Discovering collocations via data-driven learning in L2 writing

Yi-ju (Ariel) Wu, University of Taipei

Abstract

Adopting the approaches of pattern hunting and pattern refining (Kennedy & Miceli, 2001, 2010, 2017), this study investigates how seven freshman English students from Taiwan used the Corpus of Contemporary American English to discover collocation patterns for 30 near-synonymous change-of-state verbs and new ideas about the topic of “change” in the drafting stage of their essay writing. The study used a mixed-methods approach to examine the learning outcomes, learners’ corpus use, and their perceptions of the process. Results were drawn by analyzing writings in three time frames (pre-test, post-test, delayed post-test), video files of corpus consultation, questionnaires, and stimulus recall-session interviews. The results showed that the learners successfully discovered and incorporated collocation patterns in change-of-state verbs and ideas about the topic of change into their essays, although some difficulties emerged. Their performance on change-of-state verbs improved, and this improvement remained three months after the treatment. The study also demonstrated learners’ different perceptions and actualizations of the affordances offered by the corpus. While all learners used the corpus to correct collocation errors, they had diverse attitudes and uses of the corpus to address content ideas or collocation complexities in their writing. The study concludes by discussing the theoretical and pedagogical implications of the results.

Keywords: Corpus-assisted Learning, Collocation Competence, L2 Writing, Reference Resources

Language(s) Learned in This Study: English


Introduction

Collocation, which refers to the co-occurrence of word pairs that are more likely to appear together (Sinclair, 1991; Wray, 2002), is an essential component of L2 learners’ lexical knowledge. Nevertheless, L2 learners struggle to accurately use collocation in language production because they are unaware of the idiom principle (Sinclair, 1991), whereby semi-preconstructed collocations are the building blocks of language. In academic writing, change-of-state verbs are an important element for expressing changes in conditions, reasons, and results (Frodesen & Wald, 2016) in many disciplines (Swales & Feak, 2012). Yet, collocations of change-of-state verbs are challenging for L2 learners because they may not understand the transitivity of verbs (Schleppegrell & Colombi, 2002). Their collocation knowledge of verbs might be insufficient (Boers, Demecheleer, Coxhead, & Webb, 2014), especially when differentiating the collocational behaviors of near-synonyms (e.g., Chan & Liou, 2005) such as expand and extend. Lastly, learners’ collocation production can be characterized by a restricted repertoire (Durrant & Schmitt, 2009), as learners overuse familiar and underuse less familiar collocations. Corpus-based learning, in which learners consult corpora by themselves, including the use of corpus-based concordance lines (e.g., Daskalovska, 2015) and self-correction in writing (e.g., Tono, Satake, & Miura, 2014), has shown to be effective in collocation learning. To address the challenges of using collocations of change-of-state verbs and limited use of collocations and ideas in writing, through triangulation of quantitative and qualitative data, this in-depth study investigated how seven learners used a corpus in pattern hunting and pattern refining activities (Kennedy & Miceli,
(2001, 2010, 2017) to discover collocations in change-of-state verbs and new ideas about “change” in the drafting stage of their writing on the topic of “changes in the future”; learner behaviors and perceptions in the process were also examined.

**Data-Driven Learning**

Corpus-based learning, originating in Tim Johns’ argument for “data-driven learning” (hereafter DDL) (2002), refers to the use of language corpora by language learners. Corpus consultation facilitates constructivist learning (Boulton & Cobb, 2017), which cultivates learner autonomy (Vyatkina & Boulton, 2017) and learning strategies (Han & Shin, 2017), as learners play an active, conscious role in building their own language knowledge by inducing rules from authentic language data (Lin, 2016). Additionally, corpus use has brought learning gains to various aspects of language learning (Boulton & Cobb, 2017; Boulton & Pérez-Paredes, 2014), specifically vocabulary learning (Lee, Warschauer, & Lee, 2019) and writing (e.g., Cotos, Link, & Huffman, 2017).

In writing, the autonomous use of corpora refers to learners’ use of corpora to draft or revise their writing without having obtained prior feedback on the writing (e.g., Chang, 2014; Yoon, 2016). Although learners may find it difficult to incorporate discovered patterns into their writing and formulating questions can be challenging as errors are left unmarked (e.g., Park, 2012), through the autonomous use of corpora, they become independent learners (Yoon, 2008). Their writing significantly improves, and these improvements can remain months after the treatment (Li, 2017).

**Pattern Hunting vs. Pattern Refining**

To describe the autonomous use of corpora and emphasize exploration of both language patterns and content ideas, Kennedy and Miceli (2001, 2010, 2017) coined the terms *pattern hunting* and *pattern refining*. *Pattern hunting* refers to the exploration of the corpus via open-ended questions to find ideas and language patterns that enrich the content and language of a text; *pattern refining* involves searching for language patterns in which learners already know some words of the target patterns to enhance the lexico-grammatical accuracy of a text (Kennedy & Miceli, 2017, p. 3).

Kennedy and Miceli (2001, 2010) investigated how Italian learners wrote autobiographies and engaged in *pattern hunting* and *pattern defining* by consulting the Contemporary Written Italian Corpus (CWIC) corpus, a small, monolingual corpus of “Italian autobiographies” developed by the researchers. Their study showed that while observation and reasoning skills were essential, learners’ involvement, use of, and attitude toward *pattern hunting* and *pattern defining* varied due to their own “reference resource-using style” (2010, p. 40). In a follow-up study, Kennedy and Miceli (2017) demonstrated that learners could successfully develop an “observe-and-borrow chunks mentality” (p. 91) and become effective corpora users by posing open-ended questions for their data queries and remaining open-minded when observing the data.

Adopting the *pattern hunting* approach, Geluso and Yamaguchi (2014) examined how 30 lower-intermediate-level Japanese English learners looked for formulaic sequences in the Corpus of Contemporary American English through a *pattern hunting* activity and how they embedded the patterns in their speech. The results showed a high level of “naturalness” of formulaic sequences embedded in their speech. Learners positively evaluated the *pattern hunting* activity but found it challenging to implement patterns into their speech.

In sum, prior studies have shown that *pattern hunting* and *pattern refining* activities enhanced both lexico-grammatical patterns and content ideas in writing and speaking. Nevertheless, several issues have not been investigated. First, although *pattern refining* and *pattern hunting* are identified as a vigorous approach in DDL (Boulton, 2017), their potential has been insufficiently explored given the small number of empirical studies. Second, as there was no rating of the writing products (Kennedy & Miceli, 2001, 2010, 2017) nor of a pre-test included for comparison (Geluso & Yamaguchi, 2014), whether this approach brings statistically significant and enduring learning effects requires further examination. Finally, although *pattern hunting* and *pattern refining* approaches advocated searching for linguistic features and content ideas as
possibilities of corpora use, how learners differ in their perception and actualization of these affordances is unknown. This study bridges the gap by investigating how learners consulted a corpus in pattern hunting and pattern refining activities to discover collocation patterns in change-of-state verbs and ideas about the topic of “change” in the drafting stage of their writing. Specifically, the study investigates two focuses: (a) corpus use, in particular the process of pattern hunting and pattern refining, and (b) language use, including learners’ performance and improvement of collocation patterns in change-of-state verbs after corpus use, learners’ use of other collocation patterns about “change,” and ideas about the topic of “change.” It is important to note that the two dimensions are highly related and intertwined. Examples in language use exemplify learners’ corpus use, and learners’ corpus use provides a bigger picture of how language use is enhanced through pattern hunting and pattern refining. The study also discusses learners’ behaviors and perceptions of the process.

**Research Questions**

The following research questions were investigated:

1. How do pattern hunting and pattern refining activities affect learners’ performance in using collocation patterns in change-of-state verbs during the drafting stage of essay writing?

2. How do pattern hunting and pattern refining activities affect learners’ use of collocation patterns and ideas about the topic of “change”?

3. What are learners’ behaviors and perceptions of the process of pattern hunting and pattern refining?

**Methods**

**Participants**

This study took place in a year-long freshman English class in one university in northern Taiwan. The class met three hours weekly for 18 weeks in a computer-furnished room. As part of a larger study with 35 learners (see Appendix A), this study focused on seven participants with very different scopes and focuses. The larger study was a quasi-experimental study with a control group (30 students receiving traditional rule-based instruction without corpora) and an experimental group (35 students receiving corpus-based learning). The larger study aimed at comparing learners’ overall writing performance including knowledge, organization, academic style and clarity of essays (Li, 2017) through examining the learner corpus that was built based on the essays of two groups of learners. The current study intends to provide an in-depth understanding of the seven learners’ corpus use in pattern hunting and pattern refining activities, use of collocations of change-of-state verbs and ideas about “change” in their writing, as well as learners’ behaviors and perceptions of pattern hunting and pattern refining, through examining learners’ interviews, videotapes of corpus consultation behaviors, questionnaires and their essays.

The seven participants were non-native English speakers and spoke Mandarin Chinese as their first language. Before taking part in the study, the students had learned English for 10 years and had similar levels of English proficiency (between B1+ to B2 level in CEFR). The seven learners were placed in the same class as a result of a placement test administered by the university (see Appendix B for participant profiles).

The rationale for using this specific group was as follows: First, interviews with the 35 participants showed that these seven learners were particularly reflective about their corpus-consultation process, which could indicate willing and motivated corpus users (Yoon, 2016). Moreover, scholars have called for qualitative analysis and individual case studies of DDL (Godwin-Jones, 2017), and an investigation of seven participants would provide an in-depth understanding of their pattern hunting and pattern refining processes.
Teaching Target

Thirty change-of-state verbs chosen from an academic writing textbook, “Exploring Options in Academic Writing: Effective Vocabulary and Grammar Use” by Frodesen and Wald (2016), were the main teaching target of the study (see Appendix C). These verbs were selected based on the following criteria. First, they had to appear in the Senior High School 7000 words list (administered by the Minister of Education in Taiwan for senior high school students, see Appendix A) to ensure learners’ comprehension of the lexical meaning of the verbs. Second, only words with more than 50% occurrence as verbs in COCA were selected. Finally, to facilitate successful learning through induction from concordance lines in COCA, only verbs that appear with at least 20 collocates, with each collocate including over 20 concordance lines, were chosen.

Materials and Instruments

The experimental procedures comprised the following: (a) a pre-test writing, (b) a preparation phase, (c) a treatment phase (COCA activities: pattern refining, pattern hunting, your own choice), (d) a post-test writing, (e) evaluation questionnaires and interviews, and (f) a delayed post-test writing three months after the treatment.

Instruments

Writings: Pre-test Writing, Post-test Writing, Delayed Post-test Writing

This study adopted a single-group pre-test and post-test design, and three essay writing exercises on the topic of “changes in the future” were implemented in three time frames (pre-test writing: week 7; post-test writing: week 15; and delayed post-test writing: week 27).

The learners were given 90 minutes to complete the pre-test writing, titled “Fifteen changes in a century,” in week seven, without access to any reference resources (Appendix C outlines the instruction of the writing). For the post-test writing in week 15, learners wrote on the same topic. They were also asked to incorporate at least ten patterns they had collected from the treatment of three corpus activities. Finally, to test whether the effects of the corpus-consultation activities would remain, a delayed post-test writing with a similar topic about changes in the future, titled “Fifteen changes in Asia in two centuries,” was implemented three months after the treatment, using similar writing prompts (week 27). The learners were not allowed to use any reference resources. Throughout the three writings, they could not check their previous essays in order to ensure that their writing was original.

For the three essays, the learners had to choose 15 of the 30 change-of-state verbs. To ensure the students’ comprehension of the lexical meaning of the verbs, a Chinese translation obtained from English-Chinese dictionaries was provided. Yet, students were reminded that they should not rely too much on Chinese translations. They were also reminded to provide sufficient elaboration and coherence in their essays, rather than treating the exercise as a practice of “sentence making.”

Questionnaires and Interviews

The study administered two questionnaires in Chinese, with 5-point Likert-scaled questions and open-ended requests for further elaboration of the questions (Dörnyei & Taguchi, 2010). The first questionnaire (week 7) asked background questions, including English grades and learning history, technology use in language learning, and understanding of the concept of collocations. The second questionnaire (week 15), administered immediately after their post-test writing, focused on their corpus use and post-test writing and included three dimensions: (a) positive aspects of COCA use, (b) difficulty in COCA use, and (c) incorporating patterns and write-ups.

The first dimension examined the aspects that learners found helpful in COCA use, focusing on their attitudes toward using corpora to explore and collect language patterns and content ideas by investigating words that they anticipated would be useful in their writing (six open-ended questions). The second topic intended to explore the difficulties the learners encountered when inducing and selecting patterns for possible future use (15 5-point Likert-scaled questions). The last topic investigated how learners borrowed
patterns and incorporated them into their essays and how they organized their post-test writing (seven open-ended questions) (see Appendix D).

Each participant was interviewed twice. The first semi-structured, follow-up interview with lead questions based on the questionnaire results was initiated right after the completion of the second questionnaire (week 15), to further probe answers from the questionnaire and their writings. The second interview was a stimulated recall session (Park, 2012; Yoon, 2016) based on video recordings of learners’ corpus consultation and was conducted within one week after the first interview (week 16).

**Videotape Files of Corpus Consultation and Stimulated Recall Session**

The students were required to videotape their corpus-consultation behavior on their computer monitor. Each video clip lasted approximately 80 minutes, and five screen recordings were collected from each student. The researcher watched the video and made notes about the corpus-consultation process, learners’ strategies and pitfalls. This became the source material for the stimulated recall session.

**Treatment**

In total, four instructional treatments were implemented, including one preparation phase and three phases of COCA activities (*pattern refining, pattern hunting, your own choice*).

The preparation phase included instruction on change-of-state verbs, dictionary use, awareness raising of collocation and DDL, and a corpus consultation workshop. The Corpus of Contemporary American English (Davies, 2008) was chosen as the corpus tool for its large size (containing more than one billion words and updated biannually) and inclusion of mainly native-speaker data (Chang, 2014). In the corpus consultation workshop, after the search functions of COCA and concordance interpreting skills were introduced, the students were taught to use COCA for *pattern hunting* (obtaining content or ideas about “trip”) and *pattern refining* (the students wrote five sentences about their ideal trip and checked COCA for patterns). The instruction involved teacher demonstration and students’ hands-on practice.

Next, three COCA activities were conducted on searching for patterns that students wanted to include in their post-test writing, starting with a *pattern refining* activity for two weeks (80 minutes per week). For the *pattern refining* activity (shown in Figure 1), the students provided extended collocations for the target collocations containing change-of-state verbs (e.g., *to 105 degrees after temperatures soar*), with analysis of its POS (Part of Speech, e.g., n+v+prep+n) and at least three additional collocates of the target change-of-state verb (e.g., *costs, spirit, stocks corresponding to soar*). *Pattern hunting* (80 minutes per week) followed right after and continued for two weeks. The students searched COCA for the eight most-used nouns (*change, development, problem, life, population, technology, environment, Internet*) from their pre-test writing and supplied two concordance sentences for each colligation pattern of the target noun (V+N, N+N, ADJ+N, N+V, N1+of+N2), as shown in Figure 2.

In the following week, students engaged in a *your own choice* activity for 80 minutes, the divergent task at the final stage, in which they used corpus consultation to search for whatever they wanted to know for their post-test writing. The induced patterns and concordances from the three COCA activities were reported on a Google Docs template (see Figures 1 and 2 for examples) as the reference for their post-test writing (see Appendix E for the full procedure of the study).
Figure 1
Example of a Student’s Notes from Pattern Refining

1-(1) You are searching for the title of the essays in which the name of the student “Ariel Wu” is written. The target collocation used in your essay is: With the number of population soaring up rapidly, many people in the poor family die from famine.
1-(2) You are searching for the word "soar" on COCA: Search result: Soar
1-(3) You are searching for the class of the searched word: Transaction
1-(4) You are using COCA as a tool to find collocations.
1-(5) "COCA", "search", "rr": Search result: COCA, the word you have set to: Noun
1-(6) You are searching for the word "left" on COCA: Search result: Left or Right: Left
1-(7) You are searching for the number of results: COCA returned 2 texts.
1-(8) Excerpts of text found with the searched words from COCA:

Prices soar into the tens or even hundreds of millions of dollars, and the room erupts.

1-(9) Prices soar into the tens or even hundreds of millions of dollars, and the room erupts.

1-(10) The type of collocation you found: Noun + Verb

2. Prices soar into the tens or even hundreds of millions of dollars, and the room erupts.

1-(11) The type of collocation you found: Noun + Verb

2. Prices soar into the tens or even hundreds of millions of dollars, and the room erupts.

1-(12) The type of collocation you found: Noun + Verb

2. Prices soar into the tens or even hundreds of millions of dollars, and the room erupts.

2. Prices soar into the tens or even hundreds of millions of dollars, and the room erupts.
Figure 2
Example of a Student’s Notes from Pattern Hunting

Data Analysis

To answer research question one (RQ1) about the effects on learners’ performance when using collocations in change-of-state verbs before and after the pattern hunting and pattern refining activities, the scores of the pretest writing, post-test writing and delayed post-test writing were compared to investigate potential differences. The scores were obtained from three native speakers of English who independently rated, on a rating scale of 1–5 (see Appendix F), each borrowed collocation pattern of the change-of-state verbs. The inter-rater reliability reached 0.82. Scores were then analyzed using ANOVA descriptive statistics to determine whether the differences were significant.

To answer RQ2, regarding how learners changed in using collocation patterns and ideas about the topic of “change” in the pattern hunting and pattern refining, I examined the learners’ three essays and corpus-consultation notes. Then, I developed two categories: (a) the type of borrowing and (b) the type of usage. Based on each category, I compared the induced patterns incorporated into the learners’ posttest essays with the corresponding usages in their pretest essays, determined the relationship between the two corresponding usages, divided them into several sub-categories identified within each category, and
counted the number of the patterns in each sub-category.

For RQ3 regarding learners’ behaviors and perceptions of borrowing patterns in the pattern hunting and pattern refining processes, methods were drawn from previous research (Geluso & Yamaguchi, 2014; Kennedy & Miceli, 2017). I examined the transcripts of interviews and checked learners’ essays and corpus consultation videos that showed learners’ behaviors and perceptions in the pattern hunting and pattern refining processes. Later, I identified three themes that emerged as the most significant: (a) learners’ type of borrowed patterns, (b) learners’ purposes of borrowed patterns, and (c) learners’ difficulties in borrowing patterns, and completed thematic coding (Dörnyei & Taguchi, 2010). These results were then triangulated with the quantitative results to draw broader conclusions.

**Results**

**RQ1: Writing Performance on the Collocation of Change-of-State Verbs over Time**

From the seven participants, a total of 93 collocations of change-of-state verbs were identified and rated in the pre-test writing, with 113 in the post-test writing and 102 in the delayed post-test writing. The average word counts in the pre-test, post-test, and delayed post-test writing were 303, 324, and 336 words, respectively. Figure 3 presents an overview of the learners’ collocation performance using change-of-state verbs in the three writings (out of the total = 5.00). Their performance improved from the pre-test writing in week seven ($M = 3.48, SD = 1.20$) to the post-test writing in week 15 ($M = 3.91, SD = 1.08$) and remained in the delayed post-test writing in week 27 with a slightly higher score ($M = 4.05, SD = 1.17$).

Repeated-Measures ANOVA was conducted to compare the effect of time on their writing performance. A statistically significant effect of time on the scores after corpus use was found ($F(2, 20) = 5.807, p = .017$). The results showed a significant change in the scores of the seven subjects on the three tests ($p = .017$), which was confirmed by the Eta-square effect size analysis as showing a large effect size (partial $\eta^2 = 0.49$). Partial $\eta^2$ was used because we wanted to know the percentage of variance in the mean scores of the three tests (i.e., pre-test, post-test, delayed post-test). Cohen’s d was not used as it only indicates the size of the difference between the mean scores of two tests as a pair (e.g., pre-test & post-test). According to Cohen (1988), a partial $\eta^2$ value over 0.14 indicates a large effect size. A paired t-test for pairwise comparison showed a statistically significant difference in scores between the pre-test writing and the post-test writing ($p = .018$) and between the pre-test writing and the delayed post-test writing ($p = .016$), although no statistically significant difference between the post-test writing and the delayed post-test writing ($p = .24$) was found. These results indicate that the corpus activities helped improve and sustain learners’ collocation use of change-of-state verbs.

**Figure 3**

*An Overview of Learners’ Collocation Performance of Change-of-state Verbs Over Time*
RQ2: Learners’ Use of Collocation Patterns and Ideas about “Change”

Among all the collocation patterns borrowed into learners’ post-test writing \( n = 105 \), 54 items (51%) were collocations of change-of-state verbs, and 51 (49%) were collocations with no change-of-state verbs. Two aspects were examined to understand the process: (a) the type of borrowing and (b) the type of usage.

First, regarding the type of borrowing, single two-word collocation borrowing was the most common (frequency = 71 items, 68%), followed by longer phrase/clause borrowing (24 items, 23%) and longer sentence(s) borrowing (10 items, 9%).

Second, regarding the type of usage incorporated in the post-test writing, I identified three types, in which change-of-state verbs and eight nouns designated in the pattern hunting activity were taken as the node words. Same usage refers to the formulaic patterns of node words incorporated into the post-test writing that shared the main collocate (noun when examining change-of-state verbs and verb when investigating eight nouns) with the ones in the pre-test writing, such as “financial economy shrinks” in the pre-test writing and “the world economy shrinks” in the pre-test. New usage refers to a different main collocate used in the pre-test and post-test writings of the same node words, such as “industry loses benefits” in the post-test writing and “people lose their health” in the pre-test writing. Change of transitivity refers to a collocation pattern of the same node word and main collocates in both writings, but the transitivity was changed, such as “slashes spending” and “spending slashes.”

Among 105 borrowed items that appeared in the post-test writing, 84 items also showed up in the pre-test writing. Of those 84 items, more than half (57 items, 54%) were new usages, compared to 23% (24 items) that were the same usages that appeared in their pre-test writing. Only 3% (three items) of the observed changes involved a change of transitivity.

RQ3: Learners’ Behaviors and Perceptions of the Pattern Hunting and Pattern Refining Processes

In examining learners’ behaviors and perceptions of the pattern hunting and pattern refining processes, three dimensions emerged as the most significant: (a) learners’ type of borrowed patterns, (b) learners’ purposes of borrowed patterns, and (c) learners’ difficulty in borrowing patterns.

(a) Learners’ type of borrowed patterns

First, learners differed in the type of patterns borrowed, contingent on their preferences for the novelty of induced patterns, the familiarity of vocabulary constituents, and their efforts to map meaning. Their choice of patterns was divided into three types: familiar patterns with familiar vocabulary constituents, novel patterns with familiar vocabulary constituents, and novel patterns with unfamiliar vocabulary constituents.

Learners such as Yen and Hao borrowed mainly familiar patterns with familiar vocabulary constituents derived from the pre-test and modified through corpus consultation. They paid little attention to new and unfamiliar usages and were suspicious of borrowing new patterns in their essays because “those items beyond my [their] mastery of vocabulary are [were] too risky to use...more errors could be made accordingly” (Hao, final interview). For example, Hao searched the corpus to modify the chunk “accelerate the speed of aging population” in his pre-test writing. He corrected the pattern and changed it to “a rapid aging population” in his post-test writing, which is the pattern he reported knowing but forgetting in the pre-test writing.

Learners such as Wei and Chun preferred finding “novel combinations” of familiar vocabulary constituents, although they also avoided choosing patterns with unfamiliar vocabulary and rarely spent time with other reference resources. For example, Wei elicited the pattern “undergo a sex change” in the concordance “teenager who killed himself when his parents objected to his desire to undergo a sex change” and incorporated it into the topic of Gay Pride in his post-test writing. He described the unexpected finding as “the excitement of learning something new effortlessly from something old” because he knew the meaning of each constituent of this newly induced pattern.
Three learners, Ting, Yue, and Xin, favored the last type: choosing novel patterns with unfamiliar vocabulary constituents. These learners viewed corpus practice as “a precious learning opportunity” (Xin, final interview) and devoted time to consulting other reference resources to clarify the meanings of patterns. They tended to incorporate unfamiliar patterns with difficult vocabulary into their post-test writing (e.g., “population dwindled” in Table 2) and favored longer clauses or complete concordance lines (e.g., “villages disappear as the value of coastal land skyrocketed” in Table 4).

(b) Learners’ purposes of borrowed patterns

Additionally, learners also displayed a wide array of purposes, identified as (1) enhancing collocation accuracy, (2) collocation complexity, and (3) enrichment of content ideas when borrowing patterns into writings.

(1) Collocation accuracy

First, all learners expressed positive evaluations of how the corpus helped them to find accurate collocations. Most learners’ corpus consultations reflected their high awareness of transitivity and precise collocates, as many learners alternated their observations of “left” or “right” of the searched verb to elicit the use of verbs as transitive or intransitive. Table 1 illustrates how Wei’s wrong use of “slash” as an intransitive verb in the pre-test writing was corrected through corpus consultation, and it remained correct in the delayed post-test writing.

Table 1
Example of Wei’s Correction of Transitivity

<table>
<thead>
<tr>
<th>Test</th>
<th>Corresponding sentences in writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test writing</td>
<td>The garbage worldwide will dramatically slash…</td>
</tr>
<tr>
<td>Post-test writing</td>
<td>The price plunges with the costs being slashed…</td>
</tr>
<tr>
<td>Delayed post-test writing</td>
<td>To slash their costs, business owner tend to…</td>
</tr>
</tbody>
</table>

(2) Collocation complexity

The learners also stated that the corpus tool informed them of advanced patterns or longer phrases with greater sophistication, as evident in their post-test writing. Table 2 shows how Yue’s use of change-of-state verbs to describe “population” improved in both accuracy and complexity after the corpus use, as she not only corrected a wrong usage but also used the advanced change-of-state verbs (“accelerate”, “dwindle”) to collocate with “population”.

Table 2
Example of Yue’s Sentences with “Population”

<table>
<thead>
<tr>
<th>Pre-test Writing</th>
<th>Corresponding sentences in post-test writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Although the population will slowly reduce…</td>
<td>…many animals have diminishing habitat and their population dwindled to 10%.</td>
</tr>
<tr>
<td>…the growing of the population will gradually drop down</td>
<td>As the growth of population accelerate on the earth II…</td>
</tr>
</tbody>
</table>

(3) Enrichment of content ideas
Some learners found that induced patterns helped them to generate new ideas for writing. Table 3 shows how Chun was inspired to incorporate a new topic about “the Internet” in the post-test writing, after consulting COCA on the use of “gain” and finding concordances about “the hackers,” which was evident by her note “I can write about hackers” that she left for herself. She expressed her gratitude by saying, “COCA is like a magic wand which activates my imagination... I have many new thoughts now to be included in my writing” (Chun, first interview).

Table 3
Example of Chun’s Enrichment of Content Ideas

<table>
<thead>
<tr>
<th>Test</th>
<th>Corresponding sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test writing</td>
<td>In order to gain more resource and expand their occupation...</td>
</tr>
<tr>
<td>Concordance lines and note</td>
<td>The FBI is warning that hackers may try to gain control of a cockpit's navigation system ... (可以寫到 hackers!)</td>
</tr>
<tr>
<td>(Translation of the note: I can write about hackers!)</td>
<td></td>
</tr>
<tr>
<td>Post-test writing</td>
<td>The hackers are like soldiers at that time, hackers may try to gain control of other country by hacking other country’s internet system. If one control the internet, the probability he wins the world will significantly climb.</td>
</tr>
</tbody>
</table>

Table 4 shows the change in Xin’s depiction of her hometown. In the pre-test writing, her depiction was micro-oriented, plainly describing her experience of buying bread in a convenience store as an analogy of rising prices and changes in society. In the post-test writing, she elevated her depiction to a macro-oriented, societal level by using newly induced patterns borrowed from concordances, including “religious life” and “villages disappear” and “the value of coastal land skyrocketed.” She indicated that borrowed patterns made her “feel empowered as a university student who could write sentences of ‘higher level of complexity and sophistication’” (Xin, first interview).

Table 4
Example of Xin’s Enrichment of Content Ideas

<table>
<thead>
<tr>
<th>Test</th>
<th>Corresponding sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test writing</td>
<td>When she walked into the store, clerks didn’t say “hello” to the customer, the warm and love images in Ann’s mind toward the convenience store slashed. As she looked the products on the shelf, she couldn’t believe what she saw. Compared to the past, the prices skyrocketed because no one wanted to be a farmer in villages.</td>
</tr>
<tr>
<td>Concordances</td>
<td>(1) Religious life has passed through far more difficult days than the present</td>
</tr>
<tr>
<td></td>
<td>(2) He had seen shrimping villages disappear as the value of coastal land skyrocketed.</td>
</tr>
<tr>
<td>Post-test writing</td>
<td>Traditional religious life is going not to exist anymore, and the village, which full of versatile villagers and good images, shrinks rapidly. The most depressing truth is that the village disappears as the value of coastal land considerably skyrocketed and they even don’t gain any attractions.</td>
</tr>
</tbody>
</table>

After their corpus use, the learners also used more precise and advanced language to embody a fuller, more sophisticated description of their life experiences. Table 4 shows the change in Xin’s depiction of her hometown. In the pre-test writing, her depiction was micro-oriented, plainly describing her experience of buying bread in a convenience store as an analogy of rising prices and changes in society. In the post-test writing, she elevated her depiction to a macro-oriented, societal level by using newly induced patterns borrowed from concordances, including “religious life” and “villages disappear” and “the value of coastal land skyrocketed.” She indicated that borrowed patterns made her “feel empowered as a university student who could write sentences of “higher level of complexity and sophistication” (Xin, first interview).

While learners unanimously praised corpus use for enhancing their collocation accuracy, their attitudes
varied regarding whether the corpus helped with collocation complexity and idea development. Learners such as Xin, Yue, and Chun explored “advanced and unfamiliar patterns” in the corpus to enhance their essays and incorporated complex and advanced patterns (Table 2), new topics inspired by the induced patterns (Table 3), and greater sophistication in depiction (Table 4), but learners such as Yen and Hao mostly consulted the corpus to check the accuracy of old usages in their pre-test writing or their assumptions about collocations.

(c) Learners’ difficulties in borrowing patterns

Although the learners differed in terms of their preferences for borrowed patterns and their purposes of borrowing patterns varied, their perception of borrowing patterns was similar: borrowing patterns was easy and familiar because of their experiences of incorporating patterns into their essays from a collection of “good usages” provided by instructors in their English classrooms. Nevertheless, when examining the longer clause(s), sentence(s) or paragraph(s) where borrowed chunks were incorporated into their essays, numerous pitfalls were found.

First, the learners might have induced the patterns correctly, but when they extended the induced collocation patterns into longer and holistic units, the extended collocations were problematic. Example (1) in Table 5 shows that although Wei successfully induced the pattern “diminish the value” from (1a), his implementation of the pattern with the extended collocation phrase “diminish ‘humane’ value” in (1b) was incorrect.

Second, the learners failed in “making the patterns their (learners’) own” (Kennedy & Miceli, 2017, p. 5) by recontextualizing the corpus concordances in their writings. Several patterns in the post-test writings were borrowed without appropriate adaptation, such as reorienting the pronouns and verb tenses of the borrowed clauses to the sentences they were writing. Example (2b1) in Table 5 shows that Yue failed to change the past tense in the concordances into the future tense that the writing required. Likewise, she failed to provide clear pronoun referents when she resituated the addressee of the induced patterns into those suitable for the sentences she wrote, as Example (2b2) shows.

Third, although some longer sentence(s) borrowings were carefully adapted and incorporated into the learners’ essays, some borrowing beyond the sentence level was characterized by inappropriate textual borrowing, including lack of elaboration and plagiarism (Li & Casanave, 2012). Example (3) in Table 5 demonstrates that Xin presented a “laundry list” of items in her writing (3b), with patterns directly copied from the concordances (3a). She did not elaborate on any of the items in her sentences, nor did she provide logical or temporal connectives to explicitly blend the borrowed sentence into the sentence she generated.

Finally, the inappropriate textual borrowing also resulted in another serious issue, plagiarism, which raises ethical concerns. Example (4) in Table 5 illustrates how Xin’s use of the pattern in her sentence (4b) included copying the whole paragraph of the concordance line from the corpus in (4a). Nevertheless, Xin was shocked to learn that her copying of sentences, which she had learned from all her English teachers was a “model of good usage,” was now viewed as misbehavior with serious consequences: “… Plagiarism was copy and paste of others’ assignment…but not modeling on good usages like what I did” (Xin, the final interview).
Table 5
Examples of Learners’ Difficulties in Incorporating Patterns

<table>
<thead>
<tr>
<th>Types of Difficulty</th>
<th>Student Name</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors in extended collocation</td>
<td>Wei</td>
<td>(1a) Supporters of the law said the phonies <strong>diminish</strong> the <strong>value</strong> of the prestigious awards.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1b) Robots <strong>diminish</strong> humane <strong>value</strong></td>
</tr>
<tr>
<td>Failure to recontextualize the concordances</td>
<td>Yue</td>
<td>(2a1) As of 2012, 82 percent of U.S. <strong>households had access to high-speed Internet</strong>…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2b1) Third, <strong>households had access to high-speed internet</strong>, so there’s no need to worry about the slow rate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2a2) Therefore, the social constructivist environment includes activities where students <strong>experience their level of understanding and seek assistance to get to the next level</strong>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2b2) Parents love to send their children to schools which can allow them to <strong>experience their level of understanding and seek assistance to get to the next level</strong>.</td>
</tr>
<tr>
<td>Insufficient elaboration</td>
<td>Xin</td>
<td>(3a) Psychological needs often include <strong>issues of crisis intervention, personal hygiene, mental health, substance abuse, self-esteem, and a lifestyle allowing for safe living</strong>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3b) And, some groups emphasize on the mental health. They deal with <strong>issues of crisis intervention, personal hygiene, mental health, substance abuse, self-esteem, and a lifestyle allowing for safe living</strong>. Besides, some even notice the severity of the significantly steep population.</td>
</tr>
<tr>
<td>Plagiarism</td>
<td>Xin</td>
<td>(4a) Many congregations and nongovernmental organizations are at the cutting edge of creative social engagement: developing community projects focused on sustainable agriculture and water quality.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4b) Although the government is noticeably poor at addressing difficulties. Fortunately, <strong>many nongovernmental organizations are at the cutting edge of creative social engagement developing community projects focused on sustainable agriculture and water quality</strong>.</td>
</tr>
</tbody>
</table>

Discussion

This study investigated whether the combination of **pattern hunting** and **pattern refining** helped learners to
draft their academic writings. Through data triangulation that connected learners’ writing performance in three time frames, video files of corpus-use behavior, and learners’ perceptions through questionnaires and follow-up interviews, the study provided an in-depth picture of how learners’ behaviors and perceptions in pattern hunting and pattern refining, focusing on discovering collocations of change-of-state verbs, was associated with their immediate and sustained improvement in writing. It also shed light on how learners prepared and collected collocation patterns to describe “changes” in this process. The study complements the findings of Kennedy and Miceli (2017) by addressing their methodological restriction, as they failed to map learners’ perceptions, writing performance, and corpus use.

Learning Change-of-State Verbs and Other Collocations about “Change”

First, through a rating measurement on change-of-state verbs in writing exercises conducted over three time frames, the statistically significant results of this study demonstrate how pattern hunting and pattern refining enhanced students’ collocation use in academic writing, compared with prior studies that did not include pre-tests for comparison (Geluso & Yamaguchi, 2014; Kennedy & Miceli, 2001, 2010, 2017). The learners in this study differentiated the collocation use of near-synonyms (Laufer & Waldman, 2011) in change-of-state verbs and incorporated correct collocations with a higher level of complexity (Huang, 2014). These findings demonstrate learners’ heightened awareness of the “idiom principle” (Sinclair, 1991) and “chunk-aware mentality” (Kennedy & Miceli, 2017, p. 14) in language production.

Second, the learners in this study not only autonomously consulted corpora to draft their writing without having obtained prior feedback (e.g., Park, 2012; Li, 2017; Yoon, 2008), but by inducing and selecting patterns for possible future use, the learners were also proactive by exploiting the corpus to prepare and collect language patterns and content ideas prior to their actual production, as evident in the unusual post-test writing results. For example, 30% of the textual borrowing went beyond “two-word collocation,” and the learners incorporated two times more new usages than old usages, compared with their pre-test writing. This result contrasts with those of prior studies in which learners mostly confirmed assumptions and rarely elicited new usages (e.g., Yoon, 2008).

Thirdly, the learners in this study did not merely copy and paste induced patterns into their writing for language accuracy per se; they further transferred the patterns originally addressing collocation errors to develop and enrich the ideas in their writings, as evident in Chun’s “hackers may try to gain control of other country.” This indicates that the learners not only showed the “observe and borrow chunks mentality” (p. 91) used by effective learners in Kennedy and Miceli’s (2017) study; they further transferred chunks to fulfill multiple affordances of corpora.

Finally, the learners in this study not only improved their collocation use of change-of-state verbs in the immediate post-test writing, but their performance improved slightly in the delayed post-test writing. The improvement reflected the noticing hypothesis of Schmidt (2001), such that the learners’ conscious attention to linguistic input enhanced their acquisition of input. In the study, the input from the concordances was enhanced through noticing (Flowerdew, 2015), i.e., learners’ active attention to recurrent phrases in concordances in the three COCA activities, including conscious comparison of the corpus input and the learners’ output (Li, 2017), exploration of the lexical and grammatical environments of collocations (Thomas, 2015), and learners’ implementation of induced patterns in their post-test writings. Those practices entailed deep, thoughtful mental processing of language input, which ultimately manifested as “linguistically longer-term benefits of DDL” (Boulton, 2011, p. 1) through the learners’ intake (Schmidt, 2001) of collocation patterns of change-of-state verbs in the delayed post-test writing three months after the treatment.

Learners’ Various Uses and Perceptions of the Multiple Affordances of the Corpora

Learners’ types of borrowed patterns varied significantly, as evident in their various preferences toward the novelty of induced patterns, the familiarity of vocabulary constituents, and their efforts to map meaning. Moreover, their distinct choices derived from various purposes of borrowed patterns: enhancing collocation accuracy, collocation complexity, and enrichment of content ideas, identified as multiple affordances of
corpus (Leńko-Szymańska & Boulton, 2015). These results demonstrate not only the learners’ autonomy in corpus use but also the new dimensions of individual differences in DDL. First, while learners differed in their correction rates when they implemented induced patterns to self-correct writing errors (Tono et al., 2014; Wu, 2016), their preferences of the types of patterns borrowed also varied. Furthermore, learners did not only differ in their corpus-consultation behaviors (e.g., Yoon, 2016) such as their individual “reference-resource-using style” (Kennedy & Miceli, 2010, p. 40), their use of corpora in relation to other reference resources (Lai & Chen, 2015), and their evaluations of corpus use (Lee & Swales, 2006); they also displayed diverse perceptions and actualizations of the multiple affordances offered by corpora (Hafner & Candlin, 2007; Yoon, 2016).

Learners’ Lack of Awareness of Pitfalls in Borrowing Patterns

Finally, the results illustrate that the learners lacked awareness of some of the pitfalls of borrowing patterns into their essays. Contrary to findings from prior studies (e.g., Geluso & Yamaguchi, 2014; Park, 2012), the learners in this study did not find borrowing patterns difficult because “modeling good usages” from authoritative sources was a familiar literary practice in English classrooms (Li & Casanave, 2012). Nevertheless, the learners still encountered some difficulties incorporating patterns into their writings, including inappropriate textual borrowing, erroneous extended collocation, no recontextualization of concordances, and plagiarism. This indicates that they were not fully capable of authenticating the corpus data (Mishan, 2004) by making the use of the patterns they induced in pattern hunting and pattern refining in their own essay writings. The results, which enumerated specific types of pitfalls in borrowing patterns, also shed light on the gap between learners’ perceptions and their actual use of a corpus (Wu, 2015). Specifically, learners showed much higher awareness of the difficulty in inducing patterns from a corpus than incorporating patterns into their essays, but they encountered more difficulties in the latter.

Conclusion

The findings suggest that in the pattern hunting and pattern refining activities, learners were proactive in exploiting corpora to “prepare and collect” language patterns and ideas about changes in preparation for writings, although learners differed in their perceptions and actualization of multiple affordances of corpora (Leńko-Szymańska & Boulton, 2015). An examination of the learners’ writings revealed that, although they encountered some difficulties in incorporating induced patterns into essays (Geluso & Yamaguchi, 2014), their collocation use in writing improved in terms of both accuracy (Li, 2017) and complexity (Huang, 2014), which showed their heightened awareness of the idiom principle (Sinclair, 1991) and chunk-aware mentality (Kennedy & Miceli, 2017). Specifically, learners differentiated the collocation use of near-synonyms (Laufer & Waldman, 2011) in change-of-state verbs in both the post-test and delayed post-test. This finding provides support for Schmidt’s (2001) theoretical construct of noticing, as it indicates that pattern hunting and pattern refining enhanced learners’ noticing of input about collocations of change-of-state verbs from concordances and helped them to “intake” it, as the improvement was sustained three months later. These findings, as well as those of prior studies, show that suggestions regarding appropriate learner training and guidance are needed to illicit positive learning effects.

The first pedagogical implication of this study echoes Kennedy and Miceli’s (2001, 2010, 2017) warning that pattern hunting for content and idea development should not be peripheral to pattern refining for linguistic accuracy. In the study, learners’ top concern in correcting linguistic errors drove some of them to explore the corpora only to address linguistic accuracy. Thus, it is suggested that their possible negligence of corpus affordances of enhancing language complexity and content ideas should be preempted. Learners need to be taught and guided to exploit the full array of the multiple affordances of a corpus, including enhancing linguistic accuracy, linguistic complexity, and content enrichment. Secondly, learners should also be encouraged to actively exploit the potential of induced patterns, because induced patterns originally used to address errors could further spark the development of ideas.

Finally, given that scholars have emphasized the importance of learner training in pattern induction (e.g.,
Han & Shin, 2017), this study further advocates learner guidance for incorporating patterns into writing, such as useful strategies and pitfalls to avoid, as shown in this study. Specifically, for undergraduate non-
English majors, even those with intermediate proficiency like the learners in this study, insufficient training in academic writing could result in numerous pitfalls in incorporating induced patterns.

Although the study shed some new light on the under-researched approach of DDL, *pattern hunting* and *pattern refining* (Boulton, 2017), there were some limitations which lead to suggestions for future research. First, although the learners demonstrated that they could “prepare and collect” language patterns and content ideas in corpus activities and ultimately incorporated the patterns into their essays, we do not know if the corpus literacy developed in those activities was transferrable to new tasks. It would be intriguing to investigate whether the same group of participants could apply the corpus consultation skills they learned for both language patterns and content ideas in this task to another new writing task with similar writing prompts. Second, since this study focuses on how learners consulted corpora in *pattern-hunting* and *pattern-refining* activities in the drafting stage of writing, it would be useful for future research to explore the process and learning effects of *pattern-hunting* and *pattern-refining* activities in different stages of the writing process, such as the revising stage of writing.

**Acknowledgements**

I would like to thank the LLT editors, the anonymous reviewers, Professor Zhao-Ming Gao, Professor Hsien-Chin Liou and Professor Yu-Ju Lan for their valuable comments.

**References**


**APPENDIX A. Information about the Larger Study and Vocabulary List**


## APPENDIX B. Participant Profiles

<table>
<thead>
<tr>
<th></th>
<th>Yen</th>
<th>Hao</th>
<th>Yue</th>
<th>Xin</th>
<th>Ting</th>
<th>Chun</th>
<th>Wei</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Female</td>
<td>Female</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Age</td>
<td>19</td>
<td>20</td>
<td>19</td>
<td>19</td>
<td>18</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Field of study &amp; Degree Pursued</td>
<td>Physical therapy (BS)</td>
<td>Pharmacy (BS)</td>
<td>Physical therapy (BS)</td>
<td>Nursing (BS)</td>
<td>Physical therapy (BS)</td>
<td>Physical therapy (BS)</td>
<td>Occupational Therapy (BS)</td>
</tr>
<tr>
<td>Knowledge of collocation &amp; resource used</td>
<td>Yes/none</td>
<td>Yes/none</td>
<td>None</td>
<td>Yes/Longman Dictionary of Collocation</td>
<td>None</td>
<td>None</td>
<td>Yes/none</td>
</tr>
<tr>
<td>Prior experience with corpora</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

*Note. All participant names are pseudonyms*
APPENDIX C. Instruction of the Pre-writing “Fifteen Changes in a Century”

(1) Please write an essay about the changes that you think will happen in a century. You can be very creative and include things that are not likely to happen.

(2) Make sure you provide enough elaboration of each “change” you incorporated. Also, coherence is important. Remember, this is not a “list of sentence-making” but a comprehensive essay.

(3) Please choose fifteen change-of-state verbs out of the thirty change-of-state verbs listed in the table.

<table>
<thead>
<tr>
<th>1. accelerate 加快, 增长，增加</th>
<th>16. lower 放下，降下，减低，减弱</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. climb 爬, 攀登, 上升, 向上爬</td>
<td>17. multiply 成倍增加，繁殖</td>
</tr>
<tr>
<td>3. contract 縮小, 收缩</td>
<td>18. peak 達到高峰</td>
</tr>
<tr>
<td>4. decline 下降, 下跌；减少，衰退</td>
<td>19. plunge 下降，急降，下傾</td>
</tr>
<tr>
<td>5. diminish 减少，縮減，被貶低</td>
<td>20. proliferate 增殖，激增，擴散</td>
</tr>
<tr>
<td>6. drop 下降，變弱，滴下，掉下</td>
<td>21. raise 增加，提出，引起，豎起</td>
</tr>
<tr>
<td>7. enlarge 擴大，擴展，放大</td>
<td>22. reduce 減少，變弱，減輕，降級</td>
</tr>
<tr>
<td>8. escalate 逐步上升，增強，擴大，升級</td>
<td>23. rise 上升，上漲，高聳，起立</td>
</tr>
<tr>
<td>9. expand 展開，張開，膨脹，擴大</td>
<td>24. skyrocket 往上衝，猛漲，高升</td>
</tr>
<tr>
<td>10. extend 延長，延伸，擴大，擴展</td>
<td>25. shrink 收縮，縮短，變小，變少</td>
</tr>
<tr>
<td>11. fall 落下，下降，減少，陷落</td>
<td>26. sink 下沉，衰弱，滲透，降低</td>
</tr>
<tr>
<td>12. gain 得到，獲得，贏得，增添</td>
<td>27. slash 大幅度減少，減低，減少</td>
</tr>
<tr>
<td>13. grow 成長，增大，成熟</td>
<td>28. spread 伸展，散布，分布</td>
</tr>
<tr>
<td>14. intensify 增強，強化，變激烈</td>
<td>29. swell 腫起，增長，增大，高漲</td>
</tr>
<tr>
<td>15. lose 輸掉，失敗，丢失</td>
<td>30. soar 猛增，暴漲，飛騰，昂揚</td>
</tr>
</tbody>
</table>

APPENDIX D. Second Questionnaire

(Note. Three out of the six questions in Part 1, 12 out of the 15 questions in Part 2 and five out of seven questions in Part 3 were included as they had the greatest relevance to the aim of the study.)
(1) Positive aspects of COCA use

1. Do you think corpus searching helps you improve the collocation accuracy in your writing? (As in, turning incorrect into correct collocations) Why or why not? Please give me examples from your essay.

2. Do you think corpus searching helps you enhance the collocation complexities in your writing? (As in, changing correct into advanced collocations) Why or why not? Please give me examples from your essay.

3. Do you think corpus searching helps you increase the content ideas in your writing? Why or why not? Please give me examples from your essay.

(2) Difficulty in COCA use

Please identify the following difficulties when you used COCA by clicking

   __ 1 strongly disagree __ 2 disagree __ 3 neutral __ 4 agree __ 5 strongly agree

1. I had difficulty using COCA because the interface was complicated 1 2 3 4 5
2. I had difficulty using COCA because it took a long time to find one pattern 1 2 3 4 5
3. I had difficulty using COCA because there were too many concordances 1 2 3 4 5
4. I had difficulty using COCA because there were too few concordances 1 2 3 4 5
5. I had difficulty using COCA because of unknown cultural contexts 1 2 3 4 5
6. I had difficulty using COCA because of unknown professional knowledge 1 2 3 4 5
7. I had difficulty using COCA because the query itself was difficult 1 2 3 4 5
8. I had difficulty using COCA because of the cut-off sentences 1 2 3 4 5
9. I had difficulty using COCA because of unfamiliar vocabulary 1 2 3 4 5
10. I had difficulty using COCA because I needed to induce patterns by myself 1 2 3 4 5
11. I had difficulty using COCA because of strong uncertainty 1 2 3 4 5
12. I had difficulty using COCA because I was not sure what I would use in my future writing 1 2 3 4 5

(3) Incorporating patterns and write-ups

1. How did you organize your writing?
2. How did you select patterns induced from COCA activities into your writing? Did you find it difficult? Why or why not?
3. How did you incorporate patterns into your writing? Did you find it difficult? Why or why not?
4. How did you like finding the patterns as preparation for your writing?
5. Please write down any other difficulties you have encountered.
APPENDIX E. Procedure of the Study

<table>
<thead>
<tr>
<th>Time</th>
<th>Objectives</th>
</tr>
</thead>
</table>
| **I. Pre-test Writing** (Week 7) | (1) Background and technology understanding questionnaire  
(2) Pre-test writing titled as “fifteen changes in a century” |
| **II. Preparation** (Week 8–9) | (1) Change-of-state verbs instruction  
(2) Dictionary use workshop  
(3) Awareness raising of collocation and DDL  
(4) Corpus consultation workshop (COCA) |
| **III. Pattern Refining** (Week 10-11) | (1) Students consulted COCA to collect collocation patterns of eight change-of-state verbs chosen from their pre-test writing |
| **IV. Pattern Hunting** (Week 12-13) | (1) Students discovered the collocation patterns of eight most frequently-used nouns in their pre-test writing (change, development, problem, life, population, technology, environment, Internet) via consulting COCA |
| **V. Your Own Choice** (Week 14) | (1) Students searched for anything they wanted to know from COCA |
| **VI. Post-test Writing** (Week 15) | (1) Students incorporated at least ten patterns from their corpus activities into their post-test writing titled as “fifteen changes in a century”  
(2) Evaluation questionnaire and interview probe-up  
(3) Stimulated recall session interview |
| **VI. Interviews** (Week 16) | (1) Students incorporated at least ten patterns from their corpus activities into their post-test writing titled as “fifteen changes in a century”  
(2) Evaluation questionnaire and interview probe-up  
(3) Stimulated recall session interview |
| **VII. Delayed Post-test Writing** (Week 27) | (1) Students wrote the delayed post-test writing titled as “fifteen changes in Asia in two centuries” |
### APPENDIX F. Rating Scale for Collocations of Change-of-State Verbs

<table>
<thead>
<tr>
<th>Scale</th>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Correct collocation &amp; appropriate embedment</td>
<td>The collocation is perfectly correct, and it is appropriately embedded in the sentence, e.g., <em>sea level will rise.</em></td>
</tr>
<tr>
<td>4</td>
<td>Correct collocation &amp; problematic embedment</td>
<td>The collocation is perfectly correct, but it can have multiple interpretations because of insufficient elaboration or clarification, e.g., <em>raise people’s awareness to solve this problem.</em></td>
</tr>
<tr>
<td>3</td>
<td>Correct collocation &amp; problematic transitivity</td>
<td>Correct collocate, but it has problems in the use of transitive and intransitive, e.g., <em>the spending slashes</em></td>
</tr>
<tr>
<td>2</td>
<td>Incorrect collocation, but it is intelligible with some guesswork</td>
<td>e.g., <em>20% of the land will diminish.</em></td>
</tr>
<tr>
<td>1</td>
<td>Incorrect collocation, and it is unintelligible</td>
<td>e.g., <em>the competition in Taiwan will drop.</em></td>
</tr>
</tbody>
</table>

### About the Author

Yi-ju (Ariel) Wu is an assistant professor in Department of English Instruction at University of Taipei, Taiwan. She received her PhD in Education (Applied Linguistics emphasis) from University of California, Santa Barbara. Her research expertise includes corpus linguistics, L2 writing, virtual reality and English for Specific Purposes. She has published in leading international journals such as *Language Learning & Technology* and *Educational Technology & Society*.

**E-mail:** yjarielwu@gmail.com