

## 「論 文」

# A Meta-analysis of Data-driven Learning Approach in the Japanese EFL Classroom

Atsushi Mizumoto and Kiyomi Chujo

## Abstract

In this study, a meta-analysis was conducted targeting the studies that employ data-driven learning (DDL) approach in the Japanese EFL classroom context. After a thorough literature search, 32 effect sizes from 14 primary studies that took place in the Japanese EFL classroom were retrieved, coded, and calculated. The synthesized results, based on the classification of the outcome measures, showed that the DDL approach worked well particularly for learning vocabulary items (Level 1: lemma). It also worked positively for basic grammar items (Level 2: category) and noun and verb phrases (Level 3: phrase). For a proficiency measure, the combined effect size was small. Accordingly, the results of the current meta-analysis would provide further support for the use of DDL approach in the classroom, which could be an alternative methodology for facilitating the learning of lexico-grammatical items. Suggestions for further research and pedagogical implications are provided.

## 1. Introduction

The development of corpus linguistics as a discipline, especially since the end of the 20th century, has had a tremendous influence on the field of applied linguistics (Hunston, 2002). In particular, applied domains of corpus linguistics such as lexicography, pedagogic grammar (e.g., Biber, Johansson, Leech, Conrad, & Finnegan, 1999), phraseology, and discourse analysis have benefited significantly from the very large corpora (Myles & Mitchell, 2004). Development of these applied domains within the field of corpus linguistics in turn has affected other areas in applied linguistics in general. Consequently, almost all introductory books on second language acquisition (SLA), language teaching, and language testing include sections on corpus linguistics or its applied domains (e.g., Loewen & Reinders, 2011; Long & Doughty, 2011; Mackey & Gass, 2012; Shohamy & Hornberger, 2008). For more specific pedagogical purposes, teaching and materi-

本論文は、石田知美・杉浦正利（2012）「日本人英語学習者による連語表現の言語的特徴—判別分析を活用して」『英語コーパス研究』第19号：1-14. の訂正論文である\*。

## 「訂正論文」

# 日本人英語学習者による連語表現の言語的特徴 —判別分析を活用して

石田 知美・杉浦 正利

## Abstract

The purpose of this study is to investigate the characteristics of repetitive word combinations in the written English of Japanese EFL learners compared to those characteristics in the written English of native speakers. We focused on types, tokens and the Guiraud index (GI) of the repetitive combinations, and implemented a quantitative comparison and discriminant analysis.

The quantitative comparison of the data revealed that the number of repetitive word combination types, except for bi-gram types, and tokens in the learner corpus was significantly higher than in the native corpus. The GI was higher in most of the learner corpus data as well. The analysis also found that learners used longer word combinations compared to native writers. The result of the discriminant analysis showed that tokens of 2-word and 3-word combinations were always selected as one of the linguistic features to distinguish between learners' essays and those of native speakers. Qualitative observations of the 4-word to 6-word combinations of EFL learners compared to sentences paraphrased by a native speaker showed an insufficient awareness of register and a lack of variation in expressions among the learners.

## 1. はじめに

英語学習者の中間言語の特徴や「不自然さ」は、連語表現にも表れる。Allerton (1984: 39) は、文法的にも意味的にも問題は無いが、母語話者が使用することのない表現を、学習者は用いると述べている。また、学習者の連語表現に関してではないが、Pawley and Syder (1983: 193) は、文法的に正しい表現は色々あ

## 「研究ノート」

# 学習支援用日英例文パラレルコーパス SCoRE の構築における課題—日本語対訳例文の訳出に焦点を当てて

若松 弘子・石井 卓巳・中條 清美

## Abstract

The Sentence Corpus of Remedial English (SCoRE) is a collection of level-appropriate example sentences which was created to provide freely accessible grammatically categorized materials for effective data-driven learning in Japanese EFL classrooms. To benefit the target population, the example sentences were prepared by native speakers of English who included most relevant grammatical items and verified appropriate vocabulary level and sentence length. Each example sentence has a Japanese translation so that students are able to focus on grammatical structure and do not struggle with meaning. The English to Japanese translation initially aimed at realizing an easily-read direct equivalent; however, this occasionally proved to be challenging because of the contrastive differences between English and Japanese (e.g., cognitive semantic and structural features). For instance, *come* and *go*, when used as verbs of motion, are not always semantically equivalent between English and Japanese. In addition, the meaning of *you* in English represents *what you say*, as in *I understand you*, but this does not hold in Japanese. In this paper, these types of examples are presented and discussed.

## 1. はじめに

学習者に言語使用の用例を提示し、言語の規則性やパターンなどを探らせる指導法は、80年代半ば、実践者である Tim Johns によってデータ駆動型学習 (data-driven learning: DDL) と名付けられて注目を浴び (Johns, 1991), 以降、教育現場での DDL の普及やその効果が期待されてきた。それから30年ほど経った現在、日本の教育現場における DDL は当初の期待ほどに普及しているとは言い難い。その理由として、日本の英語学習者の大半にとって、既存のコーパスに含まれる英文の難易度が高すぎる事が挙げられる。効果的な語学学習には学習者のレベ

## 「講 演」

# タスク駆動型のコーパス構築と情報処理技術 —「分類」を活用した新技術開発

田中 省作

## Abstract

In the academic field of corpus studies, techniques for information processing, such as language processing, are commonly not regarded as tools for construction of corpora, although they are actually also useful for it. Since recent corpus studies require more sophisticated and complicated linguistic features, such as semantic information, it is becoming more important for linguistic and informatics studies to effectively construct new kinds of corpora using information processing techniques. This paper presents two case studies that build linguistic data as a corpus in the broad sense in parallel with a new information technique based on “classification.” The first develops a technique for estimating the quality of the linguistic features of scientific papers and constructs a corpus of scientific papers with such information from the Web. The second develops a technique for analyzing a sentence based on the school grammar of English and constructs a corpus of sentences with such grammatical information.

## 1. はじめに

コーパス研究において、言語処理などの情報処理技術は、一部の言語情報の付与をのぞき、コーパスの分析手段として位置づけられることが多い。一方、情報学研究では、研究課題を解決する過程で、新たな言語データの構築が必要となることが少なくない。その効率的整備に情報処理技術が活用されることもある。近年、コーパス研究でも、研究対象となる言語現象や取り扱うべき言語素性が多様化、複雑化しつつあり、このような情報処理技術を活用したコーパス構築の効率化は、ますます重要となることは間違いない。本稿では、情報学研究における新技術開発とそれに資するデータ構築を共進した2つの研究事例を示す。これらの事例は、著者らの比較的小規模な研究組織による、「分類」に基づいた新技術開

## 「シンポジウム」

### 英語教育・研究のための教材コーパスの構築と利用 —実践例と課題

石井 康毅

#### まえがき

学習者のアウトプットデータである学習者コーパスは近年ますます注目を集めている。対照的に、インプットデータである教材コーパスについては、著作権等の問題があってデータを公開できない場合が多いこともあり、個人的にデータを作成して利用するという域を超えるのが難しく、教材コーパスの構築と利用についてのノウハウは広く共有されているとはいいがたい。

しかし、教材コーパスやそこから作られるデータは、教材の特徴を質的・量的に分析するなどの研究目的以外に、教育目的でも活用が可能である。特に、英語教育における小中高大の連絡・連携が重要視されるようになってきた昨今、各学校段階で利用されている教材を分析し、その結果を踏まえて教育活動の改善につなげることは重要な課題である。

このような背景のもとに、2014年10月4日に熊本学園大学で開催された英語コーパス学会第40回大会において、表題のシンポジウム（司会：石井・藤原）が行われた。各講師が教科書を中心に、広く「教材」と呼べるもののコーパスを作り、利用するという統一テーマのもとで実践例を報告し、労力のかかる教材コーパス構築の意義と課題を共有することを目指した。なお、本シンポジウムは、教材コーパスの具体的な構築プロセスと活用を実際に体験することを目的として同日午前で開催されたワークショップ（講師：藤原・石井）と連動するよう企画された。

以下に掲載する4編の論文は、各講師のシンポジウムでの発表内容とその後の質疑応答・全体討論をもとに、加筆・修正したものである。

本シンポジウムの論文ならびにワークショップが、実際に教材コーパスを構築したり利用したりする際に参考になれば、企画した司会者・講師一同にとって幸いである。