様式Ｃ－１９

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 6 月 11 日現在

機関番号：15401
研究種目：若手研究(B)
研究期間：2010～2011
課題番号：22720213
研究課題名（和文）日本人英語学習者のコロケーション知識の保持と運用に関する基礎的研究
研究課題名（英文）
A Fundamental Study on Retention and Performance of English L2 Learners' Collocational Knowledge
研究代表者
阪上　辰也（SAKAUE TATSUYA）
広島大学・外国語教育研究センター・特任講師
研究者番号：60512621

研究成果の概要（和文）：
本研究は，英語学習者のコロケーション知識がどのように保持され，また運用されているかを明らかにすることである。平成 22 年度には，ライティングの過程を記録できるシステムを開発・導入し，平成 22 年度後半から平成 23 年度前半にかけて，1 時間程度で産出されたライティングのデータを収集した。データを産出過程・産出時間・産出された表現といった観点から分析した結果，コロケーション知識をまとまりとして保持し運用していることが分かり，ひとまとまりとしてコロケーションを処理している可能性が示された。

研究成果の概要（英文）：
The purpose of this research was to reveal how English L2 learners retain and use collocational knowledge. In 2010, for this purpose, a new real-time logging system was developed to record the writing process on the scale of milliseconds. From late 2010 through early 2011, using the logging system, the participants were then asked to write an essay about school education in one hour and the data were collected. The analysis was conducted from three aspects: the production process, production time, and produced expressions. The results confirm that L2 learners have and use collocational knowledge and indicate the collocations they used were processed as single whole units.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
2010 年度	800,000	240,000	1,040,000
2011 年度	500,000	150,000	650,000
年度			
年度			
年度			
総　計	1,300,000	390,000	1,690,000

研究分野：第二言語習得研究
科研費の分科・細目：言語学・外国語教育
キーワード：第二言語習得，コロケーション，学習者コーパス，ライティング

１．研究開始当初の背景

研究開始当初の背景として，従来のコーパスを用いた研究の問題点が 2 点挙げられる。１つは，直接実験によって産出の過程を探ることが困難であるという点であり，過程を記録した学習者コーパスは見られない。もう１

つの問題点は，コーパスを利用して抽出されたコロケーションが，認知的側面からの分析が行われていないという点にある。コーパスを用いての計量言語学的アプローチでは，コーパス中の「頻度」のみが分析の基準となり，頻度が高く，かつ，言語学的に意味のある表現のみがコロケーションとして認定される傾向がある。しかし，頻度のみを基準として認定された表現が，学習者の脳内に知識として記憶されているとまでは言い切れない。従来のコーパス利用による研究では，こうした点が十分に考慮されることなく論じられており，認知的側面も含めたコロケーション研究を行う必要がある。

こうした問題点に加え，日本人英語学習者が，単語という一語から成る語だけではなく，複数の語を組み合わせたコロケーションを産出していることが分かってきているが，こうした研究では，最終的に出来上がった英文のみを分析した結果であり，文を産出する過程を観察した上で，どのようなコロケーションが出現するのかは分析されていないという問題点も残っていた。

そこで，まず，英語学習者のライティングの過程をリアルタイムで記録できるシステムを開発し，英文の産出過程を含めたデータを収集する必要があると考えた。記録システムの開発により，ライティングの過程の記録と分析が可能となる。結果として，コーパスを用いた従来のコロケーション研究で重視されてきた計量的側面からだけでなく，認知的側面から見たコロケーション知識の保持と運用に関する研究が実現できると考えた。

２．研究の目的

本研究は，日本人英語学習者の産出単位を分析可能にする新たな学習者コーパスの構築し分析することにより，日本人英語学習者がコロケーション知識をどのように保持し，運用しているのかを明らかにすることを目的としている。

３．研究の方法

本研究の方法は，以下の３点に大きく分類することができる。

(1) ライティング過程記録システム開発

まず，ライティング過程のリアルタイム記録システムの開発を行った。ライティングシステムは，２段階の入力をする必要があり，最初の段階では，入力速度を求めるために，表示された英文をできるだけ速く入力する。次の段階では，英作文のテーマが提示され，学習者は英文を入力する。いずれの段階でも，一文字が入力されるごとに，その入力とその際の時間をミリ秒単位で記録できるようになっている（図１）。

図１．記録システムの画面

(2) ライティングの実験実施

記録システムが完成した後に，作文テーマの検討を行い，世界的に広く利用されている学習者コーパスである ICLE（International Corpus of Learner English）のテキスト分析を行うことで，実験参加者が無理なく書ける課題として，「学校教育」というテーマを選定した。

テーマの選定後に，予備実験を行い，データの記録に問題がないかをチェックした。実際には，初期型の記録システムには，正確な時間を記録できない場合があるという不具合が発見され，最終的には，HSP（Hot Soup Processor）という言語を用いてシステムの改修を行った。

システムの改修後に，本実験を行い，大学学部生 17 名に対して，１時間のライティング課題を与えて，英作文データおよびその過程データの収集を行った。

(3)データ分析

データの分析は，以下の手順で行った。

①ライティングタスクで記録されたデータから，各個人の一語を産出するのに要した速度（産出速度）を求める。

②産出速度を基に，ポーズ時間を求め，ポーズが入るまでの語句を認知的側面から見たコロケーションとして抽出する。

③従来のコーパスによるコロケーション研究で用いられてきた手法のひとつである「N-gram モデル」を利用し，計量的側面から 2-gram から 4-gram までのコロケーションを抽出する。

④N-gram モデルにより得られたコロケーションと，産出過程の分析から得られたコロケーションを比較し，共通する点と異なる点を記述する。

４．研究成果

本研究の成果として，以下の３点を挙げる。

１点目は，「全く新しい学習者コーパスが構築された」という点である。これまでにも，産出結果を基にした学習者コーパスは構築されているものの，ライティングの過程を含めた学習者コーパスはほぼ皆無であり，本研究により収集されたデータを基にしたコーパスが，既存のコーパスとは全く異なる種類の言語習得研究用のデータベースとなるであろう。

２点目として，「ライティングの産出過程の分析手法を開発できた」点である。これまでは，作文の過程を記録することさえも十分にできなかったことから，分析手法は，これまでにほとんど提案されていない。本研究を進めることで，過程の分析手法を追求することになり，結果として，新しい研究手法を提示できるようになった。

３点目に，日本人英語学習者がどのような単位でコロケーション知識を保持して運用しているのか」という研究課題に対して，計量的側面と認知的側面という両面からの結果を得られた点である。ライティングの産出過程を記録・分析することは，方法論上困難とされてきたわけであるが，その実験調査を実現することができた。

なお，実験の成果は，主に２つの国際会議で発表され，１つは，2011 年４月に，アメリカ合衆国アラバマ州で行われた国際シンポジウム「Exploring the Boundaries and Applications of CORPUS LINGUISTICS」にて，「A comparative study of multi-word units: Written results and writing process by English learners.」というタイトルで発表を行った。具体的には，学習者の産出した表現が，文頭の副詞と主語と動詞がひとまとまりで抽出されるなど，機械的に産出単位を抽出する N-gram モデルとは異なる産出単位・言語表現が観察されたことを報告した。

加えて，2011 年９月には，ベルギー王国で行われた「Learner Corpus Research 2011」にて，「A new learner corpus for SLA research: Dynamic Corpus of English Learners」というタイトルで新規に構築されたコーパスとその分析結果を報告した。具体的には，本研究では，17 名の日本人英語学習者を被験者として，産出データを収集し，産出過程のデータを動的コーパスとして分析し，コロケーションを抽出した。抽出手順は，1）個々の被験者のキー入力の時間から平均産出速度を求めて，2）内境界点を元に閾値を求め，3）コーパス中の閾値を超えた値をポーズとみなし，ポーズ間の表現をコロケーションとするというものである。本コーパスから得られたコロケーションの数は 90 個で，最終的な産出結果から得られる 2-gram 表現および 3-gram 表現を比較した結果，上位２種類の表現は，動的コーパスから得られた表現と一致していたが，その他の表現は異なっていた。それと同時に，動的コーパスから得られるコロケーションは，2-gram や 3-gram のものと同じであっても，その頻度が少なくなる傾向が見られた。つまり，最終的な産出結果から得られたコロケーションと，本研究で新たに構築した動的コーパスから得られたコロケーションの間には，その種類や頻度に違いが見られることが分かった。発表後，海外の研究者から，産出の時間や品詞の並びなどをさらに観察すると新たな発見があるだろうという助言を得ることができ，一定のインパクトを与えることができたと言える。

今後の展開としては，知識の保持や運用面だけでなく，コロケーションの「処理過程」を観察することで，日本人英語学習者はコロケーションをどのように処理しているのかを明らかにすることを目的とした新たな研究に発展させることができると考えている。

５．主な発表論文等

（研究代表者，研究分担者及び連携研究者には下線）

〔雑誌論文〕（計４件）

1. 阪上辰也 (2012). 「統計解析環境「R」を利用した言語データの処理」『外国語教育メディア学会 (LET) 関西支部メソドロジー研究部会 2011 年度報告論集』, 査読無, pp. 8-14.
2. 阪上辰也・鬼田崇作 (2011).「日本人英語学習者における定型表現の処理」『第 37 回 全国英語教育学会 山形研究大会発表予稿集』, 査読無, pp.132-133. 全国英語教育学会 第 37 回 山形研究大会 (山形大学)
3. 阪上辰也 (2011). 「学習者コーパス入門―NICE を利用して―」『外国語教育メディア学会 (LET) 関西支部 メソドロジー研究部会 2010 年度報告論集』, 査読無, pp. 74-99.
4. Matsuno, K., Murao, R., Morita, M., Sakaue, T., and Sugiura, M. (2010). Production units in English writing: A comparative study of writing fluency between native speakers and non-native speakers of English. in Hirakawa, M. et al. (eds). Studies in Language Sciences, 9, 査読有, pp. 143-159. Kuroshio Publishers.

〔学会発表〕（計 6 件）

1. Kida, S., Sakaue. T., & Tagashira, K. (2012). Processing of L2 formulaic sequences by Japanese EFL university learners. American Association for Applied Linguistics 2012 Boston, USA. (2012 年 3 月 25 日)
2. Sugiura, M., Yamashita, J. Leung, C.Y., Bnado, T., & Sakaue, T. (2012). Do L2 Learners Have the Same Collocational Knowledge as L1 Speakers? Evidence from Eye-tracking Data. American Association for Applied Linguistics 2012 Boston, USA. (2012 年 3 月 24 日)
3. 鬼田崇作・阪上辰也 (2011). 「第二言語における定型表現の処理単位：プライミング法による検討」JACET 英語語彙研究会第 8 回大会, 麗澤大学 (2011 年 12 月 10 日)
4. Sakaue, T., and Sugiura, M. (2011). A new learner corpus for SLA research: Dynamic Corpus of English Learners. Learner Corpus Research 2011, Louvain-la-Neuve, Belgium (2011 年 9 月 17 日)
5. 阪上辰也・鬼田崇作 (2011).「日本人英語学習者における定型表現の処理」第 37 回 全国英語教育学会 山形研究大会, 山形大学 (2011 年 8 月 20 日)
6. A comparative study of multi-word units: Written results and writing process by English learners. Exploring the Boundaries and Applications of CORPUS LINGUISTICS, The University of Alabama, USA. (2011 年 4 月 16 日)

〔その他〕
ホームページ等

http://lab.sakaue.info/wiki.cgi/StKaken2010

６．研究組織
(1)研究代表者
阪上　辰也（SAKAUE TATSUYA）
広島大学・外国語教育研究センター・特任講師
研究者番号：60512621

(2)研究分担者
（　　）
研究者番号：

(3)連携研究者
（　　）
研究者番号：