

令和 5 年 6 月 18 日現在

機関番号：14501

研究種目：基盤研究(C) (一般)

研究期間：2017～2022

課題番号：17K02927

研究課題名(和文) Employing corpus-based Word Sketch collocation summaries as an L2 writing reference tool: The impact on lexical accuracy and diversity of Japanese EFL compositions

研究課題名(英文) Employing corpus-based Word Sketch collocation summaries as an L2 writing reference tool: The impact on lexical accuracy and diversity of Japanese EFL compositions

研究代表者

QUINN Cynthia (Quinn, Cynthia)

神戸大学・国際文化学研究所・非常勤講師

研究者番号：00368474

交付決定額(研究期間全体)：(直接経費) 2,400,000円

研究成果の概要(和文)：学習者はコーパスを用いてコロケーションの誤りを修正することに概ね成功した。これらの誤りを修正するために学習者に課された解釈上の要求がその成功の度合いに重要な役割を果たした。結論に至るまでに解釈や分析が必要であればあるほど学習者が修正に成功する確率は低くなります。このことは分析力を鍛えることの重要性を指摘するだけでなく学習者がコーパスを用いてどのようなエラーを修正することができるのかということも示している。学習者にとっては時間と分析に大きな投資を必要としないエラー訂正が最も生産的であると考えられる。このバランスを取ることが外国語教室にコーパス参照をうまく導入するための重要なポイントかもしれません

研究成果の学術的意義や社会的意義

この研究は、言語学習リソースとしてコーパス参照を導入することを支援するものである。そうすることで、学習者が言語使用を研究するための代替アプローチ(従来の辞書参照に対して)と、生徒がコロケーションを研究するための手段を見つけることをサポートします。

研究成果の概要(英文)：Findings indicate that learners were generally successful in correcting their collocation errors with a corpus; however, the interpretative demands placed on them to address these errors played an important role in their degree of success. Overall, the more interpretation and analysis required to come to a conclusion, the less likely learners were to make a successful correction. This not only points to the importance of practicing analytical skills, but also what types of errors learners are capable of correcting with a corpus. In short, error corrections that do not require a significant investment of time and analysis are likely most productive for learners. Striking this balance may be an important aspect in successfully introducing corpus referencing to the foreign language classroom.

研究分野：L2 writing

キーワード：L2 writing corpus referencing error correction

1 . 研究開始当初の背景

For years, the role of language in the writing classroom has held a tenuous position in L2 writing pedagogy. Early L2 writing classrooms were characterized by an emphasis on grammatical form and error avoidance, but dissatisfaction with this prescriptivist orientation led to the introduction of process writing, an approach imported directly from L1 composition studies (Silva & Leki, 2004). While much has been gained from this approach in terms of understanding the complexity of the writing process and how writers manage it, a consequence of its widespread use has been the subordination of learners' linguistic needs (Hyland, 2011). This insufficient attention to language leads to the larger issue of how to address L2 learners' written accuracy within the overall writing process. Aside from giving feedback to students on an individual basis, many teachers struggle to address their learners' linguistic needs at the curricular level and to incorporate other approaches to written accuracy. One major reason is the lack of time, considering the enormous challenge of adequately covering language, content, and composition instruction along with responding to individual writers' needs. With these factors in mind, corpus referencing has been advocated as a promising resource for the L2 writing classroom.

2 . 研究の目的

The role of corpus referencing in the L2 writing classroom is investigated as a resource for student writers to resolve lexicogrammatical errors in their texts. As a classroom-based study that employs qualitative research methods, the students' corpus-based error corrections were examined in terms of how the learners applied corpus research findings to their writing and what factors influenced their error correction decisions.

With few corpus referencing studies systematically investigating the learner language that results from corpus-based error correction, this study contributes to the current body of research by offering a detailed descriptive account of how learners go about integrating corpus data into their own language production -- a critical juncture in the error correction process. Many corpus-based error correction studies have discussed the effectiveness of corpus referencing largely in terms of the learners' success rates, focusing more on correction outcome than on the correction process. As for qualitative studies conducted on corpus referencing, these have generally involved small groups of learners; in contrast, the current study tracks the correction choices made by a class of 24 students over 15 weeks and across three writing assignments, enabling the researcher to identify recurring correction tendencies and create a more systematic linguistic composite of the learners' behavior.

At the same time, this research contributes to the L2 writing literature by offering a learner-centered approach to error correction in contrast to the extensively researched teacher-centered corrective feedback perspective. Through the application of corpus research methods, this study encourages discussion of other pedagogical approaches to error correction that can bring more responsibility to the learner and thereby create more opportunities for language learning in L2 writing classrooms.

Finally, by exploring what the learners' correction decisions suggest about their ability to make use of corpus data, this study may contribute to a better understanding of how corpora can be employed as a pedagogical resource, an issue for many teachers that has limited its widespread application to language learning classrooms (Chambers, 2019; Frankenberg-Garcia, 2012b). By combining research perspectives from both L2 writing and corpus linguistics, this research aims to increase interest in and broaden acceptance of corpus referencing beyond its current corpus linguist audience (Chambers, 2019), seeking to increase mutual understanding among corpus experts, writing specialists, and language teachers.

3 . 研究の方法

(3a) Data Sources and Participants

This study emerged out of several years of preliminary research, including a two-year grant-supported learner perceptions study conducted in 2014- 2015. The research reported in this thesis draws primarily on data collected in the author's 2016 semester-long writing course and focuses on the learners' texts: 72 sets of original and revised essays written by 24 Japanese learners of English. The learners' error corrections were analyzed item-by-item to examine the quality of their linguistic choices based on the corpus data. In this way, the researcher investigated to what degree learners could apply the corpus data accurately to their writing as well as what kinds of linguistic issues they faced during the corpus-based correction process.

(3b) Writing Curriculum

Students were introduced to corpus research through a basic training module and taught the essay genre being used (reaction/response essay). Each writing assignment proceeded through the following cycle. First, students read and annotated the class-assigned article at home and answered comprehension check questions. Then students annotated the reading with their reactions, which were also shared in groups. The class worked together to transform their ideas and reactions to the article into viable essay topics, based on their annotations and group discussion. After this, students composed their reaction essay drafts at home. Once the teacher received the students' drafts, she prepared written feedback. Essays were then returned to students for revision and error correction with the corpus. During this revision stage, learners responded to the teacher feedback (on a variety of issues) and corrected designated errors with the corpus along with tallying their corrections on a correction log. Once the draft was fully revised and corrected, students re-submitted their final drafts to the teacher again along with the correction logs.

This writing assignment cycle took place three times, each with a different topic and class reading. Throughout the writing process, students also practiced correcting errors with the corpus. Each class typically started with a corpus research task, which ranged from practice with various search tools and approaches to identifying and analyzing patterns in the corpus data to completing error correction exercises. In this way, the researcher aimed to continue developing the learners' corpus research skills regularly throughout the course and within the thematic context of the individual writing assignments.

4 . 研究成果

(4a) Overall findings

Table 1. Corpus-based error correction success rates ranked by error type

Rank	Error Type	Total Corrections (N)	Successful Corrections (N)	Success Rate	
1	Preposition omissions	104	88	84.6%	<i>more fixed/formulaic less interpretive narrower corpus searches</i> ↑↓ <i>less fixed/formulaic more interpretive broader, multi-layered corpus searches</i>
2	Preposition errors	115	93	80.9%	
3	Phrases	84	64	76.2%	
4	Collocations	238	173	72.7 %	
5	Lexical usage	424	295	69.6%	
TOTALS		965	713	73.9%	

Based on this general categorical ranking, a clear pattern can be seen: the more fixed or formulaic the item, the greater the success with error correction; on the other hand, the less formulaic, the less successful. Prepositions, for example, have a generally fixed patterning, which makes it easy for learners to formulate corpus queries and to identify suitable corrections. Furthermore, for a given preposition query, there are relatively few correction options to choose from compared to an error in lexical usage, for instance, which may involve selecting from any number of alternative word choices. The fixed nature of preposition combinations makes the lexical patterning more salient in the corpus data and facilitates data analysis for the learner. To a lesser degree, this is generally true for the phrases and collocations as well, given that these error categories are made up of formulaic sequences that also display systematic usage patterns.

On the other hand, as mentioned above, lexical usage errors could be remedied in a number of ways, given that errors in this category had several correction options. As with any error correction, the writer must analyze the corpus data for each correction possibility in order to come to a good decision. Given the greater range of choices for addressing lexical usage problems, selecting an appropriate correction would require greater skill with corpus referencing than the other more formulaic error types: the writer must weigh various alternatives against their original context while exploring the meanings and usage patterns specific to each word choice. For these reasons, it is not surprising that success rates are lowest for the lexical usage errors.

(4b) Categorical qualitative findings

Although the rates of success in correcting preposition, phrase, and collocation errors were good, further examination of the quality of these corrections reveals that the learners employed a narrow range of correction strategies to address the errors and engaged in a limited degree of corpus data analysis. In addition, the analysis showed that learners had difficulties:

- comprehending the nature of their errors
- parsing the language of the concordances as well as of their own texts, and
- revising their writing based on the patterns discovered in the concordance data.

Importantly, these are factors that are not only necessary to reference a corpus, but also essential for learners to successfully correct their errors, as they both require a good degree of linguistic awareness and critical analysis. Consequently, the linguistic challenges reported in this study offer implications for improving L2 writing pedagogy, particularly in terms of providing learners with better language support throughout the writing process. In order for learners to acquire the self-editing skills they need to address their error tendencies in the long-run, the study's findings highlight the importance of understanding and analyzing lexicogrammatical relations for developing L2 writers.

(4c) Overall conclusions

By examining the learners' error corrections both quantitatively in terms of outcome category and qualitatively in terms of individual process, this study provides detailed insight into the nature of corpus-informed learner language use. While the current study reinforces the general assumption that error type is an important factor to successful corpus referencing, the findings further imply that the degree of interpretation necessary to resolve a particular error also impacts the learner's success.

As mentioned above, the more fixed the patterning of an error, the more successfully it could be addressed by the learners. This outcome is a consequence of factors related to the corpus referencing process, where researching more formulaic language involves more straightforward search methods, produces more salient patterns in the corpus data, and results in fewer correction options overall. In contrast, when researching corrections that display greater variability in patterning, the process becomes more interpretive, and therefore more demanding, as the learner must distinguish the usage of various correction options. This research outcome highlights the fact that error types which tend to be less fixed will involve more interpretation overall, from data analysis to data application, thus requiring more skill from the learner. Therefore, the interpretative demands placed on the learner for a given type of error are an important factor to consider in corpus-based error correction tasks.

The study also revealed that the successful corrections tended to be local errors corrected through lexical substitution and word combinations that did not entail much rephrasing of the learners' original written context. On the other hand, unsuccessful correction attempts involved less predictable language and were therefore more complex to correct, requiring learners to notice a number of linguistic features in the corpus data and apply these features accurately to their writing. Through the study's correction analysis, it became apparent that learners tended to analyze the corpus data vertically, looking for words that could replace their erroneous items, which demonstrates a paradigmatic approach to data analysis. In contrast, corpus linguists are primarily concerned with syntagmatic relations, reviewing concordances horizontally to explore the phraseology and preferences of specific lexical items (Flowerdew, 2009). As language examples to support learners' production, corpus data highlights collocational and colligational behavior, providing input that raises the learners' awareness of morphosyntactic and distributional properties, helping them to achieve accurate usage (Frankenberg-Garcia, 2014). However, in this study, rather than examining the co-text of the error corrections in order to identify alternative means of expression, the learners' approach was more like a writer referencing a thesaurus to find alternative word choices.

This outcome reveals the limits of the learners' data analysis skills, illustrating to what degree they were able to explore the language of their correction options, while also revealing the demands placed on their linguistic abilities. Phraseological analysis is known to be extremely challenging for foreign language learners (Lenko-Szymanska, 2014; Wray, 2002), making the application of such patterning to the learners' own language production at least equally as difficult. A case in point is the test-based assessments in Jones and Haywood's (2004) study that showed learners could improve their awareness of formulaic sequences and their ability to produce such phrases in controlled situations, but when it came to using these phrases in their own writing, no overall improvement was shown. In another study, Frankenberg-Garcia (2012a, 2014) found that when it came to production, the availability of multiple corpus examples helped learners correct the use of words that they understood but often misused, thus reinforcing the value of data-driven learning. However, this success was partly facilitated by the error-relevant data provided to the learners by the researcher. This highlights the fact that for learners referencing corpus examples on their own, a critical juncture is their ability to isolate error-appropriate examples that will enable them to make use of the

language data.

Despite the learners' apparently narrow use of the corpus in the study, they do in fact engage a number of important skills -- skills that are fundamental not only to corpus research, but also to their growth as writers. Students learn to formulate queries based on their individual errors, to sort data in terms of its relevance to their errors, and to make linguistic decisions that are appropriate to their texts. Such skills are important for using most any reference tool or technology, as they require learners to understand the linguistic features of their written context well enough to be able to make use of the language resource. At the same time, developing learners' language analysis skills challenges them to critically assess their own writing and to reflect on their linguistic choices, both of which are essential to successful writing.

(4d) Implications

The results of the study highlight the importance of narrowing correction tasks to specific error types and contexts in order to make corpus-based error correction manageable for learners. The intermediate learners in this study were not familiar with many high-frequency, salient word combinations, and with the corpus they could efficiently research potential corrections, allowing them to test their linguistic hypotheses (Gilquin & Granger, 2010).

Moving beyond this basic level of corpus research to examine syntagmatic relationships encourages learners to view their lexical choices as members of fuller phraseological units and to identify their associated meanings and usage patterns. In the current study, the difficulty of this type of analysis was particularly evident with the moderate phrase corrections, to which learners could make some improvements on their errors but were not able to transfer the patterns from the data to their own text with full accuracy. Even though they were able to identify the error-relevant data, the learners were not able to manage the layers of correction necessary to completely resolve the problem. These partial corrections illustrate the challenges of attending to several aspects of a particular pattern for learners in order to make accurate use of it in their writing, demonstrating that error correction at the phraseological level quickly becomes difficult.

Considering the widely reported challenges of corpus research, placing inappropriate demands on learners who are new to corpus research will only increase these difficulties, effectively discouraging long-term use of corpora beyond the initial classroom experience. In all likelihood, this is at least one reason why corpus technology has not been adopted by teachers or learners on a broader scale: requiring learners to take on substantial data analysis without adequate scaffolding or referencing experience does little to build confidence and sustain motivation. Certainly, the medium itself is an issue to contend with, considering that corpus systems are typically designed by researchers for researchers. However, regardless of what improvements are made in the technology to create more accessible resources for learners, the data interpretation and application skills that are central to data-driven learning do not change: learners must be able to make use of instructive examples in order to improve their written accuracy. With the corpus's main advantage being a phraseologically instructive one, learners need to be able to make use of language samples in order to exploit its unique capabilities for the benefit of their written accuracy.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 1件/うちオープンアクセス 1件）

1. 著者名 Cynthia Crosby Quinn	4. 巻 1
2. 論文標題 L2 writers referencing corpora to address accuracy: A qualitative analysis of learners' lexicogrammatical error corrections	5. 発行年 2021年
3. 雑誌名 University of Birmingham PhD thesis	6. 最初と最後の頁 1-248
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕 計3件（うち招待講演 0件/うち国際学会 3件）

1. 発表者名 Cynthia Quinn
2. 発表標題 Referencing corpus data to address lexical error in the L2 writing classroom
3. 学会等名 The 40th ALAK international conference (Applied Linguistics Association of Korea), Sogang University, Seoul, Korea (国際学会)
4. 発表年 2018年

1. 発表者名 Cynthia Quinn
2. 発表標題 Corpus referencing to address lexical error L2 writing
3. 学会等名 2018 Korea TESOL international conference 2018 (KOTESOL), Sookmyung Women's University, Seoul, Korea (国際学会)
4. 発表年 2018年

1. 発表者名 Cynthia Quinn
2. 発表標題 Correcting lexical errors with corpus data
3. 学会等名 The 44th Japan Association of Language Teachers (JALT) national conference, Shizuoka Convention & Arts Center, Shizuoka, Japan (国際学会)
4. 発表年 2018年

〔図書〕 計1件

1. 著者名 Cynthia Crosby Quinn	4. 発行年 2020年
2. 出版社 University of Birmingham, PhD thesis	5. 総ページ数 248
3. 書名 L2 writers referencing corpora to address accuracy: a qualitative analysis of learners' lexicogrammatical error corrections	

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------