



Data-driven learning of academic lexical bundles below the C1 level

*Keith J. Lay, Cyprus International University**

*Mehmet A. Yavuz, Cyprus International University**

Abstract

This study investigates the possibility and efficacy of paper-based, in-class, data-driven learning (DDL) of academic lexical bundles below the C1 level of proficiency described by the Common European Framework of Reference (CEFR; advanced high ACTFL). A two-stage experimental design involving three groups ($n = 41$) and 24 two-to-four word academic items was implemented. First, the question of whether this type of learning works with these items below the C1 level is addressed through a nonequivalent-groups quasi-experimental design covering a five-week period. The results indicate that this technique is effective at the B2 level, but not at the A2-B1 level. Next, an equivalent-groups experimental design compares this style of learning to conventional techniques at the B2 level. The results of this stage suggest that paper-based, in-class DDL is more effective than conventional learning with academic lexical bundles at the B2 level.

Keywords: *Academic lexical bundles, Corpus, Data-driven learning (DDL), English for academic purposes (EAP)*

Language(s) Learned in This Study: *English*

APA Citation: Lay, K. J., & Yavuz, M. A. (2020). Data-driven learning of academic lexical bundles below the C1 level. *Language Learning & Technology*, 24(3), 176–193. <http://hdl.handle.net/10125/44741>

Introduction

As increasing numbers of lower proficiency learners are applying for and entering tertiary level education in English Medium Instruction (EMI) contexts (Coxhead, 2011), the need to innovate in order to help those lower proficiency learners cope with English for Academic Purposes (EAP) is greater than ever before. As high academic performance is more likely from those who write with a broader vocabulary (Cooper, 2017), the challenge to compete with native speakers is significant. This is further illustrated by Csomay & Prades (2018), who found essay scores to be higher in writers with a broader academic vocabulary.

One form of academic vocabulary that has been shown to be both particularly useful and particularly challenging is that of the academic collocation. Works such as the Academic Collocations List (Ackermann & Chen, 2013) and the Academic PhraseBank (Morley, 2017) are testimony to the place of academic collocations in EAP development, an area in which second language learners at the university level may need greater assistance. However, academic multi-word items are difficult for even native speakers to learn and start using in a short amount of time, as shown by Cortes (2007). Therefore, it seems that introducing academic collocations as early as is feasible would be desirable, yet therein lies a problem: How can lower proficiency learners cope with academic collocations?

One innovation that may help these learners comes in the form of corpus study for the purpose of language proficiency development, commonly referred to as data-driven learning (DDL). In the last decade, studies such as those by Boulton (2008a, 2009, 2010), Chujo et al. (2012), Vyatkina (2016), and Brown (2017), among others, have served to map the possibilities of DDL with lower-level learners. In addition to proving successful with such learners, as a tool for vocabulary learning, DDL answers many of the central issues regarding vocabulary learning successfully. As a result, there is potential to explore EAP areas in such

*Cyprus International University, Nicosia, Northern Cyprus, Mersin 10, Turkey

contexts. This paper reports on a two-stage study aimed at investigating the possibility of incorporating DDL into the multi-word EAP vocabulary development of learners below the C1 (Advanced-high ACTFL) threshold.

Literature Review

Academic Lexical Bundles

Academic collocations represent an important and sizeable area of EAP vocabulary and those collocations come in a variety of forms. Resources such as the Academic Collocations List (ACL) compiled by Ackermann and Chen (2013) demonstrate that a great number of academic collocations are in the form of noun collocations; often adjective-noun combinations, such as ‘contextual factors’ and ‘significant impact’. A further resource, the Academic PhraseBank (Morley, 2017), consists of phrases that are of a usually greater length (often several words), such as ‘An example of X is...’, representing a collocation type more in line with terms such as *lexical phrases* (Li & Schmitt, 2009, DeCarrico & Nattinger, 1988), *lexical bundles* (Biber et al., 2004), *formulas*, and *formulaic sequences* (Ellis et al., 2008), among others. These collocation types differ from the typical noun combinations noted above as they perform a lexicogrammatical function above the level of merely labelling. In order to distinguish the item types focused upon here and denote their more common usage within academic discourse, the term *academic lexical bundles* is selected, fitting with the *recurrent formulaic sequences* that Pan et al. (2016) describe lexical bundles as representing.

Collocations that perform functions, such as those described as lexical bundles here, are significant because they have the potential to affect the coherence and cohesion of writing and as such can greatly influence performance. Academic lexical bundles are one of the defining and key factors of academic vocabulary, acting as ‘a component of fluent linguistic production’ (Hyland & Jiang, 2018). Pan et al. (2016) found that even at the highest levels of academic discourse native and non-native speakers use lexical bundles differently, with the former more likely to use phrasal type bundles and the latter incorporating a higher proportion of clausal bundles into their writing. The consequences of such differences can be significant; Csomay and Prades (2018) found that essay scores were higher in those learners who display greater academic vocabulary, while Cooper (2017) notes the contrast in the vocabulary ranges of low and high achievers.

EAP collocations pose a particular challenge to even advanced level non-native speakers of English (Ellis et al., 2008, p. 378). Byrd and Coxhead (2010) describe a number of challenges for both the teacher and learner in EAP relating to lexical bundles, including a lack of pedagogical advice, issues relating to item type, and justifying the time spent on them when the individual components may already be known. Indeed, academic collocations also pose a challenge to native speakers; a study focusing on the learning of 35 four-word bundles (selected from an 800,000 word history corpus and classified by function) during five micro-lessons in a ten week period, found that the participants’ use of the target bundles was rare both before and after the experiment (Cortes 2007, p. 395).

While it seems that a strong set of academic lexical bundles would likely improve academic scores and learner confidence, the challenge these items pose demands a learning strategy. One possible answer may come from the world of corpora.

Data-Driven Learning for Vocabulary Learning

Data-driven learning (DDL) is the term, coined by Johns (1991), which refers to language learning through the study of corpora (a collection of texts, usually computerized and vast in size). For example, the Corpus of Contemporary American English consists of over 560 million words and is searchable in a variety of ways. This allows language analysis and, as such, language learner study of the text of a corpus, or DDL. The manner of study can vary from the student having direct access to the corpus, so-called *hands-on* concordancing, to groups/classes using teacher-selected concordance lines, often referred to as *hands-off*

concordancing, and all the combinations therein.

A corpus can offer thousands of seemingly contextualized examples of an item at the click of a button. At first glance then, corpora would seem to be a magic answer in searching for meaningful, real-world examples. However, a more in-depth consideration of the central issues of vocabulary learning highlights not only positive factors, but also some potentially negative factors DDL may bring to the vocabulary learning field.

Laufer (2017) concisely defines the current debate between two schools of thought in vocabulary learning:

The first group believes that the best way to acquire many words is by reading large quantities of material. The second group claims that to meet new words 12 times during extensive reading—12 times being the average number of encounters required for acquisition to occur, according to the available body of research—L2 learners will have to read about a million words per year, which is unrealistic since they have neither the time nor the ability to do this. Hence, the argument goes, word-focused instruction is indispensable (Laufer, 2017, p. 6).

While one recent study found that 15 encounters of a collocation was a suitable number for effective retention (Webb et al., 2013), a counter argument from Laufer and Rozovski-Roitblat (2015) suggests that what is done with new lexis appears to be more significant than the number of encounters itself. Consequently, factors such as context, the style and interaction pattern that form the situation in which the item is encountered, and the spacing between encounters must all contribute to retention rates.

First, the degree to which the study of corpora provides context is disputed. While some note the benefit of concordance analysis in providing a ‘feel’ for items and enabling visualization due to the context (Boulton 2008b, p. 40), others put forward the notion that the concordance results of corpus analysis are removed from the wider context of the text and are therefore decontextualized (Flowerdew, 2005, p. 324). Many concordancers allow the selection of concordances in full sentence form and some allow the user to see the original text in full to provide the context. However, keyword in context (KWIC) concordances, a common form used in DDL, are comprised of the item and a certain number of items in each direction from that key word, limiting context potential. On the other hand, recent works have begun offering a more interactive and context-heavy interface; the TED Corpus Search Engine compiled by Hasebe (2015) being one such concordancer. Therefore, it seems likely that context will soon be less disputed as an advantage of using corpora in language learning.

In terms of style and interaction, the study of lexical items in corpus concordances involves intentional vocabulary learning. This type of learning can be described as “aiming at committing lexical information to memory” (Hulstijn, 2001, p. 272), much in line with the word focused tasks mentioned by Laufer (2017). Choo et al. (2012, p. 855), referring to Rosszell (2007), state that a combination of incidental and intentional vocabulary learning may be the best solution, with intentional study providing a boost to vocabulary learning. On that basis, paper-based, in-class DDL for specific items can be considered a potential means of providing that boost.

The retention level of new lexical items is also affected by the spacing (in terms of time) between encounters, of which Nation (2001, p. 56) makes the distinction between spaced repetition and mass repetition. The former involves spacing between encounters and the latter involves a high number of encounters at the initial stage. Ergo, specific-item DDL involves mass repetition. Early research studies, such as those by Pimsleur (1967, in Nation, 2001), show strong empirical evidence in favour of spaced repetition as opposed to mass repetition. Subsequent second language research strengthens this position of spaced repetition (for a more recent example see Hirschel & Fritz, 2013). Thus, a lack of spacing is a clear limitation. This could perhaps be somewhat countered in DDL practice by encouraging recycling, by actively recapping, or by explicitly teaching learners the benefits of spaced repetition.

The vocabulary learning benefits of DDL presented here are: The provision of a sufficient number of encounters, the potential for contextualization, the opportunity for intentional learning, and the relative ease

of access and use. The potential downsides include the nature of mass repetition, the questions regarding contextualization, and the lack of incidental or peripheral learning opportunities. As a result, with the stakes high in terms of the effect of a strong foundation of academic lexical bundles on the academic performance and confidence of learners, it is important to establish whether or not DDL is a solution for the learning of those item types, and at which level this can be done.

Data-Driven Learning with Lower Proficiency Learners

A meta-analysis of several hundred papers in a search for empirical evidence of DDL alludes to the level of participants being mostly made up of high level learners: “the majority of the studies are concerned with more or less advanced learners.” (Boulton 2007, p. 13). In the 13 years since the publication of that paper, many studies have concentrated on investigating DDL with lower proficiency learners. Although reported on according to different language proficiency descriptions, studies have focused on the whole range of lower proficiency. Here, those studies are aligned to the levels of the CEFR, a comprehensive current tool for describing language proficiency. They include: ‘basic user’ learners at the A1 (beginner) and A2 (elementary/pre-intermediate) levels (such as Vyatkina, 2016, Chujo et al., 2012, and Boulton, 2008a studied), as well as the ‘independent user’ range at the B1 (intermediate) and B2 (high intermediate) levels (studied by Boulton, 2010, 2008b, and Brown, 2017). The mapping of DDL below the high level threshold is significant because for a technique such as DDL to filter through to language learning materials, classrooms, and teacher education, there is a need for strong evidence of its value. Therefore, it is useful to consider the findings of these studies in relation to both the learner level and the language area in focus.

Some recent studies have focused on the very lowest levels. Vyatkina (2016) focused on beginner level learners of German as a foreign language in a study comparing the learning of eleven verb-preposition collocations through either DDL or through a textbook. The participants, 88 American English native speakers, were mostly at A2 level or below in German, and the results indicated that DDL was more effective than traditional learning. Chujo et al. (2012) also found evidence of the superiority of DDL over conventional methods with beginner level learners, this time focusing on noun phrases.

Moving up in level, a study of 132 learners ranging from elementary to intermediate according to a TOEIC test, investigated the possibilities of DDL with linking adverbials. The adverbials ranged in length, with seven one-word items, and one each of two-word, three-word, and four-word adverbials. The findings showed DDL to again be effective, although the overall scores were ‘very low’ at just 13.92% (Boulton, 2009, p. 9), prompting the question, were the items too difficult for the level, regardless of the method of learning?

Further positive results with lower levels come from a study focusing on phrasal verbs, measuring the performance of 113 lower intermediate participants with paper-based DDL (Boulton 2008a). As with the study above, the scores were considered low; only two items were focused upon (‘look up’ and ‘pick up’), yet the overall score was just 57% (where 50% should be returned by chance), although again the comparison between DDL and conventional methods favoured DDL (Boulton, 2008a, p. 9–10). These studies seem to suggest a pattern that DDL, while more effective, is not a miracle cure for the problem of early exposure to difficult language items. However, some respite may come when the focus turns to problematic areas that are learner-generated.

Boulton (2010) undertook a further study with lower proficiency learners and similarly to that described above, this was done in a hands-off, paper-based style of DDL. A selection of 15 problematic areas in the writing of the participants (produced in a pre-test session) were identified and selected for inclusion in the treatment, meaning the items were not being learned in the study, but developed. As the author notes, these items were typical problematic areas for French learners of English (Boulton, 2010, p. 7). This is a significant move away from the focus on intentional learning in the studies discussed above. Although the researcher again noted low scores, this study also found DDL to be more successful than conventional learning (Boulton: 2010, p. 16).

When collating the evidence, it seems that while DDL is more effective than conventional methods with

lower proficiency learners across levels and item types, what can and cannot be done at lower levels is still limited, which could be a reason for the slow progress of DDL with regards to becoming incorporated into mainstream teaching.

As the academic performance of non-native speakers of English relies on their ability to perform well in terms of vocabulary (Cooper, 2017, Csomay & Prades, 2018), one area of EAP vocabulary that may aid the cohesion and coherence that learners perceive to be problematic (Al Badi, 2015) is that of academic lexical bundles. The earlier they can start acquiring these valuable items, the greater their chance of success later on will be, and as items such as these pose a considerable challenge (Cortes, 2007), identifying the most efficient way of learning them is desirable. However, while even at low levels DDL has proven to be more effective than conventional methods (Vyatkina, 2016, Boulton, 2008a, 2009, 2010, Chujo et al., 2012), those studies have often reported low scores. Therefore, it seems prudent to establish at which level these items become a realistic option for DDL instruction in the EAP classroom.

This study was carried out in two stages. The first stage was designed to answer the first two research questions, establishing the lower limit of this type of learning with these item types.

Research Questions:

Stage One

RQ1. Can B2 level learners learn academic lexical bundles to a significant degree using paper-based in-class DDL?

RQ2. Can A2-B1 level learners learn academic lexical bundles to a significant degree using paper-based in-class DDL?

The results of the first stage justified a second stage in which this technique was compared with conventional learning at the B2 level.

Stage Two

RQ3. Can B2 level learners learn academic lexical bundles to a significant degree using conventional methods?

RQ4. Is paper-based, in-class DDL more effective than conventional learning with academic lexical bundles at the B2 level?

Methodology

Design

Participants

Three groups of participants ($n = 41$), all studying at a large EMI university in Cyprus, took part in this study; groups A and B participated in the first stage and group C in the second. Groups A and C consisted of students in their first year of undergraduate study and group B consisted of students learning (foundation) English prior to faculty entry. In order to involve participants of the same level and stage of learning as the first treatment, the second stage was conducted in the following academic year.

The participants, 29 females and 12 males with a mean age of 20.17 and predominantly of Turkish nationality, had the following first languages: Turkish (35), Arabic (3), Kurdish (2), and French (1).

Group A ($n = 12$) and group C participants ($n = 17$) had already passed a general proficiency test pegged to the CEFR and had an average CEFR level of B2 according to the mutual opinion of three of the instructors (experienced language teaching professionals) of those classes. Group B participants ($n = 12$) were studying at the pre-intermediate level of a foundation programme pegged to the CEFR and are therefore best described as A2-B1. The faculty participants were studying eight courses, all delivered in

English, while the foundation programme participants were studying English in 25 separate 50-minute lessons each week. The usual form of English instruction at the university is a focus on skills development and communicative competence. With very few classes being monolingual, English is the language of instruction in language lessons too. DDL is not usually carried out in the institution at which the research was conducted and academic lexical bundles are not a feature of courses taken by the participants, except for the occasional item occurring during normal learning conditions. The sessions of this research were incorporated into their programmes of study for the purposes of vocabulary building and English language development.

Mechanics

Six 50-minute training sessions were conducted with each group, giving a total in-class treatment time of 300 minutes per group. For the first stage, a nonequivalent-groups quasi-experimental design was chosen, incorporating a pre-test, training sessions, and post-test (the same for both groups to allow statistical comparison). The first stage took place over a period of five weeks. The second stage utilized an equivalent-groups quasi-experimental design. In order to use the results of group A as the experimental data, a control group (group C) was required. Participants in the control group were given the same pre and post-test, and received the same amount of instruction time with the same lexical bundles as groups A and B. However, the control group was taught in a conventional manner, which is described in more detail later in this section.

Planning

A total of 24 academic lexical bundles were included in the pre and post-test stages, chosen through the implementation of a selection and piloting process. The lexical bundle selection process involved five items (chosen through a randomizer) from each of the seventeen function categories listed on the Academic Phrasebank (Morley 2017) being subjected to a cross-reference search by hit per million words in the British Academic Written English (BAWE) corpus (see [Appendix A](#)). The number of hits for each lexical bundle in each category was then calculated to determine the most common functions according to the selected lexical bundles. The resultant functions were causality, signalling transition, compare and contrast, and giving examples. A second search was then conducted on each lexical bundle in the Phrasebank (Morley 2017) for those selected functions (see [Appendix B](#)), resulting in the selection of 24 lexical bundles according to the most common six per function.

A gap-fill with a phrase bank was designed for the pre/post-test (sample in [Appendix C](#)). This style of test was chosen as it is similar to the materials used in the DDL sessions of the study as well as to tasks included in existing language learning materials. Also, the degree of control provided by a gap-fill task reduced the chances of validity and reliability issues that may have arisen from alternative tasks types, such as a written task. To this end, the test items were initially divided into two 12-item gap-fill tasks for the pilot test.

A piloting process was then conducted to ascertain the most desirable experimental design. A panel of experts consisting of three academic professionals in the field of English teaching (PhD holders) were consulted prior to and following the administration of a pilot test and the re-selecting of lexical bundles. The pilot test participants ($n = 7$) were of approximately the C1 level of the CEFR and were in their final year of undergraduate study. Despite the participants of the pilot test being above the level of the intended study participants, this sample was accepted as a convenience sample owing to the benefit of the material for the participants and the assumption that should they be able to complete the task, a test on learners below their level would be worthwhile.

This process confirmed that changes were desirable regarding the test design as well as both the functions and lexical bundles. The pilot test participants said they thought it would be better if the items were split into smaller sets, which was then done to form four sets of six items. As for the functions and lexical bundles, with the basis of discourse at all levels involving the basic functions performed by the words “and,” “but,” “so,” and “because,” something voiced during discussion with the panel of experts, it was agreed

that the four functions would be adding information, compare and contrast, cause, and effect. These functions and the selected items are shown in [Table 1](#).

Table 1: Lexical bundles of the experiment by function

Function	Lexical bundles
Adding information	<i>and furthermore, in addition , as well as, and what is more, not to mention, further to</i>
Compare/contrast	<i>is similar to, in contrast to, in the same way, significantly different from, is comparable to, on the other hand</i>
Causality cause/effect	<i>which leads to, may cause, which causes, as a consequence, have contributed to, may affect</i>
Causality - effect/cause	<i>stems from, is linked to, caused by, due to, owing to, as a result of</i>

Materials

The style of DDL used here was hands off concordancing. That is to say, the participants did not interact with a corpus themselves, but were exposed to corpus extracts selected by the researchers. The corpus from which the extracts were taken is the BAWE corpus, a 6.9 million word collection of written academic texts in British English. The number of citations given per item was 12 and a sample can be seen in [Appendix D](#). It is important to note that the citations were given in full sentences, not KWIC format concordances. Although there are advantages to KWIC format concordances, such as the increased noticing of surrounding words and the potential for ‘vertical or paradigmatic reading’ (Ballance, 2016, p. 1208), full sentences were chosen to allow more familiarity in early stage DDL. The citations in the concordance sheets were selected randomly and then checked for any inappropriate content in line with ethical considerations.

Implementation

Stage One: Experimental Groups

The pre and post-tests (see [Appendix C](#) for a sample) were carried out in test conditions. Participants were informed that they would complete four tasks as part of a 20-minute test. It was stated that they should answer every question, even if they were unsure of the answer. The participants were also told not to worry about the difficulty of the task and that they should not try to understand every unknown word. Although the post-test was conducted in identical circumstances, the participants were of course fully aware of what was expected of them having completed a pre-test and training sessions.

Each session followed an “observe, hypothesize, experiment” (OHE) procedure ([Appendix E](#)) based on the suggestions put forward by Lewis (1993). This procedure is in line with the “identify, classify, generalize” procedure of Johns (1991) and the “look, familiarize, practice, create” procedure described by Thurston and Candlin (1998). As can be seen by the prompts in the plan used ([Appendix E](#)), the aim was to encourage learner-generated noticing and hypothesizing from the outset. Oral instructions were used to achieve this ([Appendix E](#)).

Four items were included in the first training session, with instructions given prior to the first sheet of concordances being distributed. Participants were asked to circle the lexical bundle in each sentence (which was already highlighted in bold). Participants were then given time to observe the concordances. This was followed by a period in which the group voiced their observations and were encouraged to hypothesize. These observations and hypotheses were discussed and debated by the group. The second sheet was then handed out and participants repeated the process. Following the second item, participants were asked to try and form a simple sentence (orally) to show how the item works. Participants were then given the opportunity to discuss their sentences in terms of accuracy of meaning and use, with a particular emphasis

on the item meaning. The first item was then subjected to the same treatment and participants had now understood the process of OHE that was desired. The process was then followed for items three and four. The session ended with participants forming example sentences (orally) of all four items. Four items were also provided in session two, with the now-established process being followed for all items.

In the subsequent four sessions, participants of both groups experimented by making simple sentences for problematic items (such as ‘I like chocolate *in the same way* I like pizza’) and checking with other participants. The instructor also confirmed some of these examples as being correct in meaning or use, when participants expressly requested instructor feedback.

Stage Two: Control Group

The second stage was carried out one year after the first stage in order to include a group with a comparable academic and English background to group A. The pre-test and post-test were conducted in identical conditions to those of the first stage and the same items were involved in each session.

The method of learning, in order to give a measure of conventional learning, was in line with how modern course books present items. The first three sessions followed a deductive “present, practice, produce” procedure, moving from controlled to semi-controlled tasks. The students were allowed and encouraged to use reference materials when they felt it necessary. However, it should be noted that learners do not instinctively turn to dictionaries in the setting of this study (this applies to all participants) due to learner training away from a reliance on total comprehension at all times and towards facilitating general comprehension. This is in contrast to the type of learning described as typical in France by Boulton (2010).

In the second three sessions, participants were given context (teacher-provided, spoken example sentences of the items and an academic reading task) and encouraged to inductively work towards understanding (in both cases these activities were supported by matching tasks). Meaning was clarified and again reference materials were consulted where participants so desired. Participants then produced examples and practiced substituting one-word connectors with the academic lexical bundles in focus here. In all six conventional sessions, there was a focus on meaning, form (including pronunciation and spelling), and use, and learners were allowed and encouraged to ask questions of the teacher at any time.

In order to carry out statistical analysis of the data, the tests were marked. Each question was marked as correct if the original bundle was used. It should be noted that there was the possibility of a few of the items being somewhat interchangeable within the selected sentences, but it was decided that “correct” would mean completing the original sentence here. The data sets were then put through a series of tests, the results of which are reported below.

Results

Paper-Based, In-Class DDL with Academic Lexical Bundles below the C1 Level

Paired-samples *t*-tests were conducted for both experimental groups as a measurement of performance between the pre-test and post-test. Further to these, where significant results were found, effect sizes were calculated using the relevant Cohen’s *d* formula. These results are shown in Table 2 (group A) and Table 3 (group B) respectively.

RQ1. Can B2 level learners learn academic lexical bundles to a significant degree using paper-based in-class DDL?

Table 2: Group A pre-test & post-test results

	<i>N</i>	<i>M</i>	<i>M (%)</i>	σ	<i>t</i>	<i>p</i>	$\bar{x}_2 - \bar{x}_1$	Cohen’s <i>d</i>
Pre-test	12	6.92	28.82	11.16	4.0524	0.0019	10.07	1.05
Post-test	12	9.33	38.89	7.61				

N: Number of participants, *M*: Raw mean, *M*(%) Mean as %, σ : Standard deviation, *t*: *t*-test value, *p*: level of significance, $\bar{x}_2 - \bar{x}_1$ = the difference in mean values

The mean of the pre-test scores of group A is 28.82% and the post-test mean is 38.89%. This difference, a 34.94% increase in performance, is statistically significant at the alpha $p < .01$ level. The effect size calculation gives a Cohen's *d* figure of 1.05.

The results represent an increase that is significant to a strong level (.01), a confidence level of 99.8%, coupled with an effect size that is above the 'large' effect size level of 0.8 put forward by Cohen (1988), and that of 1.00 suggested by Plonsky and Oswald (2014, p. 12) for second language research studies, indicating that the answer to the first research question is affirmative; B2 level learners can learn this item type to a significant degree through paper-based, in-class DDL.

RQ2: Can A2-B1 level learners learn academic lexical bundles to a significant degree using paper-based in-class DDL?

Table 3: Group B pre-test & post-test

	<i>N</i>	<i>M</i>	<i>M</i> (%)	σ	<i>t</i>	<i>p</i>	$\bar{x}_2 - \bar{x}_1$
Pre-test	12	6.0	25.0	6.15	-0.1205	.9062	-0.35
Post-test	12	5.92	24.65	7.84			

N: Number of participants, *M*: Raw mean, *M*(%): Mean as %, σ : Standard deviation, *t*: *t*-test value, *p*: level of significance at $p < .05$, $\bar{x}_2 - \bar{x}_1$ = the difference in mean values

The group B pre-test mean score is 25% and the post-test mean is 24.65%. The difference, a 1.4% decrease in performance, is not significant at the alpha $p > .05$.

This result suggests that A2-B1 level learners cannot learn academic lexical bundles with paper-based in-class DDL. Consequently, it can be seen that the combination of findings related to RQ1 and RQ2 imply that B2 is a cut-off point for this item type when combined with this form of learning.

Conventional Learning of Academic Lexical Bundles at the B2 Level

A paired-sample *t*-test for the pre-test and post-test data of group C is detailed in Table 4 with a Cohen's *d* calculation of the effect size in order to address the third research question.

RQ3: Can B2 level learners learn academic lexical bundles to a significant degree using conventional methods?

Table 4: Group C pre-test & post-test

	<i>N</i>	<i>M</i>	<i>M</i> (%)	σ	<i>t</i>	<i>p</i>	$\bar{x}_2 - \bar{x}_1$	Cohen's <i>d</i>
Pre-test	17	7.29	30.39	9.17	1.1863	0.2528	3.17	0.29
Post-test	17	8.06	33.58	12.45				

N: Number of participants, *M*: Raw mean, *M*(%): Mean as %, σ : Standard deviation, *t*: *t*-test value, *p*: level of significance at $p < .05$, $\bar{x}_2 - \bar{x}_1$ = the difference in mean values

The group C pre-test mean score is 30.39% and the post-test mean is 33.58%. The difference is not significant at the alpha $p > .05$. The effect size calculation returns a Cohen's *d* figure of 0.29, slightly above Cohen's (1988) small effect threshold of 0.20, but below the small effect size of 0.40 recommended for second language research by Plonsky and Oswald (2014, p. 12).

The results suggest that B2 level learners cannot learn academic lexical bundles to a significant degree through conventional methods.

Comparing DDL & Conventional Learning

Finally, an independent-samples *t*-test to compare the differences in means of the pre-test and post-test scores of groups A and C together with a Cohen's *d* calculation of effect size are detailed in Table 5 in order to answer the final research question.

RQ4: Is paper-based, in-class DDL more effective than conventional learning with academic lexical bundles at the B2 level?

Table 5: Comparison of Group A & Group C

	<i>N</i>	<i>X</i>	\bar{x}	σ	<i>t</i>	<i>p</i>	$\bar{x}_2 - \bar{x}_1$	Sig	Cohen's <i>d</i>
Group A	12	2.41	10.07	8.61	1.7993	.0832	6.90	> .05	0.70
Group C	17	0.77	3.17	11.08					

N: Number of participants, *X*: difference in raw mean pre/post test, \bar{x} : Difference in mean % pre/post-test, σ : Standard deviation, *t*: *t*-test value, *p*: level of significance at $p > .05$, $\bar{x}_2 - \bar{x}_1$ = the difference between group A and group C

Table 5 shows the results of an independent samples *t*-test of the differences in performance of participants in groups A (experimental group) and C (control group) from pre-test to post-test. The difference of group A is 10.07, a 34.94% improvement, while group C has a 3.17 difference, a 10.5% increase. Although both groups show an increase, with that of group A greater than that of group C, the result is not significant at the .05 level.

In terms of effect size, the calculation of Cohen's *d* returns a figure of 0.70. This figure is between the 0.50 medium effect and 0.80 large effect that Cohen (1988) tentatively identified, with 0.70 being closer to the large effect. However, as Plonsky and Oswald (2014 p. 12) point out, effect size interpretation varies according to discipline, and according to their recommendation, 0.40 indicates a small effect size, 0.70 a medium effect size, and 1.00 a large effect size for L2 research. According to this interpretation, there is a firm medium effect size. In sum, it can be said that the improvement in performance of group A was better than that of group C, although a larger statistical difference would have allowed more confident generalization.

Discussion

This study set out to investigate the learning of academic lexical bundles below the C1 level through paper-based, in-class DDL. The results suggest that this is possible at the B2 level, but not so below that point. A comparison then found that this type of learning is more efficient than conventional learning for academic lexical bundles in keeping with the many studies that have found DDL to be more effective than conventional learning, as reported by Boulton and Cobb (2017: 34). However, the scores achieved in all conditions were low, and the effect sizes varied. These two areas are therefore in need of further scrutiny.

Boulton and Cobb's (2017) meta-analysis of DDL, which covered two and a half decades up to 2014, provides a great deal of insight. That study consists of a 64-study, 88-sample meta-analysis of DDL research, with effect size calculations made for many combinations of variables. A sub-group of 13 studies with a sample size less than 20 and a pre/post-test design returned a mean effect size of $d = 1.17$, which is somewhat stronger than the effect size found for group A in this study ($d = 1.05$). Also, 13 studies with a pre/post-test design were paper-based and had a mean effect size of $d = 1.06$, very close to that of group A ($d = 1.05$). Therefore, it can be seen that the effect size for group A was within expectations and is particularly consistent with previous findings from studies with similar designs. Furthermore, the effect size calculation of 0.70 for the comparison of group A and group C is larger than that of the mean effect size of another subset of fifteen studies which were paper-based and involved control and experimental groups ($d = 0.52$) reported in that meta-analysis.

However, when compared to the overall effect sizes for pre/post-test designs of $d = 1.50$, the effect size of

group A ($d = 1.05$) is considerably lower. The 0.70 effect size for the difference between group A and group C is also lower than the overall experimental/controlled design effect size of 0.95 calculated by Boulton and Cobb (2017). With the combination of results and effect sizes of this study being somewhat limited, a closer look at the scores may shed more light on the situation.

The scores were low in both the pre-tests and the post-tests of all three groups. The pre-test mean scores of 28.82% (group A), 25% (group B), and 30.39% (group C) represent an overall correct response rate of 28.35%. The post-test mean scores of 38.89%, 24.65%, and 33.58% return a mean score rate of 32.52%. This is in keeping with past studies. However, as mentioned earlier, the group B participants' scores may have been further subdued by the difficulty of the test used. An alternative test design may have yielded higher scores all round, although that may have also impacted upon the validity of the test were it to be too easy or not include sentences similar to those used in the training sessions.

In terms of the content of the concordances used in both the test and the DDL sessions, it could be said that the selected sentences were heavy going. One solution could have been to incorporate the use of a scaled down level of English, such as a graded reader corpus. On the other hand, the advantage of using a graded reader corpus for DDL is not proven (see results reported in Hadley & Charles, 2017) and the likelihood of finding enough academic lexical bundles in such materials is low. Another possibility would naturally be the careful selection of sentences by reading difficulty tests such as Gunning's Fog Index (discussed in Gunning, 1969), but this would inevitably cause difficulties in practice due to time constraints. Nonetheless, the content could have affected performance, although low scores are not uncommon.

Several studies have reported low scores in lower proficiency DDL. Boulton (2008a, 2009, 2010) has reported several studies of DDL with lower proficiency learners, each demonstrating the place of DDL as superior to conventional learning. However, those studies also carry a further commonality in having low scores. The earliest study involved two items and so 50% would be expected by chance only, yet the overall score returned was 57% (Boulton, 2008a, pp. 9–10). The following study involved linking adverbials and produced what the author described as very low scores with a mean of 13.92% (Boulton, 2009, p. 9). The next study returned overall mean scores of 48.53% in the pre-test and 57.97% in the post-test (Boulton, 2010, pp. 11). In contrast, Chujo et al. (2012) reported scores which were considerably higher with pre-test mean scores ranging from 43.6% to 65.2% and post-test means ranging from 64.6% to 88.5%. Of course, the scores that participants achieve depends on how they are tested and by no means negates the beneficial findings of these studies. However, teachers and learners alike prefer the appeal of higher scores to lower ones and this could be detrimental to encouraging the uptake of DDL.

Another likely factor to the low scores is the language item type. Previous works have noted the difficulty of academic collocations (such as Cortes, 2007 and Ellis et al., 2008) as they require a depth of knowledge (Cooper, 2017) that puts a strain on the learner and academic vocabulary is already a more challenging prospect than general vocabulary.

On a more positive note, several observations can be made about the treatment in terms of benefits to the participants. Participants were subjected to intentional vocabulary learning—the kind of word-focused learning noted by Laufer (2017)—which is one side of the current debate in vocabulary learning. Intentional vocabulary learning means that language educators are able to find and provide concrete examples that Cooper (2017) suggests are needed to target learner idiosyncrasies. The participants were also able to benefit from encountering the bundles more than the 12 times recommended by Laufer (2017) and the 15 times needed for collocation learning stated by (Webb et al., 2013).

When one considers the nature of the DDL that took place, participants showed signs of the type of detective work detailed by Johns (1991). Participants of both groups quickly noticed that the items performed functions and were therefore able to experiment relatively quickly, connecting items together and experimenting with replacing some items with others, often those which performed the same function. An example of this was a Group B participant, who gave an example sentence comparing the weather in his country (Libya) with that of Cyprus. He initially used

the bundle “is comparable to,” but wanted to express that Cyprus was in fact hotter than Libya. The participant then asked if he could say “Libya weather it is good, on the other hand Cyprus is hot.” Another strategy participants adopted involved focusing on key semantic clues; the lexical bundles that contained key content words signifying the core meaning seemed easier for learners to recall. Items such as “in the same way” and “is comparable to” seemed easier than items such as “due to” or “owing to,” despite the number of words. However, this was not borne out in the results of the post-test, and may have been a false strategy.

Conclusion

Several limitations of this study require mention. First and foremost, the sample of the study can be criticized in a few ways. The size of the groups ($n = 12$, $n = 12$, and $n = 17$) could clearly have been larger and more diverse culturally and linguistically to allow for a more generalizable result. The pre/post-test instrument is a limitation in terms of test difficulty and this difficulty could, of course, have affected the lower level learners to a greater extent than other participants. It also lacks productive use of language, although this would have heightened the effect of test difficulty on the lower level learners. Finally, the acceptance of only the original bundles as being correct may have allowed well-formed sentences to be deemed wrong. Further research could address these limitations with a larger, more diverse sample being tested in both receptive and productive ways through a test that is of a more suitable degree of difficulty.

There are some other possible research implications that arise from this study. It may be useful to conduct a hands-on version of this study. Based on Boulton and Cobb’s (2017) meta-analysis findings, an increase in effect size would be expected. This might be particularly useful with a context-heavy corpus such as the TED Corpus of Spoken English compiled by Hasebe (2015).

This study has demonstrated that learners at the B2 level, but no lower, can learn academic lexical bundles through paper-based, in-class DDL more effectively than through conventional methods. As the demand for EAP at lower proficiency levels continues to rise, can word-focused vocabulary learning of academic lexical bundles therefore be incorporated into the B2 level curriculum through means of paper-based in-class data-driven learning? If stakeholders start listening to effect sizes, perhaps. On the other hand, should they focus on scores, as is the nature of teachers and learners alike, perhaps not.

References

- Ackermann, K., & Chen, Y. H. (2013). Developing the Academic Collocation List (ACL)—A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes*, 12(4), 235–247.
- Al Badi, I. A. H. (2015). Academic writing difficulties of ESL learners. In *The 2015 WEI international academic conference proceedings* (Vol. 1, No. 1, pp. 65-78).
- Ballance, O. J. (2016). Analysing concordancing: A simple or multifaceted construct? *Computer Assisted Language Learning*, 29(7), 1205–1219.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371–405.
- Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534–572.
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(01), 37–54.
- Boulton, A. (2008a). Looking (for) empirical evidence of data-driven learning at lower levels. *Corpus Linguistics, Computer Tools, and Applications: State of the Art.*, 581–598.

- Boulton, A. (2008b). DDL: Reaching the parts other teaching can't reach? In Frankenberg-Garcia, A. (ed.), *Proceedings of the 8th Teaching and Language Corpora Conference*. Lisbon: Associação de Estudos e de Investigação Científica do ISLA-Lisboa, 38–44.
- Boulton, A. (2007). But where's the proof? The need for empirical evidence for data-driven learning. In M. Edwardes (Ed.), *Proceedings of the BAAL annual conference* (pp. 13–16). London: Scitsiugnil Press
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393.
- Brown, M. H. (2017). Using the Sentence Corpus of Remedial English to introduce Data-Driven Learning tasks. *Kanda Academic Review*, 1(1), 1286. Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, 5(5), 31–64.
- Choo, L. B., Lin, D. T. A., & Pandian, A. (2012). Language learning approaches: A review of research on explicit and implicit learning in vocabulary acquisition. *Procedia-Social and Behavioral Sciences*, 55, 852–860.
- Chujo, K., Anthony, L., Oghigian, K., & Uchibori, A. (2012). Paper-based, computer-based, and combined data-driven learning using a web-based concordancer. *Language Education in Asia*, 3(2), 132–145.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Cooper, T. (2017). Students' use of academic vocabulary in comparison to that of published writers: A corpus-driven analysis. *Stellenbosch Papers in Linguistics*, 47, 133–152.
- Cortes, V. (2007). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17(4), 391–406.
- Coxhead, A. (2011). The Academic Word List 10 years on: Research and teaching implications. *TESOL Quarterly*, 45(2), 355–362.
- Csomay, E., & Prades, A. (2018). Academic vocabulary in ESL student papers: A corpus-based study. *Journal of English for Academic Purposes*, 33, 100–118.
- DeCarrico, J., & Nattinger, J. R. (1988). Lexical phrases for the comprehension of academic lectures. *English for Specific Purposes*, 7(2), 91–102.
- Ellis, N. C., Simpson-vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *Tesol Quarterly*, 42(3), 375–396.
- Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes*, 24(3), 321–332.
- Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, 6(2), 3–13.
- Hadley, G., & Charles, M. (2017). Enhancing extensive reading with data-driven learning. *Language Learning & Technology*, 21(3), 131–152. Retrieved from: <https://www.lltjournal.org/item/3014>
- Hasebe, Y. (2015). Design and implementation of an online corpus of presentation transcripts of TED talks. *Procedia-Social and Behavioral Sciences*, 198, 174–182.
- Hirschel, R., & Fritz, E. (2013). Learning vocabulary: CALL program versus vocabulary notebook. *System*, 41(3), 639–653.

- Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge, UK: Cambridge University Press
- Hyland, K., & Jiang, F. K. (2018). Academic lexical bundles. *International Journal of Corpus Linguistics*, 23(4), 383–407.
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. In T. Johns, P. King (Eds.), *Classroom concordancing*. English Language Research Journal Volume 4, (pp. 1-16), Birmingham University Press, Birmingham, UK.
- Laufer, B. (2017). From word parts to full texts: Searching for effective methods of vocabulary learning. *Language Teaching Research*, 21(1), 5–11.
- Laufer, B., & Rozovski-Roitblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, 19(6), 687–711.
- Lewis, M. (1993). *The Lexical Approach*. Hove: Language Teaching Publications.
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18(2), 85–102.
- Morley, J. (2017). *The academic phrasebank: An academic writing resource for students and researchers*. The University of Manchester: Manchester, UK.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge University Press, Cambridge
- Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes*, 21, 60–71.
- Pimsleur, P. (1967). A memory schedule. *The Modern Language Journal*, 51(2), 73–75.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912.
- Rosszell, H. R. (2007). *Extensive reading and intensive vocabulary study in a Japanese university*. Temple University.
- Thurston, J., & Candlin, C. N. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes*, 17(3), 267–280.
- Vyatkina, N. (2016). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL*, 28(02), 207–226.
- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63(1), 91–120.

Acknowledgements

The data in this study come from the British Academic Written English (BAWE) corpus, which was developed at the Universities of Warwick, Reading and Oxford Brookes under the directorship of Hilary Nesi and Sheena Gardner (formerly of the Centre for Applied Linguistics [previously called CELTE], Warwick), Paul Thompson (Department of Applied Linguistics, Reading) and Paul Wickens (Westminster Institute of Education, Oxford Brookes), with funding from the ESRC (RES-000-23-0800).

Appendices

Appendix A:

Selection process in stage one: Result samples from two of the 17 categories on **Morley's Academic Phrasebank (2017)**.

Describing Quantities - Five most common lexical bundles by hit on the BAWE corpus

Item	Total Hits
5 Items Selected (Below)	38
<i>the highest proportion</i>	3
dropped from * to	4
<i>half of those</i>	2
<i>lower than average</i>	1
in the range of	28

Explaining Causality - Five most common items by hit on the BAWE corpus

Item	Total Hits
5 Items Selected (Below)	1570
as a result	1287
as a consequence of	83
Because of this	103
may have contributed to	14
owing to	83

Appendix B:

Selection process in stage two sample.

Giving Examples

Item	Total Hits	Frequency Per Million
A/An * example of * is	20	2.4
For example	3348	401.6
This is exemplified	12	1.44
Another example of	66	7.92
Is further exemplified	1	0.12
An example of this is	55	6.6
Has been exemplified in	0	0
Seen in the case of	6	0.72
Is a good illustration of	2	0.24
Illustrates this point	3	0.36
Can be illustrated	26	3.12
By way of illustration	5	0.6
Illustrate that	134	16.08
Such as	6926	830.8
Supports the view that	23	2.76

This * demonstrates 23 2.76

Appendix C:

Sample from the pre/post-test

a) stems from	b) is similar to	c) and furthermore	e older
d) is linked to	e) may cause	f) in contrast to	

- This concern with the implications mimesis has for the search for truth _____ Plato's hierarchical conception of society.
- Never apply ice directly onto the skin, as this _____ tissue damage.
- What should be noted, however, is that experiences were diverse, with differing experiences between, as well as within states, _____ between different social groups.
- The accessibility awareness at the Hippodrome _____ that at the Playhouse.
- It is though very unlikely that it _____ human selection pressures, as human-to-human transmission does not occur.

Appendix D:

Concordance for 'which leads to'.

Possible hypotheses for this concordance include: The item is used to connect ideas, the first idea causes the second idea (desired hypothesis), the ideas are/are not opposed, the second idea adds detail, the second idea changes the first and so on.

An advantage of using qualitative methods is that we can study behaviour in the context of which it occurs which leads to a greater understanding to the meaning behind the behaviour (Billig, 1997).

There is a reduction in an individual's self-restraint and normative regulation of behavior which leads to a newfound obedience to any form of authority.

The trophoblasts invade the tubal wall, weakening it which leads to bleeding.

However, we can never observe all instances of a phenomenon, which leads to infinity.

He employs such verbs as 'punir...réfuser...détourner...abuser' to describe her behaviour and his sentiments include 'tendre...l'amour...l'amitié'. This is an effective technique and Madame de Tourvel genuinely wrestles with her conscience and eventually capitulates which leads to her perfidy.

Thus it is not surprising that activation of the phototransduction pathway functions in a similar way to inhibition of the olfactory transduction pathway, which leads to rapid desensitisation of olfactory neurons in case of high intensity or prolonged presence of the odourant.

However, this gradient only holds true for the continental crust, a vital error which leads to an overestimation of the temperature at the centre by a magnificent factor of forty.

If the economy is actually facing deflation and is in a liquidity trap, output below the natural level leads to more deflation over time, which leads to a further increase in the real interest rate and leads to a further shift of the IS curve to the left.

This shift leads to a further decrease in output, which leads to more deflation, as shown below.

This has been proven by blocking adenosine receptors which leads to the increase in the rate of ethanol release to the water (Nilsson, 2001).

There are a number of methodological flaws when examining "Islam" which leads to limitations in its findings and conclusions.

They find that 60% of file sharers use downloading to sample music, which leads to them buying at least one CD.

Appendix E:

OHE Based Plan

Warm-up: Greet learners, introduce the task of the lesson – show sheets that will be studied. Use instructions as a guide (below).

Observe: Give learners first sheet, encourage observations.

Hypothesize: Allow students to hypothesize on their own first then share their ideas. Encourage debate and rationalization.

Experiment: Encourage learners to produce utterances that test their hypotheses. Allow learner to learner correction and commenting. Only clarify when they absolutely need.

Repeat process for each item.

Instructions Guide

(given orally, teacher monitoring and repeating instructions to individuals to check comprehension and compliance)

Before they look:

I am going to give you some lists.

The lists have examples of some of the phrases from the original test.

The language in the examples is very high level academic English.

Please don't worry about words you don't know or examples you don't understand.

Please don't use dictionaries or the internet to look up the phrases or words you don't understand.

After giving:

Look at the examples.

Think of as many different things that you notice.

Are the examples similar?

Do similar words appear?

Do similar parts of speech appear (nouns, verbs, prepositions, etc.)?

How are the parts that go before the phrase similar?

How are the parts that go after the phrase different?

If you wanted to change the phrase for a different phrase or word, what would you change it to?

Do you think the phrase shares the same/a similar meaning to something you already know?

How could you describe the phrase to help a friend understand it?

About the Authors

Keith John Lay is a teacher educator at Cyprus International University in Northern Cyprus. He gained his BA in Criminology from the Department of Sociology, School of Social Sciences, Nottingham Trent University, in 2000. He earned his CELTA at Gloucester College, UK, in 2010, marking the start of his teaching English career. Earning his MA in English Language Teaching in 2016 and being awarded his PhD in 2020, both at Cyprus International University, his areas of interest include data-driven learning, corpus analysis, language assessment, higher education, and nonverbal communication.

Email: klay@ciu.edu.tr

Dr. Mehmet Ali Yavuz is a professor of linguistics and language teaching at Cyprus International University in Northern Cyprus. He earned his degrees (BA, MA, and PhD) in English Language Teaching. He taught at Dokuz Eylül University in Turkey. His areas of interest include error analysis, contrastive analysis, syntactic analysis, and semantic and pragmatic issues in human communication.

Email: mayavuz@ciu.edu.tr