How Does Data Driven Learning Affect the Production of Multi-Word Sequences in EAP Students’ Academic Writing?

Melissa Larsen-Walker 1

1 University of South Florida, Tampa, FL 33620, USA
ml Larsenw@mail.usf.edu

Abstract. Multi-word sequences (MWSs) have been found to occur with high frequency in academic writing [2, 25]. MWSs are recurrent expressions, which a writer retrieves from his/her long-term memory in order to construct utterances. In written discourse, such MWSs serve to refer the reader to previous research, organize the sections of texts and discourse within them and position the writer as knowledgeable [2, 3]. Previous research suggests that L2 writers frequently misuse these forms, resulting in disfluent written discourse [8, 20, 23]. Nevertheless, Hyland [14] suggests that use of appropriate and sophisticated MWSs helps to establish the writer as a member of an academic discourse community. The current, corpus-based and quasi-experimental study investigates the effectiveness of using Data Driven Learning (DDL) in conjunction with teaching MWSs. Key MWSs have been selected from Simpson-Vlach and Ellis’ [25] Academic Formulas List (AFL), specifically from the Referential Function and two sub-categories of the Stance Function, Hedges and Epistemic Stance. The researcher used an objective pretest-posttest to ascertain how DDL affects students’ receptive knowledge of AFL-MWSs and used the first and final drafts of an argumentative essay to assess students’ ability to produce them. A statistically significant difference between pretest and posttest scores for the treatment group supports the assertion that DDL positively impacts students’ receptive knowledge of AFL-MWSs. Discussion includes comparison between students’ self-generated inductions regarding each AFL-MWSs and how they used them within their essays.

Keywords: Data Driven Learning, Multiword Sequences, Academic Formulae.

1 Introduction

L2 writers, endeavoring to craft appropriate essays and research papers, must conform to the academic expectations of colleges and universities in the U.S. Despite many strengths, one aspect of L2 undergraduate students’ writing that appears to be a weakness is their use of multi-word sequences (MWSs). Research suggests that in many cases, they use more MWSs than native speakers, whether novice or expert [13, 19-20, 22, 24]. However, they tend to use MWSs that are considered to be too informal
and conversational for the essay task. Other aspects of writing, such as lucid content, vocabulary choice and organization may be relatively more important to the overall effectiveness of a student’s written composition. Nevertheless, specific academically-appropriate MWSs influence the quality of the writing in terms of cohesiveness and positioning of the writer as an expert on the topic [14]. Such MWSs frequently occur in the writing of published academic authors [3, 21-22, 27]. Use of these MWSs has also been associated with higher grades on the Common European Framework of Reference [18]. In spite of numerous studies that document the importance of MWSs, no consensus exists as to how best to teach these formulas.

The current research investigates the possible effectiveness of using Data Driven Learning as a means to teach MWSs, specifically MWSs selected from Simpson-Vlach and Ellis’ Academic Formulas List (AFL) which is described below. Data Driven Learning (DDL) is an inductive approach wherein learners search a corpus or corpora under instructor guidance, in hopes of finding a pattern in the data. It has been compared to discovery learning in that the instructor guides students and acquaints them with how to use the material, i.e. learner website and corpus, but refrains from pre-teaching the rule or generalization regarding the target form [15-16]. Based on previous research, the researcher infers that DDL could be a viable alternative to traditional direct instruction. Despite ample documentation for the importance of using appropriate MWSs to academic writing, most studies on MWSs do not contain any instructional component [2, 7, 11, 22, 24]. Even rarer are classroom based studies on MWSs that include DDL. The current research aims to address this gap.

Awareness of the tendency of L2 students to select inappropriate MWSs motivated Simpson-Vlach and Ellis [25] to compile a list of MWSs or academic formulas that frequently occur in academic writing and would be valuable pedagogically. The corpora used were: Hyland corpus of research articles; (spoken) academic portion of Michigan Corpus of Academic Spoken English (MICASE), British National Corpus (BNC) spoken academic corpus; and Lee’s genre category of the BNC. For comparison, they consulted the FROWN corpus (informal American English speech). Simpson-Vlach and Ellis used multiple measures, qualitative and quantitative methods, combining information from log likelihood, mutual information score, and formula teaching worth (FTW). First, they measured simple frequency in the target and comparison corpora. They were concerned about the fact that lists of MWSs which have been compiled based solely on frequency, such as that composed by Biber, Johansson, Leech, Conrad and Finegan [4], tend to contain MWSs which from a pedagogical standpoint are unimportant [19-20, 26]. For example, one of the was found to be the most frequently occurring lexical bundle in the initial search, yet these individual words were extremely common throughout both the expert and comparison corpora. In order to correct for this tendency, without going to the other extreme of selecting MWS solely on the researchers’ intuitions, Simpson-Vlach and Ellis used an alternative way to analyze MWSs, the mutual information (MI) statistic. The MI statistic, which originated in the field of information science, rates the likelihood that two (or more) words did not occur together by chance. As such, relatively low frequency collocations such as blue moon have a high MI score.
The current study applies the formulas from the AFL, i.e. AFL-MWSs to the community college, EAP setting, testing whether DDL can be effective in teaching these AFL-MWSs to L2 English under-graduates. Specifically, this study explores the following questions:

**Research Question #1:** To what extent does Data Driven Learning affect receptive knowledge of academically appropriate MWSs, operationalized as key formulas from Simpson-Vlach and Ellis’ *Academic Formulas List* (AFL). Receptive knowledge would be measured by an objective test of MWSs (i.e. academic formulas).

**Research Question #2:** To what extent does DDL affect productive knowledge of academically appropriate MWSs? Productive knowledge would be measured by accurate usage of academic formulas within the argumentative/persuasive essays.

## 2 Literature Review

### 2.1 MWS Studies Including an Instructional Intervention

Cortes [8] endeavored to teach MWSs to L1 undergraduates in a history course. Direct instruction including presentation of the formulas was followed by practice activities throughout her five visits to the classroom. Results confirm the null hypothesis; i.e. the students did not produce more MWSs in their assigned writing after the Cortes’ instruction. The small number of participants (n=12) may have influenced the results.

### 2.2 Review of Empirical Studies on DDL

**Linking adverbial studies.** Numerous studies have investigated the effectiveness of DDL in teaching linking adverbials (LAs), a construct related to MWS [5, 9-10, 12]. Linking adverbials connect sections of text, either within a sentence or across longer sections of text. These transitional devices may be single word (however) or multi-word (on the other hand). Boulton [5] explored paper-based DDL as a means to elicit appropriate use of LAs among L1 French undergraduates, writing in English. Results suggest that DDL was effective despite the fact that the participants were intermediate level. Garner [12] used both the COCA and the MICUSP to investigate DDL and LAs. His quasi-experimental study aimed to elicit use of academically appropriate MWSs and to elicit correct use of LAs. The measure of the treatment’s effectiveness were the pretest and posttest essays written by the participants. Garner found a statistical difference between correct use of LAs in pretest and posttest essays, from which he infers the effectiveness of the treatment. Cotos [9] aimed to find out whether the inclusion of a learner corpus would impact the effectiveness of DDL in teaching LAs to L2 English undergraduates. She instructed two classrooms via DDL, but only the LLD group used a learner corpus of essays written by unidentified course students. All participants used a native speaker corpus comprised of articles from his/her discipline. Her instruments were both an objective LAs test and essays written pre- and
Post-treatment. The NSC and LDD groups showed a statistically significant difference between the pre- and post-test scores on LAs recognition. As such, Cotos claims the DDL intervention including a learner corpus was effective, and the effects of instruction persisted beyond the course limits.

Tangible learning gains were made in studies wherein DDL was used to teach vocabulary and/or collocation of prepositions. Koosha and Jafarpur [17] study aimed to investigate the effectiveness of using DDL to teach collocation of prepositions to L1 Farsi undergraduates. Their quasi-experimental included 200 participants at two university campuses in Iran. Participants in the experimental group used the Brown corpus under instructor guidance, while the control group used the Longman Dictionary online. Koosha and Jafarpur used an objective test of target collocations for their pretest-posttest. Results suggest that DDL was effective in eliciting knowledge of correct collocation of prepositions, based on both a statistical difference between pretest and posttest scores for the experimental DDL group and a difference between the experimental and control groups. Using a similar research design, Celik [6] investigated the effectiveness of using DDL to teach collocation of prepositions to L2 medical students in Ankara, Turkey. His results were less impressive than Koosha and Jafarpur’s as far as the difference between control and treatment groups on the immediate posttest. Nevertheless, on the delayed posttest there was a statistical difference between the mean scores for the DDL and comparison groups, with the former achieving the higher mean score. These vocabulary studies appeared to be effective although the outcome measure was an objective test rather than a measurement of students’ use of MWSs within pretest and posttest writing assignments.

In a study by Park [21], the researcher explored the effects of corpus consultation on errors in essays and a microanalysis of the students’ revision process. Analysis of results shows corpus consultation resulted in a correct revision in 56% of cases.

3 Methods

The design of the study is quasi-experimental and will include four classrooms across the course of three semesters. This short paper addresses the first phase of data collection, which included on treatment group only. The instructional sequence lasted two weeks, including the time the administration of the objective pretest and posttests. All instructional materials were included as part of an instructor designed website. http://mws-afl-eap-blueлотусfeet.sitey.me/ As part of the process of guiding students to ascertain a generalization for how the phrase is used, the instructor modeled inductive reasoning, using the phrase, claim that. After this modeling, the students continued to work on induction independently, completing the second task associated with the website, which required them to induce a rule for how each of the key AFL-MWSs (shown in Table 1) should be used. To address RQ#1, the researcher designed and administered an objective test, described below, to measure knowledge of the target MWSs. To address RQ#2, the researcher analyzed the first and second drafts of
an argumentative essay, with a researched component, the first drafts of which each student wrote prior to treatment. Participants revised and re-submitted them as final drafts after the DDL treatment.

### 3.1 Target Formulas and Instructional Sequence

The target formulas in this study are MWSs from the AFL, the most frequently occurring from within the Referential functional category, such as *the fact that*. The rationale for this is the fact that previous studies have shown that academic writing by experts contains relatively more referential MWSs [2, 7, 19]. Along with referential, the proposed study will include two of the subcategories of stance, hedges, such as *are likely to be* and epistemic stance, such as *we assume that*. Hedges qualify or mitigate the writer’s assertion, leaving them open to the reader’s attempts to falsify such claims [25]. Use of MWSs from AFL contributes to the coherence of the text, positioning the writer as one who builds upon the scholarly research of others while qualifying the claims for his/her own possible contribution. For that reason, the researcher will provide instruction via the website in the Hedging subcategory of the Stance function prior to students’ independent searches in the corpus. The table below displays the specific MWSs for which the participants will seek, classified within stance or referential functional categories.

**Table 1.** Target MWSs from the Academic Formulas List, Including Functional Category.

<table>
<thead>
<tr>
<th>Major Functional Category</th>
<th>Sub-Category</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential</td>
<td>Intangible Framing Attributes</td>
<td><em>the fact that</em></td>
</tr>
<tr>
<td></td>
<td>Quantity Specification</td>
<td><em>in some cases</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>a series of</em></td>
</tr>
<tr>
<td></td>
<td>Contrast and Comparison</td>
<td><em>on the other hand</em></td>
</tr>
<tr>
<td>Stance</td>
<td>Hedges</td>
<td><em>are likely to</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>to some extent</em></td>
</tr>
<tr>
<td></td>
<td>Epistemic Stance</td>
<td><em>we assume that</em></td>
</tr>
</tbody>
</table>

### 3.2 Data Analysis

The researcher will tally the pretest scores on the receptive test (RQ#1) for the comparison group and the treatment group classrooms. The score will be based on a possible 35 points, one point for each correct response. After calculating the mean for each group, the researcher would subject these to an independent samples T-test. This statistic will enable the researcher to infer whether the two groups are equivalent prior to instruction. After treatment, students will take the posttest (i.e. retake the same test), with the items appearing in a different order. A paired T-test will be used to find
out if there is improvement in the students’ receptive knowledge of academic formulas as evidenced by their posttest scores.

The researcher investigated RQ2 by coding the MWSs used by each student in his/her essay, first by classifying them according to the functional categories in the AFL and then assessing them for appropriateness.

4 Results

The results of the objective test suggest an improvement in the students’ receptive knowledge of AFL-MWS. A paired T-Test was used to analyze the difference between pretest and posttest scores for the experimental group. With the p value set to p = 0.05, the obtained T value (t = 5.299) was found to be higher than the critical T (t = 2.26). Based on this analysis, the researcher rejects the null hypothesis. From this result, the researcher cautiously infers that DDL positively impacted students’ receptive knowledge of AFL-MWSs.

Table 2 below categorizes the types of MWSs used in students’ posttest essays by functional category. While the change in the number of AFL-MWSs between pre- and post-test essays was insignificant, the micro-analysis of the formulas used and the induction process provides ample opportunity for comparison. Additionally, misuse of the formulas was not observed. It appears that except for the fact that and in some cases, the students were unfamiliar with these phrases prior to treatment. The appropriacy of target AFL-MWS use within the essays and the inclusion of other relevant academic phrases such as claim that may suggest non-quantifiable learning gains.

<table>
<thead>
<tr>
<th>Category</th>
<th>Formula</th>
<th>Pretest Essay</th>
<th>Posttest Essay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential – Intangible Framing Attributes</td>
<td>the fact that</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Referential – Quantity Specification</td>
<td>in some cases</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>a series of</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Referential – Contrast and Comparison</td>
<td>on the other hand</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Stance – Hedges</td>
<td>are likely to</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>to some extent</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Stance – Epistemic Stance</td>
<td>we assume that</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

5 Discussion and Conclusion

Analysis of the essays revealed that while students used AFL-MWSs sparingly, they used them accurately. Detailed qualitative analysis of MWSs within student essays reveals a connection between the student’s induction and how s/he used the AFL-

83
MWSs. In the examples shown in Table 3, the student’s induction, based on task 2 from the website, was entirely consistent with how s/he employed it in the essay. The greatest uptake was on the noun phrase MWS *the fact that*.

Writing by each of the participants seemed to follow the self-discovered rule for that particular AFL-MWS. If we view the essay excerpts beside the induction, we can see how closely they align, especially for participants 23, 29, 34. The research cited by participant 28 suggest “high probability” of the youth skipping classes and having low self-esteem. Participant 29 used induction to ascertain that the fact that begins a Noun Clause and introduces a significant point, both of which were reflected in the essay. The study by Common Sense Media, cited by participant 34, contradicts the argument that social media negatively impact youth, at least according to and as a result, in the students’ essays may be attributable in part to their textbook, *Final Draft 4* [1]. Every chapter had a section on academic formulas and included in these were the two academic phrases above.

Previous research suggests that students struggle with induction [5, 27]. However, it appears that if induction is modeled and practice activities reinforce it, then students can induce a rule for how an MWS should be used, as evidenced by their objective test scores. In some cases, students take the next step, applying the induction to their academic writing. If researchers conduct future studies that have a greater number of participants and use both treatment and control groups, they should be able to confirm or disconfirm the conclusions drawn from the small exploratory study.

**Table 3. Comparing Student Inductions with Excerpts from Essay.**

<table>
<thead>
<tr>
<th>Induction</th>
<th>Short Context from his/her Essay</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Part. 29 &lt;the fact that&gt;</strong> is the beginning of the noun clause. Most of the time it comes before the pronoun or noun. The author uses the MWS <em>the fact that</em> to put a significant point on the case.</td>
<td>Today over 60% of young people between 13-18 years old have a profile on social media. <em>The fact that</em> children spend too much time on social media make them vulnerable to predatory adults.</td>
</tr>
<tr>
<td><strong>Participant 28 &lt;are likely to be&gt;</strong> after a noun, and before an infinitive verb. This phrase means having a high probability of occurring or being true.</td>
<td>The 2014–2015 School Crime Supplement indicates that, nationwide, about 21% of students ages 12-18 experienced bullying. When kids are cyberbullied, they are likely to have lower self-esteem, be unwilling to attend school.</td>
</tr>
<tr>
<td><strong>Part. 33 &lt;to some extent&gt;</strong> is usually used as an object or a preposition phrase. A writer uses the MWS, <em>to some extent</em> which means partly.</td>
<td>Therefore, they do not have enough time to do other important tasks, such as studying, exercising, real communicating. In the long run, it has a negative effect on the physical development as well as the spirit of youth <em>to some extent</em>.</td>
</tr>
</tbody>
</table>
References

27. Yeh, Y., Liou, H., & Li, Y. Online synonym materials and concordancing for EFL college writing. CALL, 20(2), 131-152. (2007)