adj.)	balfour**ist**	(industrial)
degat**able**	(changeable)	acklon**ic**	(rude)	gumm**ful**	(afraid)

One suffix changed the stem into an adverb (~ly)

quorant**ly** (carefully).

Five prefixes modified the meanings of the stems (un~, in~, re~, pro~, ex~)

unwray	(destroy)	**in**opie	(dishonest)	**re**berrow	(vote again)
prochanning	(pro-war, militant)	**ex**troke	(empty, take out)		

Cite this article: Laufer, B. (2024). Understanding L2-derived words in context: Is complete receptive morphological knowledge necessary? *Studies in Second Language Acquisition,* 46: 200–213. https://doi.org/10.1017/S0272263123000219