Asia Pacific Journal of Corpus Research
Vol. 4, No. 1, 61-71

http://www.icr.or.kr/apjcr
e-ISSN 2733-8096
https://doi.org/10.22925/apjcr.2023.4.1.61

Effects of Corpus Use on Error Identification in L2 Writing

Yoshiho Satake
(Aoyama Gakuin University)

Satake, Y. (2023). Effects of corpus use on error identification in L2 writing. *Asia Pacific Journal of Corpus Research, 4*(1), 61-71.

This study examines the effects of data-driven learning (DDL)—an approach employing corpora for inductive language pattern learning—on error identification in second language (L2) writing. The data consists of error identification instances from fifty-five participants, compared across different reference materials: the Corpus of Contemporary American English (COCA), dictionaries, and no use of reference materials. There are three significant findings. First, the use of COCA effectively identified collocational and form-related errors due to inductive inference drawn from multiple example sentences. Secondly, dictionaries were beneficial for identifying lexical errors, where providing meaning information was helpful. Finally, the participants often employed a strategic approach, identifying many simple errors without reference materials. However, while maximizing error identification, this strategy also led to mislabeling correct expressions as errors. The author has concluded that the strategic selection of reference materials can significantly enhance the effectiveness of error identification in L2 writing. The use of a corpus offers advantages such as easy access to target phrases and frequency information—features especially useful given that most errors were collocational and form-related. The findings suggest that teachers should guide learners to effectively use appropriate reference materials to identify errors based on error types.

Keywords: Data-Driven Learning (DDL), Corpus Linguistics, Error Identification, Essay Writing, English Language Teaching

1. Introduction

Since computer technology and the Internet have become universally accessible, the application of corpora in second language (L2) learning and instruction has gained considerable attention. Data-driven learning (DDL) (Johns, 1991) is an approach that involves immersing learners in authentic language examples in corpora. Learners discover language patterns and establish generalizations through repeated exposure to these examples.

While there has been a recent surge in DDL research, the specific effects of corpus use remain somewhat unclear. For instance, while studies have shown that corpus consultation facilitates accurate correction of L2 errors (e.g., Satake, 2020; Satake, 2022; Tono et al., 2014; Crosthwaite, 2017; Yoon & Jo, 2014), there is a dearth of research focusing on the effects of corpus use on error identification. More studies examining the effects of DDL are needed to explore efficient and confident corpus use for language instruction in L2 classrooms.

This study aims to investigate the effects of corpus use on error identification when learners correct their errors in their L2 writing by referring to corpus data and clarify how corpus consultation contributes to accurate L2 error identification.

The research question of this study is ‘What effects would corpus use have on error identification?’ This study seeks to understand how corpus consultation influences error identification in L2 writing, considering the error types for which corpus consultation is either effective or ineffective.

2. Literature Review

While error identification plays a vital role in error analysis (Ellis, 2008), many studies tend to oversimplify its complexity, assuming it to be straightforward and uncontested (Hamid & Linh, 2014; Taylor, 1986). This perspective, particularly when applied to written errors, may not yield effective feedback (Hamid & Linh, 2014), given that successful error analysis depends on accurate error identification (Bartholomae, 1980).

A critical consideration is how we define errors. Ellis (2008, pp. 47–49) cites four challenges in error identification: (1) deciding whether grammaticality (well-formedness) or acceptability should serve as the benchmark, (2) distinguishing between ERRORS and MISTAKES, (3) determining if the error is overt or covert, and (4) considering whether infelicitous uses of L2 should be classified as errors. Concerning (1), an utterance may be grammatically correct but pragmatically unacceptable. Regarding (2), an error results from a lack of knowledge leading to a standard deviation, whereas a mistake is a performance failure. In terms of (3), an overt error deviates from a standard, while a covert error is a correct form that fails to convey the learner's intention. For (4), learners occasionally use expressions that, while grammatical, are not preferred by native speakers.

Burt (1975) further divides errors into global and local errors. Global errors, such as word order mistakes, affect sentence structure and can hinder understanding, while local errors, like inflection mistakes, impact a single sentence element and do not hinder understanding.

Errors can also be classified as interlingual or intralingual (Richards, 1971; Saville-Troike & Barto, 2016; Schachter & Celce-Murcia, 1977). Interlingual errors arise due to the influence of another language, often the learners' L1, while intralingual errors reflect learning processes such as overgeneralization.

Moreover, Khalil (1985) emphasizes the importance of error gravity, which relates to the seriousness of errors. Error gravity is determined by intelligibility (degree of comprehensibility), acceptability (level of acceptableness), and irritation (relating to the receiver's emotional response and error frequency).

Given the multitude of error classifications, providing a universal definition of errors proves challenging. Therefore, researchers need to articulate an appropriate error definition suited to their respective studies.

From the perspective of error identification, which is the focus of this study, it is crucial to note that some errors can be recognized by learners while others cannot. This dichotomy corresponds with the second issue Ellis (2008) highlighted concerning the difficulties in defining errors. Corder (1967) categorizes these as non-systematic mistakes made when language users incorrectly utilize an available language system and systematic errors made due to insufficient knowledge. Learners tend to repeat the latter systematic errors as they do not perceive them as errors. Thus, learners can identify mistakes with adequate grammatical or lexical knowledge about them and vice versa.

Despite various studies on error identification, there has been insufficient research on error identification in L2 writing from a DDL perspective, which is the focus of this study. More research is required to assess the effects of the DDL approach on error identification in L2 writing. Therefore, this study aims to examine the effects of corpus use on error identification in L2 writing.

3. Methods

3.1. Participants

This study included 55 Japanese university students in the department of history from the author's mandatory English writing class. Twenty-nine students participated over nine weeks in 2014, while 26 were engaged for 11 weeks in 2015. They were first-year students, ranging in age from 18 to 20, who

had received a total of six years of English education, three years in junior high school and three years in high school. The participants' average score on the EIKEN Test in Practical English Proficiency (Grade Pre-2 or Grade 2) was equivalent to A2 to B1 of the Common European Framework of Reference for Languages (CEFR).

3.2. Data Comparison and Analysis

To evaluate the participants' error identifications, the author compared corpus use with dictionary use and no use of reference resources. The comparison across these reference resources encompassed the quantity and accuracy of error identifications, and the error types the participants identified. The data from both 2014 and 2015 were consolidated for a more comprehensive analysis.

3.3. Instruments

3.3.1. Essay and Revision Tasks

Two tasks were used in this research: essay writing and revision tasks. Initially, the participants were required to complete timed essay writing tasks (25 minutes) without access to any reference resources. The participants wrote essays based on various topics from the class textbook, such as "nonverbal behavior" and "crime and punishment" (Pike-Baky & Blass, 2007, pp. 162, 186). Subsequently, the participants were given timed revision tasks (15 minutes), where they were prompted to correct their identified errors using a corpus and dictionary at least once each. After correcting the highlighted errors, the participants were encouraged to identify and correct their own errors if time permitted. The analysis of the revision tasks was previously conducted (Satake, 2020) and is not addressed in this study.

3.3.2. Error Identification

The author and the participants alternately carried out the identification of errors on a weekly basis. However, this study focuses solely on the participants' error identification to investigate how learners identify errors. The participants were required to use both a corpus and a dictionary at least once, and were provided ten minutes biweekly to identify their classmates' errors. They were instructed to highlight as many problematic expressions as possible. Errors identified through corpus use, dictionary use, or without using any reference material were highlighted in red, blue, and green, respectively. This color-coding system facilitated the evaluation of the effects of different reference resources on error identification in L2 writing.

3.3.3. Reference Resource

The Corpus of Contemporary American English (COCA) was used as a reference corpus for its extensive, balanced collection of American English and user-friendly interface. The author gave the participants a 20-minute instruction on using the corpus for error identification and correction. The handout for the instruction included four example sentences with errors for the participants to correct using keyword-in-context (KWIC) lines, and the participants were encouraged to tackle various error types to familiarize themselves with the search and analysis process.

The author permitted the students to use any dictionary they preferred, considering the impracticality of imposing a single dictionary due to the variety they possessed. Approximately two-thirds used the Genius English-Japanese Dictionary (Konishi & Minamide, 2006) with around 96,000 words.

3.4. Procedure

The following procedure was implemented for this study:

1. In-class timed essay task (25 minutes).
2. COCA instruction (20 minutes before the first revision task).
3. Error identification (10 minutes spent by the participants biweekly. The author and the participants alternated weekly).
4. In-class revision task (15 minutes). The participants were urged to identify their own errors once they had completed correcting the highlighted errors and had spare time.

In 2014, steps 1 and 4 were conducted nine times, and step 3 was biweekly for a total of four times. In 2015, this increased to 11 iterations of steps 1 and 4 and 5 biweekly instances of step 3.

5. Creation of an error-annotated learner corpus of the participants' essays.
6. Error analysis.

The participants consulted COCA by searching for problematic phrases and reading them in KWIC lines. They used dictionaries by searching for words within problematic phrases and reading their definitions and usage examples. Although students could freely consult the corpus, dictionaries, or neither, they were prohibited from using both the corpus and dictionaries to identify the same error to maintain research validity. If there was remaining time, the participants were encouraged to identify and correct any additional errors using the corpus, dictionaries, or neither.

The author compiled an error-annotated learner corpus consisting of 456 essays in total (66,251 words). Error tags were manually attached to the corrected sections based on the information provided in the revision sheets. These tags contained information such as parts of speech and error types. The error tags were adapted from the NICT Japanese Learner English (JLE) Corpus (Izumi et al., 2004), with modifications made to distinguish omission and addition errors, which are different types of errors (Lennon, 1991), common amongst Japanese learners of English (e.g., Kaneko, 2007): The error tag [oms] was used for omission errors, and [add] for addition errors. Word combination errors were classified as collocation errors. Furthermore, the author created the tag [c] for correct expressions revised by the participants, even though no corrections were needed. The final error analysis was performed based on the information from error tags.

4. Results

Table 1 presents the frequency of each type of error identified by the participants. The data from 2014 and 2015 were integrated to provide a comprehensive analysis of the error types.

Table 1. Error Identification by Reference Material

Error types	Corpus	Dictionary	No resource	Total
lexical	31	19	46	96
omission	34	8	27	69
number	11	4	26	41
agreement	8	2	20	30
addition	14	3	10	27
tense	3	1	23	27
verb form	2	3	7	12
part of speech	5	5	1	11
spelling	2	2	5	9
voice	3	1	3	7

colocation	5	0	2	7
case	1	1	3	5
order	2	1	2	5
punctuation	2	0	3	5
unintelligible	1	1	2	4
capitalization	1	0	3	4
inflection	1	1	1	3
negative verb	1	0	2	3
complement	0	2	1	3
Japanese	0	1	1	2
(in)finite	1	0	0	1
Total	128	55	188	371

The participants identified 371 errors in total. The majority of errors (188 errors, 50.67 percent) were identified without any reference material, followed by corpus use (128 errors, 34.50 percent), and the fewest were identified with dictionary use (55 errors, 14.82 percent). This distribution mirrors the pattern observed when the participants corrected their errors (Satake, 2020), suggesting that the choice of reference materials was consistent across both activities. Thus, it is reasonable to state that the nature of the error significantly influenced their selection of reference materials.

As displayed in Table 2, the participants identified 134 correct expressions as errors, accounting for more than a quarter of their 505 identifications (i.e., 371+134=505, see Table 1). They highlighted the highest number of correct expressions (69 expressions, or 51.49 percent) without any reference material. This was followed by 49 correct expressions (36.57 percent) identified with corpus use, and the lowest number of correct expressions, 16 (11.94 percent), identified using a dictionary.

Table 2. Correct Expressions That the Participants Inaccurately Identified as Errors

Part of speech	Corpus	Dictionary	No resource	Total
multi-word expressions	14	5	18	37
verb	13	0	12	25
noun	3	4	13	20
adjective	5	2	3	10
article	8	0	2	10
adverb	1	3	5	9
conjunction	2	2	4	8
preposition	2	0	5	7
auxiliary verb	1	0	4	5
pronoun	0	0	3	3
Total	49	16	69	134

Given that this ordering mirrors that observed when the participants corrected their errors (Satake, 2020), it is plausible that the choice of reference materials when misidentifying correct expressions as errors was primarily influenced by the types of expressions. This assumption aligns with the above analysis that the error types were a primary factor in choosing reference materials for error correction and identification.

Correct expressions are not errors and cannot be evaluated based on error types, so they were examined according to part of speech. The parts of speech that the participants misidentified as errors more than ten times include multi-word expressions, verbs, nouns, adjectives, and articles. The highest number of correct multi-word expressions and nouns were likely misidentified due to the lack of reference material, leading the participants to overlook their correctness. Correct expressions of verbs, adjectives, and articles were most often misidentified with corpus use. The least number of correct

expressions were inaccurately identified with dictionary use, possibly because the participants utilized dictionaries the least when trying to identify their classmates' errors (see Table 1).

Therefore, it is reasonable to state that the types of errors and expressions significantly influenced the choice of reference materials when the participants identified errors or misidentified correct expressions. The following section discusses the relationships between their choice of reference materials and the types of errors and expressions.

5. Discussion

5.1. Effects of Corpus Use on Error Identification

Regarding the effectiveness of corpus use for error identification, it can be inferred that the participants deemed the corpus more useful than dictionaries, considering that they identified more than twice as many errors using the corpus as compared to dictionary use. The cause of this difference may be associated with the error types that the participants recognized using the corpus or dictionaries. The following discussion, therefore, focuses on the types of errors they identified.

The participants highlighted the most significant number of omission and addition errors through corpus use. Below are examples of omission errors identified by the participants. The underlined words represent sections that the participants highlighted as errors:

(1) Iphone is <u>first smart phone</u> without key board.

(2) I think she is very powerful and <u>kind</u> someone.

For instances (1) and (2), which are examples of identified omission errors, it is reasonable to assume that these could be easily identified using the corpus because the missing words could be conveniently located. For example, in case (1), an article omission error, one could easily induce that "the" should precede "first" after a quick search for "first," as numerous example sentences include the phrase "the first." Likewise, in the case of (2), a preposition omission error, the participant could easily discern that "to" should follow "kind" after searching for "kind" as an adjective, as many examples include the phrase "kind to a person or people."

The participants also identified several examples of addition errors. Instances (3) and (4) are cases of such errors:

(3) So I <u>appreciate for</u> my mother.

(4) Therefore my father helps them their agriculture <u>on every weekend</u>.

Instances (3) and (4), which are examples of identified addition errors, could have been easily identified as errors using the corpus because the superfluous words could be conveniently located and removed. For instance, in the case of (3), a preposition addition error, one can quickly ascertain that "for" is redundant after "appreciate" once it is searched on the corpus since many example sentences include the phrase "appreciate something." Similarly, in the case of (4), another preposition addition error, one can easily discern that "on" is not needed before "every weekend" after searching for "weekend" or "every weekend," as many example sentences do not include the preposition "on."

Given that omission and addition errors ranked among the top ten error types corrected using the corpus and considering their high correction accuracy, it is plausible that the user-friendly nature of the corpus, which provides easy access to exact target phrases (Satake, 2020), contributed to the

identification of these errors. Hence, corpus use likely assisted the participants in recognizing omission and addition errors, as they could refer to examples of exact target phrases, look up the co-occurrence frequencies of words, and use this information to identify classmates' errors, much like they did when correcting their own omission errors and word order errors (Satake, 2020).

Furthermore, it can be assumed that this easy access to target phrases was beneficial for identifying other types of errors since the participants could conclude that a phrase was incorrect if they failed to find it in the Corpus of Contemporary American English (COCA) during their search.

Except for lexical errors, most errors are not tied to the meaning of the target word(s) but to collocations, which refer to lexical and grammatical items that frequently co-occur (Firth, 1957). Out of the top five errors corrected using the corpus, omission, and addition errors are associated with collocations. In contrast, number and agreement errors relate to the form of the target word(s). The author has already discussed collocation-related errors, such as omission errors. Let us now consider examples of form-related number errors that the participants identified:

(5) Let us compare Japanese great leaders, Nobusuke Kishi and Eisaku Satoh they are brother and prime minister in japan.

(6) He painted many unique pictures and influenced a lot of artist all over the world

In case (5), it would have been straightforward to spot "brother" as an error once the phrase was searched in COCA, as only one example of "they are brother" exists, which forms part of "they are brother and sister." As for (6), it would have been easy to recognize "artist" as an error once the phrase was searched because it quickly became evident that "a lot of artist" is not included in COCA.

These findings indicate that COCA's easy access to target phrases and frequency information proved effective for identifying both collocation errors and form-related errors, which account for most errors. As the corpus demonstrated efficacy for most errors, the participants may have decided to use the corpus more strategically and frequently than dictionaries.

However, the participants also identified the most significant number of correct expressions of verbs, adjectives, and articles using the corpus. Below are examples of correct expressions that the participants incorrectly identified as errors:

(7) After the hard work, she makes supper, washes dishes, cleans the rooms, and more.

(8) That is why in order to smoothly manage job and get a profit, . . .

In the case of (7), the author speculates that the participant deemed "makes" to be incorrect because she failed to find the phrase "make supper" during her search for the word "supper." Even though COCA includes the phrase "make supper," the participant might have come across other phrases such as "cook supper" or "have supper" more frequently, leading her to question the correctness of "make supper. " Similarly, for (8), the participant likely judged "get a profit" to be incorrect because this phrase did not appear in the concordance lines when he searched for "profit." Having found more instances of "make a profit" and none of "get a profit," he assumed "get a profit" was incorrect.

Thus, the participants misjudged correct expressions as incorrect when they used the corpus and failed to locate target phrases easily. Learners need to understand that expressions not found in the concordance lines are not necessarily incorrect.

In summary, the corpus's easy access to target phrases and frequency information played a significant role in error identification, as most errors were collocational, such as omission and addition errors, and form-related, like number and agreement errors. These errors could be effectively addressed through inductive inference based on example sentences of the target phrases. However,

educational interventions are required to prevent learners from mistakenly identifying correct expressions as errors, as the participants often considered a phrase incorrect when they could not easily locate it.

5.2. Effects of Dictionary Use on Error Identification

Regarding the utility of dictionaries, the participants identified nineteen lexical errors through their usage. Lexical errors were unique because they were the only error type where the participants identified more than nine instances using dictionaries. As lexical errors were part of the top ten error types corrected through dictionary use, it can be assumed that the primary strength of dictionaries, which is providing information on the meanings of target phrases, facilitated their identification, just as it aided in the correction of their lexical errors (Satake, 2020). Here are examples of lexical errors that the participants identified using dictionaries, with the underlined words representing the sections they flagged:

(9) He is left style, I have never seen his lose in armletheling.

(10) So many people worried by stomach cancer.

In example (9), identifying “left style” as an error would have been straightforward once dictionaries were consulted. This phrase is unlikely to be found in dictionaries, whereas correct expressions such as “left-handed” or “a left-hander” would be present. As for (10), the participant, upon looking up “worry,” could find various combinations of “worry” with a preposition, indicating that the original expression was incorrect. Compared to (9), the case of (10) appears more challenging, considering that the phrase “worried by” is included in dictionaries. To determine whether the expression was suitable for the given context, the participant would need to analyze several example sentences. Dictionaries’ strength, which lies in providing meaning information, could have assisted the participant during the examination of these sentences. To use this meaning information effectively for accurate error identification, the learner would require some knowledge about the usage of “worry.” Therefore, we can assume that identifying the error in (10) required a deeper vocabulary understanding than in (9).

However, the participants highlighted more lexical errors using the corpus and without reference material, suggesting that dictionary use may not be more effective for identifying lexical errors than either the corpus or no reference material.

The participants found dictionaries less beneficial than the corpus for error identification, as they detected less than half the errors with dictionary use compared to corpus use. This might be because most of the errors were not connected to the meaning of the target word(s), as previously discussed (see 5.1), meaning information in dictionaries often did not assist the participants in identifying their classmates’ errors.

In summary, the meaning information provided by dictionaries likely contributed to identifying lexical errors. However, the participants seemingly identified significantly fewer errors using dictionaries than the corpus because most errors were unrelated to the meaning of the target word(s) (see 5.1).

5.3. Effects of No Use of Reference Materials on Error Identification

The participants identified the most lexical, number, tense, and agreement errors without using reference materials. Below are examples of errors that the participants identified without reference material, with underlined words indicating sections highlighted by the participants:

(11) When someone of my family's birth day comes, he sings a happy birth day songs.

(12) She heard about my worries and advises me.

(13) For instance, he usually carry out his task very hard and at any rate focus on it so much.

(14) My father's favorite subject is society.

Errors in examples (11), (12), and (13) are number, tense, and agreement errors, respectively. These errors could have been easily identified using basic grammatical knowledge as they are straightforward. For example (14), which is a lexical error, the participant would not have required any reference material to identify "society" as an error, assuming that he understood the appropriate use of "society" or the correct phrase (i.e., social studies). In all instances, knowledge of the vocabulary and grammar of the target word(s) was crucial for error identification. As the errors above seem easily correctable, it can be inferred that the participants strategically decided not to consult any reference materials and identified them directly, mirroring their strategic revision of errors without reference material (Satake, 2020).

However, the participants inaccurately identified many correct expressions when not using any reference material, due to their failure to look up these expressions and confirm their correctness (see 4). Thus, error identification without reference materials proved ineffective without comprehensive vocabulary and grammar knowledge of the target word(s).

The findings reveal that error identification without reference material accounted for most identified errors, as the participants strategically highlighted many simple and easily correctable mistakes. However, the success of this method relied heavily on good grammar and vocabulary knowledge. Using reference materials such as corpora or dictionaries is recommended for error identification involving grammar and vocabulary items where learners' knowledge is insufficient.

6. Conclusions

This study examined the impact of corpus use on error identification in L2 writing. The results indicate different effects on error identification depending on the reference materials used. The participants identified the highest number of errors without any reference material, followed by corpus use, with the smallest number identified using dictionaries.

Corpus use proved particularly effective in identifying collocational errors, including omission and addition errors, as well as form-related errors, such as number and agreement errors. This efficacy is attributed to the inductive inference derived from example sentences of the target phrases provided by the corpus. On the other hand, dictionaries were beneficial for identifying lexical errors, where meaning information was particularly helpful. The results also highlighted the participants' strategic approach, identifying many simple and easily correctable errors without using any reference material. This strategy maximized the number of errors identified in a given time frame. However, this method also resulted in the largest number of correct expressions identified as errors.

Corpus use showed distinct advantages over dictionary use in error identification. Its strengths lay in the easy access to target phrases and the availability of frequency information. Given that most errors were collocational and form-related, these features were highly beneficial.

The limitation of this study is that it lacks a control group that does not use any reference materials to identify errors, as the author was not teaching any classes suitable for comparison. The author would like to examine the effects in a setting that includes a control group in future research.

In conclusion, the strategic selection of reference materials based on the error type enhances the

effectiveness of error identification, underscoring the need for familiarity with various reference materials and understanding their distinct advantages. The pedagogical implication is that teachers should guide learners in effectively utilizing appropriate reference materials for error identification, including the strategic use of corpus, based on the type of errors they encounter in L2 writing.

Acknowledgments

This study was supported by JSPS KAKENHI Grant Number JP 16K02851.

References

Bartholomae, D. (1980). The study of error. *College Composition and Communication*, 31, 253–269.

Burt, M. (1975). Error analysis in the adult EFL classroom. *TESOL Quarterly, 9,* 53-63.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5,* 161-170.

Crosthwaite, P. (2017). Retesting the limits of data-driven learning: Feedback and error correction. *Computer Assisted Language Learning, 30*(6), 447-473.

Ellis, R. (2008). *The Study of Second Language Acquisition* (2 ed.). Oxford: Oxford University Press.

Hamid, M. O., & Linh, D. D. (2014). The problematic of second language errors. *ARAL, 37*(2), 123-144.

Izumi, E., Uchimoto, K., & Isahara, H. (2004). *Nihonjin 1200nin no Eigo Speaking Corpus* (The NICT JLE Corpus). Tokyo: Alc.

Johns, T. (1991). Should you be persuaded: two examples of data-driven learning. In Johns, T., & King, P. (Eds.), *Classroom Concordancing* (*English Language Research Journal 4,* 1-16). Birmingham: ELR.

Khalil, A. (1985). Communicative error evaluations: native speakers' evaluation and interpretation of written errors of Arab EFL learners. *TESOL Quarterly, 19,* 225-351.

Konishi, T., & Minamide, K. (Eds.). (2006). *Genius English-Japanese Dictionary* (4th ed.). Tokyo: Taishukan.

Lennon, P. (1991). Error: Some Problems of Definition, Identification, and Distinction. *Applied Linguistics, 12*(2), 180-196.

Pike-Baky, M., & Blass, L. (2007). *Mosaic 1 Writing Student Book* (Silver ed.). New York: McGraw-Hill

Richards, J. (1971). A non-contrastive approach to error analysis. *ELT Journal, 25,* 204-219.

Satake, Y. (2020). How error types affect the accuracy of L2 error correction with corpus use. *Journal of second language writing, 50,* 100757.

Satake, Y. (2022). Effects of data-driven learning in on-demand classes. *Journal of Corpus-based Lexicology Studies, 4,* 1-13.

Saville-Troike, M., & Barto, K. (2016). *Introducing Second Language Acquisition* (3rd ed.). Cambridge: Cambridge University Press.

Schachter, J., & Celce-Murcia, M. (1977). Some reservations concerning error analysis. *TESOL Quarterly*, *11,* 141-51.

Taylor, G. (1986). Errors and explanations. *Applied Linguistics*, *7*(2), 144-166.

Tono, Y., Satake, Y., & Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL*, *26*(2), 147-162.

Yoon, H., & Jo, J. (2014). Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing. *Language Learning & Technology, 181,* 96-117.

THE AUTHOR

Yoshiho Satake is an Associate Professor at Aoyama Gakuin University. She received her Ph.D. from Tokyo University of Foreign Studies in 2018 in corpus linguistics. Her principal research lies in the field of data-driven learning (DDL).

THE AUTHOR'S ADDRESS

First and Corresponding Author
Yoshiho Satake
Associate Professor
Aoyama Gakuin University
4-4-25 Shibuya, Shibuya-ku, Tokyo, 1508366, JAPAN
E-mail: t31330@aoyamagakuin.jp

Received: 21 May 2023
Received in Revised Form: 3 August 2023
Accepted: 12 August 2023