

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 Japanese EFL learners' written English compared to those in native English writers. We focused on types, tokens and the Guiraud index (GI) of the recurrent word combinations, and implemented a computer-based quantitative comparison and discriminant analysis.

The quantitative comparison of data revealed that the number of recurrent word combination types and tokens in the learner corpus was significantly higher than in the native corpus, which indicated Japanese learners' tendency to use the same expressions repeatedly. The GI was higher in most of the non-native corpus data as well. The analysis also found learners used longer word combinations compared with native writers. The result of the discriminant analysis showed that tokens from 2-word combinations to 4-word combinations were always selected as one of the linguistic features to distinguish between learners and native writers. A closer look at 4-6 word combinations compared with sentences paraphrased by native writers showed insufficient awareness of register and a lack of natural expression in language used by learners.

1. はじめに

英語学習者の中間言語の特徴や学習者特有の「不自然さ」は、連語表現に顕著に表れる。Allerton (1984) は、学習者は文法的には問題のない表現ではあるが、母語話者が決して使用することのない表現をしばしば用いると指摘している。また学習者の表現に関してではないが、Pawley and Syder (1983) は、文法的に正しい表現は色々あり得るが、その中でも母語話者らしい表現はごく一部であると述べている。母語話者と学習者の使用する言語的特徴の相違は、個々の単語に着目するだけでは不十分であり、連語表現の相違を比較分析することによって、より明らかになるといえる。

本研究では上記のような問題意識に基づき、2語以上の連語表現を分析対象

アカデミック・ライティングにおける研究者のスタンス — 研究論文の Introduction における伝達動詞の時制の検証

中谷 安男

Abstract

This paper investigates the construction of stance in the introduction sections of research articles. Using a corpus-based approach, it examines the significance of the choice of tense in reporting verbs used to make references to specific research. The data are drawn from 102 representative articles in three disciplines: cultural science, social science and natural science. The results indicate several meaningful strategies for signaling the stance. The past tense is used when authors objectively report the methods or results of previous studies. The present perfect tense is used to generalize a research topic or establish a research niche. The present tense is used to support authors' stance by claiming universality and enhancing the generalizability of specific methods and findings.

1. はじめに

学術論文 (Academic Paper: AP) では、書き手には編集者や査読者を意識し、彼らを説得するための様々なアカデミック・ライティング・ストラテジー (Academic Writing Strategies: AWS) が要求される (例 Jordan, 1997; 中谷・土方・清水, 2010)。特に自分の主張や研究上の立場をわかりやすく伝えなければならない。一般に書き手の立場は「スタンス (Stance)」と定義され、自分の気持ちや、態度、価値判断や評価を表すものと考えられている (Biber et al., 1999: 966)。Charles (2006a, 2006b) は、特に AP の中で書き手が自分の研究の立場を読者に明確にするストラテジーをスタンスとしている。このスタンスは、言及した研究論文の内容や成果に賛同したり、否定的な見解を示したりする時にも使われる (Hyland, 1999)。本研究では、このようなスタンスの中でも、書き手が研究者の論文や成果に言及する「伝達動詞」(Reporting Verb) の選択 (例 Thomas & Hawes 1994; Hyland, 2004), 及び時制 (Tense) の使い分け (例 Malcolm, 1987; 中谷・清水, 2009) に注目し AWS の活用を考察する。

このようなストラテジーが最も顕著なのは、投稿された論文の採択に大きく影響を与えると考えられている Introduction の章である (Swales, 1990, 2004)。

研究ノート

コーパス構築と著作権 — Web を源とした質情報付き英語科学論文コーパス

田中 省作・安東奈穂子・富浦 洋一

Abstract

Copyright processing is one of the important requirements in building a corpus. However, under the copyright law before 2010 in Japan, even a nonprofit group cannot build a corpus for research purposes without obtaining licenses from the authors of the documents compiled into it. The amended copyright law was enacted in 2009 and enforced in 2010. In this amendment, the new article 47-7 was included in the law; this article presumably makes legal corpus building for some purposes such as corpus analysis. This paper discusses such legal problems, particularly copyright-related ones, in the case of building a corpus based on an actual project, which built a corpus of scientific papers on the Web with the quality of their linguistic features.

1. はじめに

今日、コーパス¹は言語研究に欠かすことのできない重要な研究資源の一つである。コーパスの構築には研究課題に応じたコーパス設計、データの収集、電子化作業など一般に大変な労力を要する。さらには、このような技術的问题だけではなく社会的な問題もある。その一つが著作権である。著作物性が認められるもの（著作物）²については、憲法の条文等の特殊なものをのぞき（第13条）、日本ではそれらが作成された時点で著作権が自動的に発生する。したがって、コーパス構築に際し、それに利用される文書等の著作物については、著作権が切れた、もしくは放棄されたものを除き、原則、利用の許諾を権利の保有者（著作権者）から得なければならないことになる。しかし、コーパスは一般に非常に多くの著作物で構成され、このような著作権処理が現実的には大きな問題となる（前川, 2010）。

現在、著者らが推進している科学研究費補助金・研究課題「Web 上からの母語話者 / 非母語話者英語論文コーパスの作成・公開とその利用」（研究代表者：富浦洋一、研究種目：基盤研究（B）、研究期間：2008-2011年度）（以後、プロジェクトと略記する）は、コーパスを実際に構築することを目的としており、