

令和元年6月17日現在

機関番号：13101

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K02878

研究課題名(和文) Data-Driven Learning for English Language Beginners: An Empirically-Grounded and Evidence-Based Investigation

研究課題名(英文) Data-Driven Learning for English Language Beginners: An Empirically-Grounded and Evidence-Based Investigation

研究代表者

G. S. Hadley (Hadley, Gregory)

新潟大学・人文社会科学系・教授

研究者番号：70311008

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：This project used English books known as 'Graded Readers' to create an corpus that could be used to enhance the vocabulary and grammar of beginning students, as well as improve their reading speeds within an Extensive Reading program. Reading speeds of the experimental groups improved significantly.

研究成果の学術的意義や社会的意義

This project has established the ways that beginning students of Extensive Reading can improve through DDL. We have discovered procedures that will help teachers to successfully use this approach in their classrooms, and we have contributed to improvement of English education in Japan.

研究成果の概要(英文)：This project used English books for beginners, known as 'Graded Readers', to create an corpus so that Data Driven Learning (DDL) could be used to enhance the vocabulary and grammar of beginning students, as well as improve their reading speeds within an Extensive Reading (ER) program. Multiple cycles of testing discovered that, in terms of lexicogrammatical proficiency, overall no statistical difference between the experimental and control groups could be found. However, the experimental groups significantly improved more than the control group in terms of reading speed and total words read.

研究分野：Corpus Linguistics

キーワード：Corpus Linguistics Data-Driven Learning Extensive Reading

様式 C-19、F-19-1、Z-19、CK-19 (共通)

1. 研究開始当初の背景

Data-Driven Learning (DDL) is a method of student-centered language learning. In contrast to direct grammar instruction, in DDL, learners explore grammar and vocabulary questions using a *corpus*, which is a large database of language. The corpus is searched with software called a *concordancer*. Through this data, students learn through the exposure to many examples of words in context. With teacher guidance, DDL helps learners gain new insights into how English works. In addition, another form of language learning, called Extensive Reading (ER), uses simplified books for students to read large amounts of English. It is a widely-used methodology in Japan, with many educators believing that it is efficacious for language learning.

Educational research has found DDL to be effective with advanced and intermediate learners (Granger et al., *Computer Learner Corpora*, John Benjamins, 2002; Sun and Wang, 'Concordancers in the EFL Classroom', *Language Learning and Technology*, 16:1, 2003), but using DDL with beginners has been difficult because of the advanced levels of language in the corpora (St. John, 'A Case for Using Parallel Corpus and Concordance for Beginners', *Language Learning and Technology*, 5:3, 2001; Hadley, G. 'Data-Driven Learning for Beginners: A Three-Year Study', 2005).

2. 研究の目的

We sought to combine DDL with ER in this project. Our idea was to investigate if creating a corpus from Graded Readers would help students in ER classes to gain better grammar, vocabulary, and reading speeds than lessons using ER alone. The advantages of creating materials informed by a corpus of graded readers would have certain advantages, as students would be certain to encounter new vocabulary or grammar later on when reading other books, since all teaching materials were created from the reading materials. Would such comprehensible and recyclable material unlock the full potential of DDL for beginners? It was within this background that we sought to find out if statistically significant English language improvement through a combination of DDL and ER was possible, and to construct a methodological framework that could guide teachers into how DDL could be successfully utilized in their classrooms.

3. 研究の方法

A Pretest-Posttest Design was conducted during each semesters of the academic year (see Figure 1). We divided classes into test and control groups. The research instrument was a C-Test (Klein-Braley and Raatz, 'A Survey of Research on the C-Test', *Language Testing*, 1, 1984) created from a higher-level Bookworms reader. All students were required to read a minimum of 200,000 words, as measured by the Extensive Reading Foundation's MReader Platform (<http://mreader.org/>). Control classes engaged in a

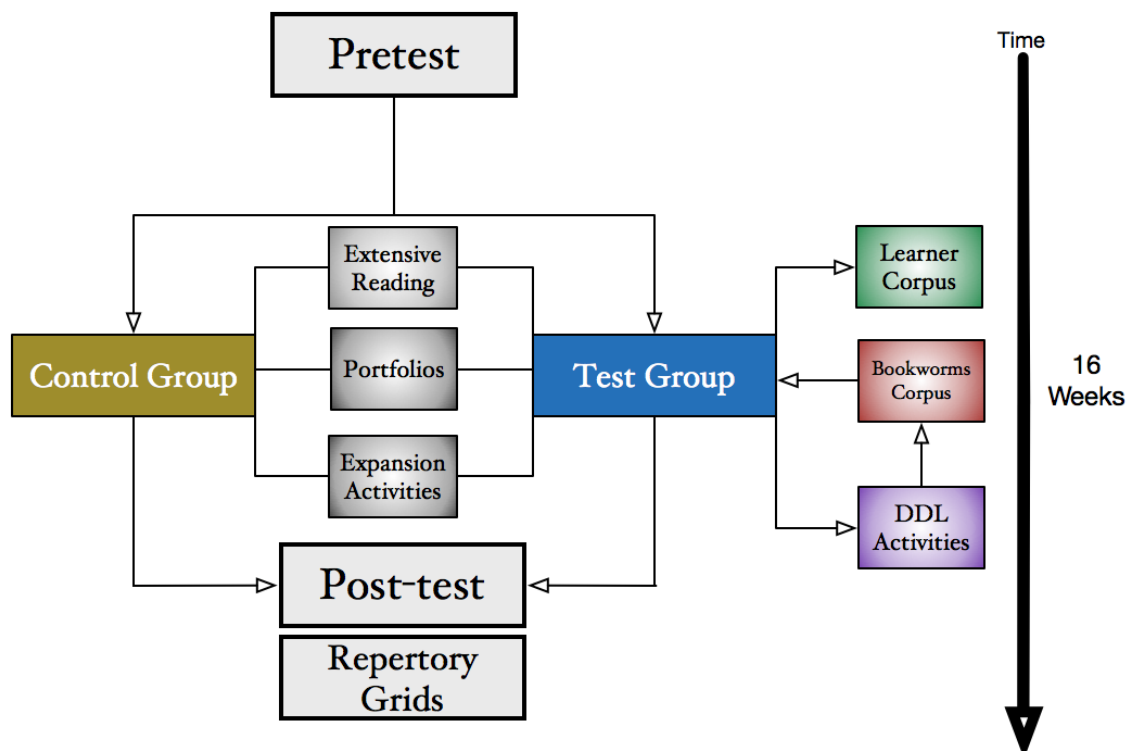


Figure 1

typical Extensive Reading class, with classroom activities from published resource books, portfolios as well as timed reading tests. The experimental group also experienced these features, but also used materials informed by corpus analysis, as well as corpus printouts and vocabulary frequency lists. Post-tests of all classes, as well as the administration of Repertory Grid questionnaires for eliciting student personal constructs about DDL, together with grounded theory-based interviews, were then used to discover the effects of the DDL treatment in the classes.

4. 研究成果

The early findings and affective response were decidedly mixed. Quantitatively the grammar and vocabulary (that is, *lexicogrammatical*) proficiency of both the experimental and control groups significantly improved over the semester, but the control group showed greater improvement in terms of vocabulary, grammatical knowledge, and reading speeds. Early qualitative data clearly indicated resistance on the part of the experimental group to the early DDL materials. Some recognized the materials as useful, but others were less certain about the purpose of the materials, and were uncertain of whether it was worth the investment of time needed to make them work. Students in the early part of this study expressed a preference for learning activities that were more in line with the classic ER approach—that is, simply reading and discussing their books with classmates.

This led us to change our approach and to embed corpus data and insight within materials which appeared to be more ‘textbook-like’. We created characters named Corpus Cat and Data Dog to aid with informative, cute, and fun activities that would supplement the extensive reading of later experimental classes over the next two years (see Figure 2). Students in subsequent classes were much more engaged and interactive as a result of these materials.

Later *t*-tests of dependent means of the post-tests for all experimental and control groups indicated that most students improved significantly by the end of term. A consolidated *t*-test of independent means for all the control and experimental groups allows for an effective overview of what took place during the five times that this experiment was

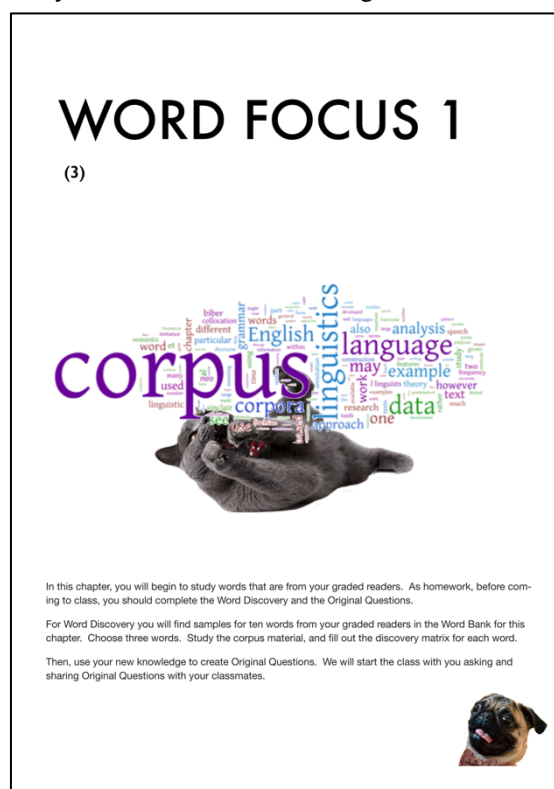


Figure 2

conducted in the following two years after the first year of experiments. Because the materials and mode of instruction for the control groups was unchanged for the entire three years, data from the first control group meeting the minimum word requirements was added to the following collated analysis. All of the data from the first experimental group, however, was excluded.

The experimental group was created from four classes (n=34). The control group (n=29) was also formed from four classes. This data was consolidated for the final presentation. The consolidated groups began their ER courses within the same statistical population, both in terms of lexicogrammatical proficiency ($t(61) = 0.724, p = 0.23; d = -0.18$), and with average reading speeds of about 133 words per minute ($t(61) = 1.273, p = 0.10; d = 0.34$), the critical value for *t* in all instances here being 1.67 ($p < .05$).

Post-test results over the successive periods yielded new surprises. In terms of lexicogrammatical proficiency, despite the new materials, no overall statistical difference between the experimental and control groups ($t(61) = 0.970, p < 0.16; d = -.28$) could be detected. However, the experimental group significantly improved more than the control group in terms of reading speed ($t(61) = 2.871, p < 0.05; d = 0.75$), and in total amount of words read (Table 1).

	<i>M</i>	<i>SD</i>
Experimental Group	314,433	180,469
Control Group	234,509	48331

Table 1 Mean and Standard Deviation of Word Amounts Read for Collated Experimental and Control Groups ($d = .59$)

Therefore, in terms of lexicogrammatical improvement, even though the enhanced materials did not prove to be superior to the materials used in the control group, neither were they inferior. Given that our first attempts resulted in materials were having a more negative effect, that the DDL materials had been improved to the point where they were having an effect that was equal to the effect of commercially published expansion material used with the control group.

The significant improvement of the experimental group's reading speeds and reading amounts through DDL is the most positive discovery from this research. More research is needed, but we hypothesize that repeated exposure to the most frequent lexical items, and becoming accustomed to skimming for patterns in the corpus readout sheets, both contributed to an improvement in reading speeds, which in turn resulted in the experimental group being able to read more than the control group. We feel the enhanced DDL materials contributed to improvements in the experimental group's reading speeds and reading amounts.

5. 主な発表論文等

[雑誌論文] (計 3件)

- ① ハドリー浩美 (2018). 「新潟大学全学英語教育カリキュラム改定の成果検証 (2)」、新潟大学言語文化研究、22 巻、25-32 ページ、査読無.
- ② Hadley, G. & Charles, M. (2017). 'Enhancing extensive reading with data-driven learning.', *Language Learning & Technology*, 21 (3), pp. 131-152, refereed.
- ③ ハドリー浩美 (2016). 「新潟大学全学英語教育カリキュラム改定の成果検証」、新潟大学高等教育研究、4 巻、1-9 ページ、査読有.

[学会発表] (計 12件)

- ① Hadley, G. (26 Nov, 2018). 'DDL and Extensive Reading: An Inclusive Approach,' JALT2018 Conference, Shizuoka.
- ② Hadley, G. (24 Nov, 2018). 'Textbooks in ELT: What are they Good for Anyway,' JALT2018 Conference, Shizuoka.
- ③ Hadley, G. (6 Jul, 2018). 'Investigating Data-Driven Learning with Receptive Skills: A Contemporary Approach,' Teaching and Language Corpora (TaLC), University of Cambridge, UK.
- ④ Hadley, G. (18 May, 2018). 'Data-Driven Insight or Information Overload? Researching Foreign and Second Language Learning with Grounded Theory,' Lynn Erler Lecture, University of Oxford, UK.
- ⑤ Hadley, G. (30 Sep, 2017). 'Investigating the Impact of Extensive Reading with Data-Driven Learning,' 英語コーパス学会 第 43 回大会, 関西学院大学.
- ⑥ Hadley, G. (16 May, 2017). 'Investigating the Impact of Data-Driven Learning with Extensive Reading,' Department of Education, University of Oxford.
- ⑦ Hadley, G. & Charles, M. (18 Mar, 2017). 'Data-Driven Learning with Extensive Reading: Language Gains and Affective Responses,' American Association of Applied Linguistics, Portland, USA.
- ⑧ Hadley, G. (28 Nov, 2016). 'The Impact of Data-Driven Learning in ER,' JALT 2016, Nagoya, Japan.
- ⑨ Hadley, G. (1 Oct, 2016). 'Data-Driven Learning with Extensive Reading: Investigating the Language Gains and Affective Responses of Beginners,' 9th Annual Extensive Reading Seminar, Nanzan University, Nagoya, Japan.
- ⑩ Hadley, G. & Charles, M. (21 Jul, 2016). 'Extensive Reading, False Beginners, and Vocabulary Development: Investigating the Impact of Data-Driven Learning,' Teaching and Language Corpora (TaLC), Giessen, Germany.
- ⑪ Hadley, G. & ハドリー浩美 (8 Jul, 2016). 'Linking Student Constructs to Educational Innovations: Dynamics within a Japanese Context,' The XIIIth European Personal Construct Association Conference, Galzignano Terme, Padova (Italy).
- ⑫ Hadley, G. (11 Jun, 2016). 'Extensive Reading, False Beginners, and Vocabulary Development: Investigating the Impact of Data-Driven Learning,' Northeast Asian Research Language Education Conference, Niigata, Japan.

[図書] (計 3件)

- ① Hadley, G. & ハドリー浩美 (In Press). 'Using Research to Inform Materials Development,' in J. Norton & H. Buchanan (eds.), *The Routledge Handbook of Materials Development for Language Teaching*, Routledge, Oxford, UK.

- ② Hadley, G. (2018). 'Learning Through Textbooks,' in A. Burns and J.C. Richards (eds.), *The Cambridge Guide to Learning English as a Second Language*, Cambridge University Press, New York, USA, pp. 298-306.
- ③ Hadley, G. (2017). *Grounded Theory in Applied Linguistics Research: A Practical Guide*, Routledge, Oxford, UK.

6. 研究組織

(1)研究分担者

研究分担者氏名：ハドリー 浩美

ローマ字氏名：Hadley, Hiromi

所属研究機関名：新潟大学

部局名：教育・学生支援機構

職名：准教授

研究者番号（8桁）：60534732

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。