International Journal of Linguistics and Translation Studies
Volume 2, Issue 4, 2021
Homepage: http://ijlts.org/index.php/ijlts/index
DOI: https://doi.org/10.36892/ijlts.v2i4.189

Morphological Awareness and Word Complexity in an EFL Context

Ahmed Zrig
University of Carthage, Tunis, Tunisia
zrigahmed@yahoo.fr

ARTICLE HISTORY

Received: 01/04/2021

Accepted: 29/09/2021

KEYWORDS

Morphological awareness, receptive vocabulary, productive vocabulary, vocabulary knowledge, word frequency, word complexity.

Abstract

This study examined the relationship between morphological awareness and word complexity (simple versus complex words) in an EFL context. The participants in this study were 100 fourth year secondary school Arts students in Tunisia. Students' morphological awareness was measured by the Morphological Awareness Test. Vocabulary size was tested using an adapted version of Nation's (2001) Vocabulary Levels Test as a receptive measure of vocabulary size. Half of the vocabulary test items were made complex to check the participants' performance on simple and complex words. The informants' scores were high on the overall morphological awareness task, and the best performance was on inflectional morphemes. This could be very useful for teachers to build on, improve, and construct better future teaching practices. Finally, morphological awareness differentiated between students' performance on simple versus complex words.

1. INTRODUCTION

Vocabulary knowledge is a language component necessary for fluent language use (Nation, 2001). Melissa et al. (2019) maintain that "Vocabulary is key to the development of comprehension in all students" (p.2). In other words, vocabulary is the sine qua non of language learning and vocabulary size is an indicator of second or foreign language learners' performances on language skills such as reading, writing, listening and speaking (Bear, Invernizz, Templeton and Johnson 2019). Having inadequate vocabulary knowledge may even hamper learners' discourse comprehension (Ellis, 1997). So, vocabulary teaching and learning is an essential activity in any language class. Learners' vocabulary learning can be enhanced by using some learning strategies, which are consciously or unconsciously learned by students to process information and enhance comprehension, learning and retention (O'Malley and Chamot, 1990). One of these strategies is employing morphological awareness to learn new lexical items.

Morphological awareness; hereafter MA, is defined as "the ability to use the knowledge of word formation rules and the pairing between sounds and meanings" (Kuo & Anderson, 2006). By the use of morphological awareness, students are able to learn morphemes and morphemic boundaries by disassembling complex words into smaller meaningful parts. The practice of dissembling-reassembling words is called *morphological analysis*.

1.1.LITERATURE REVIEW

The rise of communicative methodology in the late 1970's has led to the resurgence of vocabulary as an essential component of successful second/foreign language learning.

> Vocabulary, thus, became a central part of the language program in general and in the teaching practice more specifically. Read (2004) stated that this interest in vocabulary stems mainly from the observation that 'limitations of vocabulary knowledge hamper the learners' ability to communicate effectively in the target language, since lexical items carry the basic information load of the meanings they wish to comprehend and express. The research literature in second language vocabulary learning has shown the importance of knowing a considerable number of words in order to be able to function appropriately in the target language (Duin and Graves, 1987; Walker, Greenwood, Hart and Carta, 1994; Nation, 2001; Read, 2004; Tschirner, 2004; Zimmerman, 2005). For further details see Read (2004). There is a general consensus that around 2000 word families afford the lexical basics for learners to speak conversationally in an L2 (Nation and Meara, 2000; Schmitt, 2000).

Morphological awareness has gained a snowballing interest as a crucial strategy of vocabulary knowledge, mainly in reading. For instance, Singson, Mahony and Mann (2000) argued that morphemes have phonological, semantic and syntactic properties that communicate the function of a specific word in the reading context (e.g. –s in the verb *drives* reveals that the doer of the action is only one person. and the action takes place in the present tense). In addition, morphological awareness enhances the learner's awareness of the writing system (Kuo and Anderson, 2006). That is to say, morphological knowledge helps learners to perceive better spelling and phonological irregularities (e.g. *sign- signature*). Studies show that language learners encounter complex words at early stages of their learning (Gordon, 1989; Carlisle and Stone, 2003). The fact that students encounter many derived words in their reading has motivated researchers to explore further the contribution of morphological awareness in vocabulary improvement. Despite the importance of morphological awareness in the development of learners' vocabulary knowledge, together with literacy-related skills (e.g., reading and spelling), little research has really investigated the relationship between breadth and depth of VK in relation to MA (Al Farsi, 2008; Wang, Cheng, & Chen, 2006). In other words, more work is needed to be carried out to identify the relationship between both receptive and productive VK and different aspects of MA, and in particular, the nature of this relationship in foreign language learning.

1.2. Research Questions

1. Are measures of morphological awareness related to measures of English receptive and productive vocabulary sizes?
2. Is morphological awareness related to students' performance on simple and complex words?

2. METHODOLOGY

2.1. Sample / Participants

The participants in this study were Tunisian fourth year secondary school students. Two groups of 50 students each, representing two entire Baccalaureate of Arts classes have participated in the study.

2.2. Instrument(s)

To answer the present research's questions two widely used tests were adapted to the purposes of the study: A Vocabulary Level Test with 2 subsets: receptive and productive, and a Morphological Awareness Test with its 2 subsets: morpheme identification and morphological structure. Besides, a

questionnaire was created to provide a cross-check between what the test results show and the participants' ideas and attitudes towards the different testing instruments. Two vocabulary tasks have been used, namely Nation's Vocabulary Levels Test as a receptive measure of vocabulary size and Laufer and Nation's (1999) productive version of the Vocabulary Levels Test. Both tests were adapted for the purpose of this study. Originally, both tests consist of five levels of word frequency in English, the 2,000, 3,000, 5,000, Academic Vocabulary and the 10,000 most frequently occurring word families (90 words in total).

For level appropriateness of the informants only the 2,000 and 3,000-word levels have been used. As far as the receptive vocabulary test is concerned, each level includes 10 sections; each section comprises 10 words on the left side instead of 6 as in the original version with 5 meanings on the right instead of 3 to make the test items more representative of the total vocabulary items at each frequency level. So, 100 words are used in both the 2,000 and 3,000 levels. Half of these words are simple and the other 50 are complex words. This modification is meant to examine the relationship between morphological awareness and students' performance in simple versus complex words. The learners' morphological knowledge was tested by the *Morphological Awareness Test (*Chang et al. 2005). This test consists of two parts: a morpheme identification awareness test and a morphological structural awareness test. There is an important change made to the two tests in this present study. The test items in this study are in a written form, not oral as in Chang et al. (2005). The main reason behind this change is the practicality in administering the test.

2.3. Data collection and analysis

This study used a mixed design more specifically a "sequential explanatory design" in which the larger focus was on quantitative data. In addition to the vocabulary and morphology tests a questionnaire was created to provide a cross-check between what the test results show and the participants' ideas and attitudes towards the different testing instruments. In this study, sometimes a "between group design" was used to make comparisons between independent groups: control and experimental group. Some other times, a "within subjects design" was used to compare data from the same students: pre-test and post-test. The tests have been administered over two days to minimize fatigue. The first day of testing consisted of the VLTs. The second day of testing included the morphological awareness test. Once all of the informants; control and experimental groups have finished with this first step which is mainly aimed to answer the first research question, the researcher systematically selected half of the participants and introduced them to morphological awareness as a strategy of vocabulary acquisition (class instruction). Students of this study have little knowledge about morphemes. Also, they have never attended a class on morphology. During the two-week treatment period which consisted of 4 separate hours per week, the experimental group received explicit instruction on inflectional and derivational morphemes etc. One month after the end of the treatment period, a post-test (same test as the pre-test with the items scrambled) was administered to both groups in order to investigate the second research question. In order to answer the first research question and highlight the differences in the vocabulary knowledge that were employed by each group of participants, the results of all the participants and the separate results of each group were compared through Two-Way ANOVA. As the data in this study is quantitative; it allows performing various arithmetic operations to find statistics of the sample, and since the aim is to compare students' performance on the dependent and independent variables, the mean and standard deviation have been

used to provide more details about research question 1, which investigates the participants' receptive and productive vocabulary size, the results of the VLTs were summarized by means and standard deviations across the two different levels (2,000 and 3,000) to make the comparison possible.

As far as the second research question about students' morphological awareness is concerned, the whole group's results were reported at the beginning of the analysis for two reasons. First, results helped to have a general overview of students' performance on the different subtests of the morphological awareness test and to check possible differences between students' performances at these tasks. Since this study used a pre-test post-test non-equivalent group design: matched pair t-tests were used to compare groups' performance on two conditions: before treatment/ after treatment. After these tests, a factorial Two-Way AVOVA was carried out with repeated-measures variable: Morpheme Type: roots, derivational, inflectional and compounds, and one between- subjects: Group: experimental group and control group, to check the performances of the two groups on the morphological awareness pre-tests and post-tests. Finally, independent measures t-tests were also used to compare performance of the experimental and control group in more detail. The results of the morphological awareness test were also analyzed through the mean and standard deviation, in order to compare the participants' performances on different parts: roots, inflectional morphemes, compounds and derivational morphemes of the test. In light of research question 2, Pearson's product-moment correlation was used to investigate the correlations between each of the Vocabulary Level Tests total scores and Morphological Awareness Test. First, scores of all participants were tested for correlation between these variables before the treatment.

3. FINDINGS AND CONCLUSION

3.1. The participants' performance on receptive versus productive vocabulary tests

Table 1. Means and standard deviations for scores on the RVT and PVT (N=100)

Frequency levels	Receptive Vocabulary Test Mean	SD	Productivetive Vocabulary Test Mean	SD
2000-Word Level	16.83	4.08	13.48	4.12
3000-Word Level	13.08	3.69	9.79	3.12

Note: A maximum score at each frequency level is 24 for the RVT and the PVT

It is obvious from Table 1 that the students' performance on the receptive vocabulary test was better than their performance on the productive vocabulary test at both frequency levels. In addition, the participants' performances were highest at the 2,000-word level especially on the receptive tasks as manifested by the mean score 16.83. Based on Table 1, it is worth mentioning that the gap between students' performances on the receptive and productive vocabulary tests was wider at the 2,000-word

level where the mean values of both types of vocabulary size were significantly high (16.83 and 13.08). These values increased with dropping the frequency levels (13.48 and 9.79). In fact, these findings are supported by Webb (2008) and Hayashy and Murphy (2009).

Based on our sample data, the difference between means of both types of vocabulary size is 3.75, 95% CI [2.99, 4.51]. There appears to be a significant mean difference between students' performance in the receptive and productive vocabulary tests since zero is not captured in this interval, and since the entire interval is above zero. Furthermore, the difference between means of both frequency levels is 3.69, 95% CI [2.98, 4.4]. Here again, we are 95% confident that this interval captures the true mean difference, and that this mean difference is statistically significant as the CI does not cross the point of no difference.

As it was stated in the method section, p values emanating from experimental design studies do not tell much about the difference between the control and the experimental group. Hence, Cohen's d was used as an effect size to report the size of this difference. As far as the frequency level is concerned, results revealed that the participants performed better at the 2,000 word-level on both the receptive and productive vocabulary test: (r=.43, d=.96) and (r=.45, d=1.00) respectively.

Accordingly, the participants' performance at the receptive 2,000 word-level was 0.96 standard deviations greater than their performance at the 3,000 receptive word-level. In other words, the average participant's performance at the receptive 2,000 word-level was higher than 82% of students' performances at the 3,000 receptive word-level. So, the average participant who ranked 50th in the 2,000 word-level test ranked 20th in the 3,000 word-level test. The ES of .96 indicates a nonoverlap of 51.6% in the two distributions and that the amount of variance in the receptive vocabulary test by membership in the 2,000 word-level and 3,000 word-level is 16.8%.

3.1.1. The groups' performances on the vocabulary level pretests

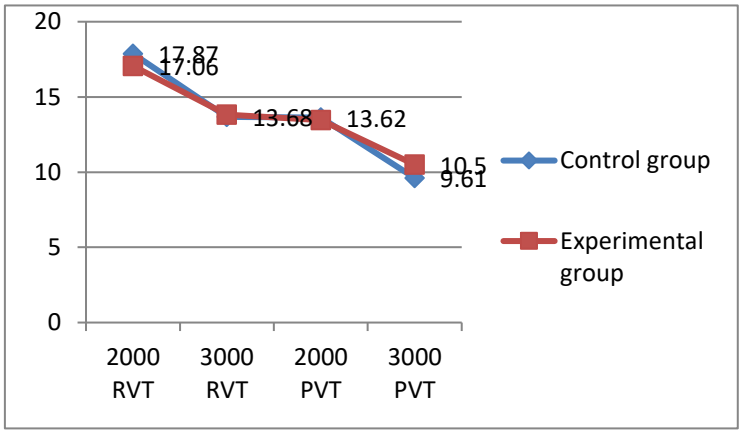


Fig 1. Line charts showing groups' mean scores on the pretest

Figure 1 shows that the experimental group was not initially better than the control group as demonstrated by their scores on the pretest. So, the groups' performances overlap greatly. It is obvious that this is a case of disordinal interaction, as the lines representing performances of both groups intersect. Most importantly, is the fact that these two lines go in the same direction asserting that, both groups performed best at the 2,000-word receptive level. The RVT data were analyzed using a factorial

analysis of variance: Two-Way ANOVA with one repeated-measures variable: Frequency level 2,000-word and 3,000-word levels, and one between-subjects variable: Group experimental and control. A main effect of frequency level was observed, $F\ (1, 92) = 8.64$ $p < 0.5$, suggesting that the two groups performed better at the 2,000-word level. However, the obtained *F ratio* was not sufficiently larger than the critical value of *F* to confirm the existence of a difference between groups $F\ (1, 92) = 1.43$, $p > 0.5$ and the *F* critical value is 4.03. Therefore, there was no significant difference between groups' performances on the receptive vocabulary level pretest. Furthermore, there was no significant Group × Frequency level interaction $F\ (1, 92) = 3.67$, $p > 0.5$.

A main effect of frequency level was observed with the PVT data, $F\ (1, 92) = 8.70$, $p < .05$, suggesting that the two groups performed better at the 2,000-word level. No group effect was observed, $F\ (1, 92) = 2.73$, $p > .05$. As such, there was no significant difference between groups' performances on the productive vocabulary level pretest. Moreover, there was no significant Group × Frequency level interaction, $F\ (1, 92) = 3.39$, $p > .05$.

Based on data from both groups, the difference between means of the 2000-word level (RVT) is 0.81, 95% CI [-0.6, 2.22]. There is no significant mean difference between performances of both groups since zero is captured in this interval, and since the CI crosses the point of no difference zero. Also, the measured difference between means of the 3000-word level (RVT) is 0.15, 95% CI [-1.26, 1.56]. Here again, we are 95% confident that this mean difference is not statistically significant since the CI includes the value zero.

As far as the productive vocabulary test is concerned, the mean difference between the performance of the control and the experimental group at the 2000-word level is 0.16, 95% CI [-1.25, 1.57]. There is no statistically significant mean difference between performances of both groups since zero is captured in this interval. In other words, the CI crosses the point of no difference. The difference between means of the 3000-word level (PVT) is 0.89, 95% CI [-0.52, 2.3]. Here again, we are 95% confident that this interval captures the true mean difference, and that this mean difference is not statistically significant as the CI crosses the point of no difference.

3.1.2. The groups' performances on the vocabulary level posttests

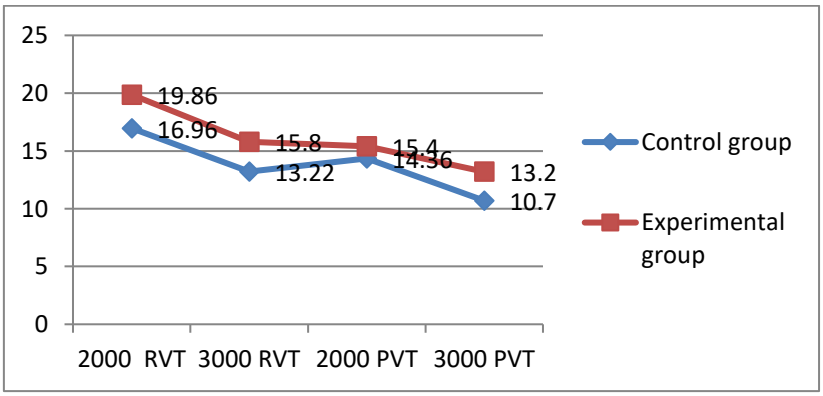


Fig 2. Line charts showing groups' mean scores on the posttest

The RVT data were analyzed using a factorial analysis of variance (Two-Way ANOVA) with one repeated-measures variable (Frequency level: 2,000-word, 3,000-word levels), and one between-subjects variable (Group: experimental and control). A main effect of frequency level was observed, $F\ (1, 92) = 8.06$ $p < 0.5$, suggesting that the two groups performed better at the 2,000-word level. Also, there was a main group effect, $F(1,92) = 4.68$ $p < 0.5$. Consequently, there was a significant difference between groups' performances on the receptive vocabulary level posttest. In addition, there was a significant Group× Frequency level interaction, $F(1, 92) = 37.72$, $p < 0.5$. Consequently, we can conclude confidently that the treatment did show a difference. In other words, the experimental group scores differ significantly from the control group scores on the receptive vocabulary posttests.

When checking the performances of the two groups on the PVT, a main effect of frequency level was observed, $F\ (1, 92) = 5.14$ $p < 0.5$. This suggests that the two groups performed better at the 2,000-word level. Also, there was a main group effect, $F(1,92) = 6.97$ $p < 0.5$, suggesting there was a significant difference between groups' performances on the productive vocabulary level posttest. Finally, there was a significant Group× Frequency level interaction, $F(1, 92) = 35.82$, $p < 0.5$. The critical value of F for both group and frequency level is 3.80. The obtained *F values* exceeded this *F critical value*. So, we can have confidence in concluding that the treatment did show a difference. In other words, the experimental group scores differ significantly from the control group scores on the productive vocabulary posttests.

The difference between means of both groups at the 2000-word level (RVT) is 2.9, 95% CI [1.35, 4.45]. There is a significant mean difference between performances of the control and experimental group since zero is not captured in this interval, and since the CI is above zero. Also, the measured difference between means of the 3,000-word level (RVT) is 2.58, 95% CI [0.85, 4.3]. Here again, we are 95% confident that this mean difference is statistically significant since the CI does not cross the point of no difference.

In line with the vocabulary language test, it was of crucial importance to make an intra-group comparison (repeated-measure) for the morphological awareness test. In other words, it was necessary to compare the control group's performances on the pre- and posttest as well as those of the experimental group to check the effectiveness of the instructional treatment.

3.2. *Performance on the Morphological Awareness Tests*

As stated previously, the morphological awareness tasks were analyzed using first matched pair t-tests in order to compare the control and the experimental group separately on two conditions (pretest/ posttest), and then independent measures t-tests to compare performance of the two groups (experimental/ control). Besides, the results of the morphological awareness test were also analyzed through the mean and standard deviation, in order to provide more details about the participants' performances on different parts (roots, inflectional morphemes, compounds and derivational morphemes) of the test.

As stated in the method, the whole group's results (100 students) were reported to get first a general overview of students' performance on the different subtests of the morphological awareness test, and to be correlated latter with results from the VLT to answer the second research question. The difference between students' performance on the different parts of the morphological awareness test was insignificant as Table 2 shows.

Table 2. Mean and standard deviation for the morphological awareness subtests (N=100)

Morphological Awareness Test	Mean	SD
Root words test	7.48	1.42
Derivational morphemes test	7.78	1.59
Compound words test	7.82	1.57
Inflectional morphemes test	8.18	1.65

Note: A maximum score at each morphological awareness subtest is 10

Students' scores at the different morphological awareness tasks were higher than 7 out of 10, exept scores at inflectional morphemes which were slightly higher than 8 out of 10. That is to say, the results on the different morphology tests were the same with insignificant differences between them. In addition, the standard deviations of the different morphological awareness subtests varied very little, which means that students' scores were rather homogeneous. In line with the vocabulary language test, it was of crucial importance to make an intra-group comparison (repeated-measure) for the morphological awareness test. In other words, it was necessary to compare the control group's performances on the pre- and post-test as well as those of the experimental group to check the effectiveness of the instructional treatment.

3.2.1. The control group performances on the pre and post morphology tests

Table 3. The control group performances on the pre and post morphology tests (N=50)

	Pre Test		Post Test	
Morphological Awareness Test	Mean	SD	Mean	SD
Roots	7.54	1.40	7.60	1.00
Derivational morphemes	7.70	1.62	7.73	1.66
Compounds	7.62	1.87	7.57	2.03
Inflectional morphemes	8.00	2.00	8.17	1.98
Total (%)	77.15		77.67	

Table 3 shows that the total pre-test mean score was almost the same as the post- test for the control group. They were ($\bar{x}$=77.15) and ($\bar{x}$= 77.67) respectively. Similar to the whole group performance, the control group's scores at the different morphological awareness tasks were higher than 7/10, exept scores at inflectional morphemes which were exactly 8/10.

Matched-pair t-tests were conducted to check the performances of the control group on the morphological awareness pretest and posttest. First, the control group's performance on the roots pre-test ($\bar{x}$=7.54 S.D= 1.40) and post-test ($\bar{x}$=7.60 S.D=1.00) reveals that students' scores improved slightly as it can be seen through the mean difference between the pre-test and the post- test (MD=0.06), with a t (29) = 0.09, p = 0.84. However, this t_{obs} value did not exceed the t critical value of 1.69. Consequently, we can conclude confidently that scores did not differ significantly from the pretest to the posttest, and that slight increase was due to chance. Actually, this may be explained by the absence of the treatment for the control group. Second, as far as the derivational morphemes test was concerned, Students' mean score between the pre-test ($\bar{x}$= 7.87, S.D=1.53) and the post-test ($\bar{x}$=7.7, S.D= 1.12) did not witness a significant improvement: t (29) = 0.03, p = 0.94. Once again, this insignificant improvement can be explained by the absence of the instructional treatment. Third, the comparison between the control group's performance in the compound-word pre-test ($\bar{x}$=7.62 S.D= 1.87) and post-test ($\bar{x}$=7.57 S.D=2.03) reveals a regression in students' scores. Though this regression is not significant, it shows that the control group did not witness any improvement, which may be explained once again by the absence of the instructional treatment for the control group and may be the difficulty of the test itself.

Finally, the comparison between the control group's performance in the inflectional morphemes pre-test ($\bar{x}$=8.00 S.D= 2.00) and post-test ($\bar{x}$=8.17 S.D=1.98) reveals that students' scores improved slightly. However, the mean difference between the pre- and the post- test (MD=0.17) was not significant at the 0.5 level. t (29) = 0.16, p = 0.74 . The following line charts illustrate the control group' mean performances on the pre- and post tests.

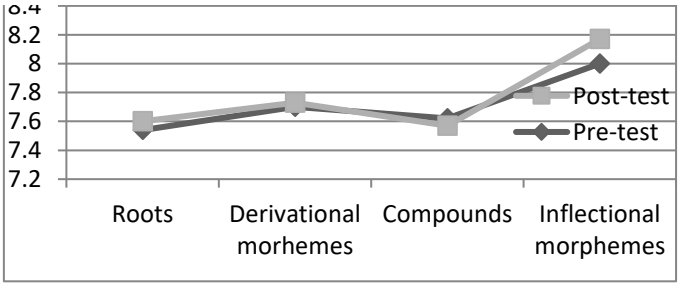


Fig. 3: Line charts showing the control group's mean scores on the pre and post morphological awareness tests

This graph better shows the consistent results of performances on the pre and post tests. In technical terminology, there is an insignificant *"disordinal interaction"* between the pretest and the posttest. That is to say, the results on the different morphology tests were almost the same with three insignificant improvements and one regression on compounds. Besides, the graph illustrates that the two lines of the chart overlap very much. This means that the control group scores on the pre and post tests are almost similar. The similarity of scores is an indication of the test-retest reliability of the

morphology test. The next part of the analysis is devoted to the performances of the experimental group.

3.2.2. The experimental group performances on the pre- and post morphology tests

First, the comparison between the experimental group's performance on the roots pre-test ($\bar{x}$=7.43 S.D= 1.45) and post-test ($\bar{x}$=7.9 S.D=0.96) shown in Table 4 below reveals that students' scores improved as it can be seen through the mean difference between the pre-test and the post- test (MD=0.47). It was necessary to calculate the *observed t* value to check the significance of this difference: t_{obs} = 1.91. In spite of this small t_{obs} value, it still exceeds the *t critical value* of 1.69. Consequently, we can conclude confidently that the experimental group scores differ from the pretest to the posttest.

Table 4. The experimental group performances on the pre and post morphology tests (N=50)

	Pre Test		Post Test	
Morphological Awareness Test	Mean	SD	Mean	SD
Roots	7.43	1.45	7.9	0.96
Derivational morphemes	7.87	1.53	7.7	1.12
Compounds	8.07	1.53	8.63	1.71
Inflectional morphemes	8.37	1.30	8.66	1.06

Second, as far as the derivational morphemes test is concerned, there was a regression in the experimental students' mean score between the pre-test ($\bar{x}$= 7.87, S.D=1.53) and the post-test ($\bar{x}$=7.7, S.D= 1.12). Though this regression was not significant, it shows that the experimental group did not witness any improvement, which may question the effectiveness of the instructional treatment and the difficulty of the test itself.

Third, the comparison between the experimental group's performance on the compound-word pre-test ($\bar{x}$=8.07 S.D=1.53) and post-test ($\bar{x}$=8.63 S.D=1.71) reveals that students' scores improved as highlighted by the small improvement shown through the mean difference (MD= 0.56). This means that, post-test results of the experimental group were slightly better than those of the pre-test as confirmed by the t-test result: t_{obs} = 1.99. Obviously, we can conclude that the experimental group

scores differ significantly from the pretest to the posttest because the t_{obs} value exceeds the t critical value.

Finally, the comparison between the experimental group's performance on the inflectional morphemes pre-test ($\bar{x}$=8.37 S.D= 1.30) and post-test ($\bar{x}$=8.66 S.D=1.06) reveals that students' scores improved. However, the mean difference between the pre- and the post-test (MD=0.29) was not significant at the 0.5 level. t (29) = 1.18, p = 0.27 . So, this small improvement does not provide insightful evidence of the effectiveness of the instructional treatment. The following graph summarizes the experimental group ın performances on the pre- and post tests.

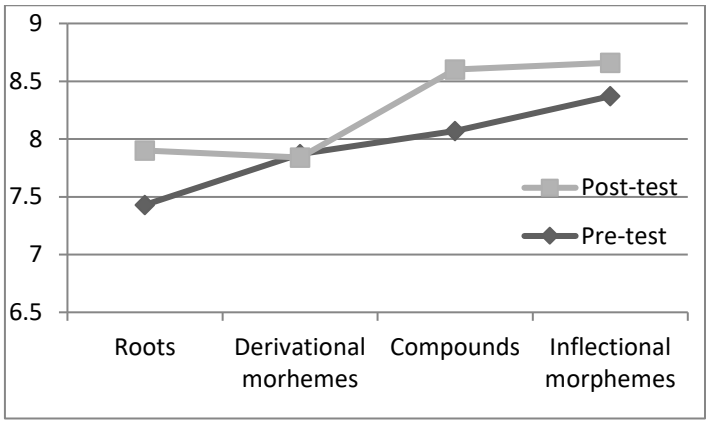


Fig. 4: Line charts showing the experimental students' mean scores on the pre and post morphological awareness tests

A significant *"disordinal interaction"* is obvious from this figure: the differences between means not only vary, but change order across the 4 levels leading the lines representing the pretest and posttest to cross. Three major facts can be concluded from these line charts. First, students' scores on roots, compounds, and inflectional morphemes witnessed a significant change. Second, performances on derivational morphemes remained the same on the pre- and post-test. Third, the participants performed best on inflectional morphemes. Concerning the first fact, this slight improvement of scores on the majority of the tasks may give an insightful evidence of the effectiveness of the treatment. As far as the second and the third remarks are concerned, they can be explained by the challenging nature of derivational morphemes and the relative easiness of inflectional morphemes (Koda, 2000; Singson, Mahony and Mann, 2000). Below are the groups' performances on morphological awareness pre-tests.

3.3. Morphological Awareness and Receptive and Productive Vocabulary Knowledge

The second research question investigated the relationship between English morphological awareness and vocabulary knowledge of the participants. Mean scores of the whole group (100participants) in the VLTs (receptive and productive) and the morphological awareness tasks were

correlated using Pearson's product-moment to assess the strength of association between the tests. Besides, the performances of each group (experimental and control) in these tasks were correlated as well.

As far as the whole group is concerned, morphological awareness tasks correlated significantly with both the RVT and the PVT, although the strength of association was not high. Pearson's product-moment correlations were .42 and .32 at $p < .05$, with $r_{crit} = 0.25$ (see Table 5). The correlation of the morphological awareness test and RVT scores was comparatively higher compared to those of the morphological awareness test and the PVT results. The findings for the group as a whole reveal that there is a significant relationship between morphological awareness and vocabulary knowledge. This is supported by Chang et al. (2005) who pointed out that since this is a correlational study, the causal relationship among the two variables could not be predicted.

Furthermore, results reveal a different pattern of association between the two groups (Control and experimental). As shown in Table 5, the morphological awareness scores obtained by students in the control group reached statistical significance when correlated with the scores of their receptive and productive vocabulary tests. On the other hand, the experimental group scores on the morphological awareness test correlated significantly with their receptive and productive vocabulary scores than the control group. Also, the table shows that on the whole, the strengh of the relationship between the different correlations can be considered as moderate.

Table 5: Intercorrelations between the vocabulary language test and morphological awareness test

	Morphological Awareness Test Whole group	Control	Experimental
Root words test	0.42	0.37	0.47
Derivational morphemes test	0.32	0.29	0.36

Note. Correlation is significant at the .05 level (1-tailed)

3.4. Morphological Awareness and Word Complexity

Question 2 examines if the performance on morphological awareness test differentiates between students' performance on simple vs. complex words on the vocabulary test. First, students' performances on simple and complex words are reported through means and standard deviations in

order to check the existence of any difference between them. After that, their performance on the morphological awareness test is correlated with their performance on simple then on complex words using Pearson's product-moment correlation.

3.4.1 Word Complexity

Table 6 presents the descriptive statistics of the students' performance on simple versus complex words at each word level of the VLT.

Table6: Mean and standard deviation of students' scores on simple words versus complex words for each word level of the vocabulary level test. (n=100)

Vocabulary Level Test	Mean	SD
2000 word level		
Simple words	20.93	3.12
Complex words	16.95	3.59
3000 word level		
Simple words	15.65	3.76
Complex words	11.78	4.23

Note: A maximum score at each frequency level is 24.

It is obvious that the participants performed better with simple words than they did with complex words at both word levels. To illustrate, the participants' best performance is found on simple words at the 2,000-word level ($\bar{x}$= 20.93, S.D= 3.12), while their mean score with complex words at the same word level is ($\bar{x}$= 16.95, S.D= 3.59). This difference in participants' scores is statistically significant since the value t_{obs} = 9.21 exceeds the *t* critical value of 2.00 at the .05 alpha level. Similarly, for the 3,000-word level, the difference between simple and complex words is statistically significant: t_{obs} = 7.58. Consequently, the null hypothesis can be rejected and we can have confidence in concluding that students performed better in simple words than in complex words. Finally, a close look at standard deviations reveals that scores get closer to the mean when we decrease the frequency level. In other words, scores became more homogeneous at the 2,000-word level.

After presenting descriptive statistics, it was necessary to report the size of this difference. Results displayed that the participants performed better on simple words at both word levels. First, as far as the 2,000-word level is concerned, results revealed that the participants performed remarkably better

on simple words than on complex words: (r=.50, d=1.18). In other words, the participants' performance on simple words was on average 1.18 standard deviations greater than their performance on complex words at the 2,000 word-level. In view of this, the average participant's performance on simple words was higher than 88% of students' performances on complex words. So, the average participant who ranked 50th in the 2,000-word level test on simple words ranked 12th in the 2,000-word level test on complex words. The ES of 1.18 indicates a nonoverlap of 62.2% in the two distributions, and that the amount of variance in the 2,000-word level test by membership in simple words and complex words is 26%: see Cohen (1988, pp. 21-23) for descriptions of measures of nonoverlap.

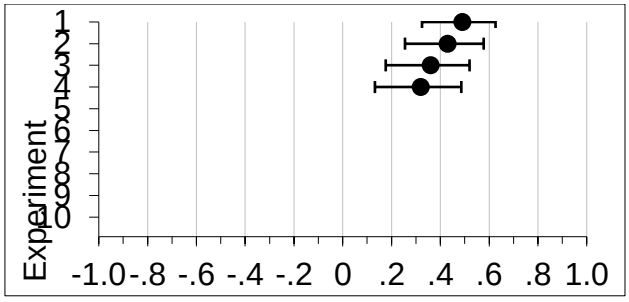
Second, concerning the participants' performance at the 3,000-word level, again, students performed better on simple words than complex words: (r=.43, d=.96). The participants' performance on simple words was on average .96 standard deviations greater than their performance on complex words at this frequency level. In other words, the average participant's performance on simple words was higher than 82% of students' performance on complex words. So, the participant who ranked 50th in the 3,000-word level test (simple words) was on a par with the participant who ranked 20th in the 3,000 word-level test (complex words). The ES of .96 indicates a nonoverlap of about 51.6% in the two distributions, and that the amount of variance in the 3,000-word level test by membership in simple words and complex words is 17%. In conclusion, after presenting statistics of students' performances on simple and complex words on the RVT, it was necessary to check the relationship between them and MAT scores.

3.4.2. Correlation between morphological awareness and word complexity

Table 6. Pearson's product-moment correlation of simple versus complex words for each vocabulary level and morphological awareness (N= 100)

Vocabulary Knowledge Test	MAT	95% Confidence Interval		
		LB	UB	MOE_{av}
***2000** complex words*	.49	..32	.62	.15
***2000** complex words*	.43	..25	.57	.16
***3000** simple words*	.36	..17	.52	.17
***3000** complex words*	.32	..13	.48	.17

Significance at the .05 level (one-tailed)



1 correlation between 2000 simple words and MAT.

2 correlation between 2000 complex words and MAT

3 correlation between 3000 simple words and MAT.

4 correlation between 3000 complex words and MAT

***Fig 3*:** Confidence intervals on the correlation between simple words, complex words and MAT

Table 6 displays the relationship between simple vs. complex words and morphological awareness of all the participants. It is worth noting that most of the students' vocabulary size falls within the 2,000-word level. Also, the morphological awareness test correlates best with the 2000 simple words .49, 95% CI [.32, .62]. The students' performance on simple as well as complex words at the 2,000-word level is positively correlated with their performance on the morphological awareness test. This relationship persists when considering the 3,000-word frequency level. Again, it can be assumed that there is a real and moderate relationship. In addition, performance on simple words at both frequency levels is slightly better correlated to the Morphological Awareness Test than complex words. Overall, it is evident from Table 6 that students who score high on the morphological awareness test tend to obtain a higher score on simple words than complex words.

4. DISCUSSION

Results reveal that students' performance on the vocabulary language test correlated positively with the morphological awareness test. Also, results revealed that morphological awareness correlated positively with both receptive and productive vocabulary tests. A striking fact is that *pearson's r* was insignificant for the control group in the productive vocabulary test. Regardless of these differences, the direction of the relationship is positive. Actually, this relationship can be best viewed via Henriksen's model: morphological awareness as part of the depth of knowledge continuum is related to the receptive-productive continuum. This means that, if a student's morphological awareness is high, his or her vocabulary knowledge is more likely to be high too. So, morphological awareness and vocabulary knowledge are both crucial subcomponents for the development of general linguistic knowledge.

Assuming that "large" effects are always more important than "small" or "medium" ones is unjustified. David Funder (2012) argued that correlation of .30 is commonly regarded as a quite modest correlation can be considered in certain contexts as medium correlation coefficient. Similarly, some educational researchers have indicated that effect sizes around 0.20 are of policy interest *when they*

are based on measures of academic achievement (Hedges & Hedberg, 2007). For example, a study with an effect of 0.20, which at first glance, might be misinterpreted as a "small" effect if one automatically applies Cohen's original conventions, can be an important outcome in some research areas. So, the interpretation of the effect sizes has to be made in relation to the context of the research in question.

Three meaningful guidelines have been used in order to evaluate effect sizes in context. The first guideline is the source or the quality of the research itself. This means the association of the study to prior and to new findings. Secondly, it is important to make comparisons across similar research conditions (measurements, study design etc.) Finally, it is not only the magnitude of the effect that is important, but also its practical or clinical value must be considered. Clinical value reflects the extent to which there has been meaningful change in participants' lives.

In conclusion, based on these guidelines the relationship between morphological awareness and vocabulary knowledge in this study can be considered on the whole as moderate. As it was stated previously, an effect size by itself can mean almost anything. A "small" or a "moderate" effect size, as in this study, can be important and have practical value since small correlations are the most common correlations in the social and behavioral sciences. The reason for this is that, most variables are affected by numerous factors. Besides, it is important to keep in mind that a large correlation is not a correlation of .90. Correlations of this size are often between two different measures of the same variable. Such large correlations often indicate not a meaningful relationship between variables, but an artificial one (Kenny: 1987).

The second research question investigated whether morphological awareness is related to simple and complex words. Actually, the participants' performance on the morphological awareness test correlated positively too with word complexity. This positive correlation is real and moderate through the 2,000 as well as the 3000-word frequency levels. It is important to put the results in the context of other studies. Results support the findings of Sternberg (1987) White, Power and White (1989) whose studies focused on college students or even school-aged children. However, the findings are inconsistent with those of Al Farsi (2008), who found that the relationship between morphological awareness and vocabulary size, and morphological awareness and word complexity could not be established due to some factors such as floor effect, task difficulty, and instruments item design. First, she claimed that the floor effect in scores at the synthetic section of the Morphological Awareness Test impacted the relationship between morphological awareness and vocabulary size and the relationship between morphological awareness and word complexity as well, since many of the participants' scores were found at the bottom of the performance scale. That is to say, in a study of 54 Omani students, eight students scored zero; two students scored 3 and two students scored 5 out of 15. In fact, the researcher related this floor effect to the other factor of task difficulty.

Second, Al Farsi (2008) claimed that task difficulty might have contributed to the failure to observe the relationship between morphological awareness and vocabulary size on the one hand and vocabulary size and word complexity on the other hand. She added that, the design of the synthesis section of Morphological awareness Test was inappropriate for the participants' level in her study. Consequently, McBride-Chang's *et al.* (2005) Morphological awareness Test should not be merely adopted, but adapted to suit the participants' level of language proficiency. Finally, the disparity between the results of this study and other studies might in fact be due to differences either in research instruments or the statistical tests (Pearson/ Spearman) used to correlate morphological awareness and word complexity.

To conclude, some of the students in the present study were not able to recognize the morphological structure of complex words. From a cross-linguistic variation perspective, Arabic morphology might have hindered the students from unlocking the meaning of English complex words. In other words, the affixes of Arabic complex words are inseparable from their roots, that is both affixes and roots are bound morphemes. So, splitting Arabic complex words into their meaningful constituents is non-sense, keeping in mind the fact that Arabic morphology is root-and-pattern morphology. Students, who were unable to appreciate the separability of bases from affixes, decoded an unfamiliar English complex word as a whole. Consequently, they could not unlock the meanings of newly encountered complex words because they fail to apply patterns and rules in solving problems. Some other students can unlock the meaning of new lexical items because they have more developed analytical thinking. Beginning Arab learner and may be intermediate, may transfer this opaque structure of Arabic morphology to English morphology. Consequently, it is necessary to explicitly raise secondary school level Tunisian students' awareness of English morphological knowledge. This need for teaching morphological units stems from two reasons. First, morphological awareness is likely to lead to better learning outcomes, since it is related to various language skills such as spelling (Bear, Invernizzi, Tempelton Templeton, & Johnston, 2008), vocabulary growth, and reading comprehension (Fowler & Liberman, 1995; Qian, 2002). Second, it has been argued that learners are able to use their morphological awareness to arrive at the meaning of complex words (Carlisle, 1995; Carlisle and Stone, 2003).

5. CONCLUSIONS

First, the results showed that the students exhibited an average overall morphological awareness of word formation rules. However, the participants performed better at the inflectional morphemes task than the other tasks. Results revealed that morphological awareness correlated positively with both receptive and productive vocabulary tests. Also, the present study displayed that there was a positive association between morphological awareness and simple word but not complex ones. These findings could be very useful for teachers to build on, improve, and construct better future teaching practices. Beck et al. (2002) argued that promoting learners' vocabulary knowledge as well as their morphological knowledge could be good predictors for academic success. That is to say, students would move from learning to read to reading to learn independently and become autonomous learners. Hence, the introduction of morphological awareness raising as a vocabulary building strategy could be of much help for students to boost their vocabulary repertoire and therefore could be included in the curriculum. In conclusion, it could be useful to duplicate this study in other schools and with a different population because the findings of this study cannot be extrapolated.

REFERENCES

Al Farsi, B. (2008). *Morphological awareness and its relationship to vocabulary knowledge and morphological complexity among Omani EFL university students*. Unpublished M.A. thesis: University of Queensland.

Bear, D.R., Invernizzi, M., Tempelton, S., & Johnston, F. (2008). Words Their Way: Word Study for phonics, vocabulary, and spelling instruction (4th edition). Upper Saddler River, NJ: Pearson Prentice Hall.

Bear, D.R., Invernizzi, M., Tempelton, S., & Johnston, F. (2019*). Words Their Way: Word Study for phonics, vocabulary, and spelling instruction* (7th Ed.). Upper Saddler River, NJ: Pearson Prentice Hall.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). Bringing words to life: robust vocabulary instruction. New York: Guilford Press.

Carlisle, J. F & Stone, C. A. (2003). The effect of morphological structure on children's reading derived words in English. In E. M. Assink, & D. Sandra (Eds.), Reading complex words (27-51). New York: Kluwer Academic.

Chang, C. M., Wagner, R. K., Muse, A & Chow, H. S. (2005). The role of morphological awareness in children's vocabulary acquisition in English. *Applied Psycholinguistics, 26*, 415–435.

Ellis, N. (1997). Vocabulary acquisition: word structure, collocation, word-class andmeaning. In N.Schmitt & M. McCarthy (Eds.), Vocabulary: Description, acquisition and pedagogy (122-139).Cambridge:Cambridge University Press.

Funder, D.C. (2012). Accurate personality judgment. Current Directions in Psychological Science, 21,177-182.

Gordon, P. (1989). Levels of affixation in the acquisition of English morphology. *Journal of Memory and Language, 28*, 518- 520.

Graves, M.F. (2004). Teaching prefixes: as good as it gets. In J.F. Baumann & E.J. Kame'enui (Eds.), Vocabulary instruction: Research to practice. New York: Guilford.

Hayashy, Y, & Murphy, V.A. (2009). *Morphological awareness and second language receptive and productive vocabulary knowledge in Japanese learners of English. International Conference on First and Second Languages: Exploring the Relationship in Pedagogy Related Contexts*, Oxford, England.

Hedges, L & Hedberg, E.C. (2007). Intra-class correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis, 29*: 60-87.

Kenny, D. (1987).Statistics for the social and behavioral sciences. Boston: Little, Brown.

Koda, K. (2000). Cross-linguistic variations in L2 morphological awareness. *Applied Psycholinguistics, 21*, 298- 319.

Kuo, L. J. & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist, 41*, 161-180.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. Applied Linguistics, 16 (3), 307- 322.

Melissa A. Gallagher, Ana Taboada Barber, Jori S. Beck & Michelle M. Buehl (2019).Academic Vocabulary: Explicit and Incidental Instruction for Students of Diverse Language Backgrounds. *Reading & Writing Quarterly*, 35(2), 84-101. https://doi.org/10.1080/10573569.2018.1510796

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19* (3), 304-330.

Nation, .P. (2001) Learning Vocabulary in Another Language. Cambridge: Cambridge University Press.

O'Malley, J. M., & Chamot, A. U. (1990). Learning strategies in second language acquisition. Cambridge: Cambridge University Press.

Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: evidence from derivational suffixes. *Reading and Writing: An Interdisciplinary Journal, 21*, 219- 252.

Wang, M., Cheng, C., & Chen, S. (2006). Contribution of morphological awareness to Chinese-English biliteracy acquisition. *Journal of Educational Psychology, 98* (3), 542.

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30(1), 79-95.

About the Author

Ahmed Zrig, PhD student at the Higher Institute of Languages in Tunis, University of Carthage, Tunisia. He has been working in the field of ELT for 15 years. He has taught at all levels of preparatory and secondary education. He is a Member of Tunisia TESOL, TAYR and the research unit LCF (Languages and Cultural Forms: ISLT). He was a finalist in Tech Age Teachers Program by IREX in 2014. His research areas in linguistics include reading, morphology and psycholinguistics.