

中国語を母語とする上級日本語学習者の「名詞＋動詞」コロケーションの 使用

—YNU 書き言葉コーパスの分析を通して—

劉瑞利（お茶の水女子大学大学院生）

1. はじめに

日本語学習者のコロケーション使用について、「ビタミンを食べる」、「気分が嫌だ」などの誤用が報告されている（大曾・滝沢，2003）。コロケーションは母語話者らしさや流暢さに強く関わっており，上級学習者にとっても難しいと指摘されている（Pawley & Syder, 1983 ; Granger, 1998 ; 西川, 2014）。一方，これまで学習者のコロケーション使用について，L2 英語を対象とした研究は数多くあるが，L2 日本語に関するものは非常に少ない。特に，日本語学習者の「名詞＋動詞」コロケーションの使用に焦点を当てた研究は管見の限り見当たらない。本研究では，金澤（2014）「YNU 書き言葉コーパス」を分析し，中国語を母語とする上級日本語学習者（以下，上級 CJL）の「名詞＋動詞」コロケーションの使用実態及び日本語母語話者（以下，JNS）との使用上の違いについて報告する。また，母語の影響だと思われる誤用に関しては，韓国語を母語とする日本語学習者（以下，KJL）に同じ産出がないかも合わせて検証した。

2. 先行研究

2.1 「名詞＋動詞」コロケーションの定義について

「名詞＋動詞」（英語では「動詞＋名詞（Verb-Noun）」）語結合は通常，①「リンゴを食べる」のように，恣意的制限がなく，意味によって自由に組み合わせられる自由結合，②「電話をかける」のように，語の選択において恣意的制限を受けるが，句全体の意味が個々の語の意味から容易に分かる制限結合，③「道草を食う」のように，語と語の結びつきが固定的であり，句全体の意味が個々の語から読み取れない固定結合の3種類があるとされている（村木，1991 ; Nesselhauf, 2003, 2005 ; Laufer & Waldman, 2011）。コロケーションの範囲については意見が分かれているが，本研究では英語学習者の「動詞＋名詞」コロケーションの使用を調査した Nesselhauf (2003, 2005), Laufer & Waldman (2011) と同様に，中間に位置する制限結合のみをコロケーションとする。

2.2 学習者のコロケーション使用に関する研究

学習者のコロケーション使用について，主に学習者コーパスを扱った研究で行われてきた（小森，2003 ; 曹・仁科，2006 ; Granger, 1998 ; Laufer & Waldman, 2011 ; Nesselhauf, 2003, 2005）。そのうち，日本語の「名詞＋動詞」コロケーションに焦点を当てた研究は管見の限り見当たらない。

英語学習者の「動詞＋名詞」コロケーションの使用を調査した Laufer & Waldman (2011) は学習者が英語の習熟度に関わらず，母語話者よりコロケーションを有意に少なく使用していると指摘した。また，-ly で終わる増副詞のコロケーション使用を調査した Granger

(1998) も同様の指摘をしている。上級日本語学習者の「名詞+動詞」コロケーションの使用において同様であるかは興味深い。

Nesselhauf (2003, 2005) も英語学習者の「動詞+名詞」コロケーションの使用を調査した研究であるが、学習者が使用したコロケーションにおいて、不適切なものは約 1/3 であり、その半分が学習者の母語の影響を受けていると報告した。前述の Laufer & Waldman (2011) も母語の影響について同様の指摘をしている。日本語学習者のコロケーション使用を調査した小森 (2003)、曹・仁科 (2006) は「名詞+動詞」コロケーションに焦点を当てた研究ではないが、母語の影響も指摘している。これらの研究はいずれもコロケーションの誤用には母語の影響が大きいと報告しているが、その結果は母語話者の直感による判断のみに頼っており、他言語話者の産出に同じ誤用がないかを調査していない。

3. 本研究の課題

以上の先行研究を踏まえ、本研究では以下の 3 つの研究課題を立てた。

RQ1: 上級 CJL における日本語の「名詞+動詞」コロケーション使用に過少使用或いは過剰使用が見られるか。

RQ1-1: JNS と比べ、全体的な使用数において見られるか。

RQ1-2: JNS と比べ、どのような共起語において見られるか。

RQ2: 上級 CJL はどのようなコロケーションを間違えやすいか。

RQ2-1: コロケーションの誤用にはどのようなタイプがあるか。

RQ2-2: 母語である中国語の影響を受けていると推測された誤用はどのくらいあるか。

RQ3: RQ2-2 で中国語の影響を受けていると推測された誤用は KJL に見られるか。

4. 研究方法

4.1 使用するコーパスについて

金澤 (2014) 「YNU 書き言葉コーパス」は JNS 30 名、同じ大学に留学している CJL 30 名、KJL 30 名を対象とした 12 種類のタスクによる書き言葉コーパスである。本研究では上級学習者であることを確保するために、日本語能力試験 (JLPT) N1 或いは旧 1 級に合格した CJL 26 名を調査対象とした。母語話者と比較する際に、JNS をランダムに 26 名抽出した。なお、母語の影響を検証する KJL のデータは多いほど説得力があると思われるため、RQ3 で検証する際に KJL 全員のデータを使用した。作文データにはオリジナルデータと補正データの 2 種類あるが、本研究ではデータベースでの検索の便利性を持っている補正データを使用した。上級 CJL、JNS 及び KJL の作文データのサイズは表 1 の通りである。

表 1 上級 CJL、JNS 及び KJL の作文データのサイズ

	上級 CJL (26名)	JNS (26名)	KJL (30名)
延べ語数	79232	67980	81121
異なり語数	3901	3354	3357

4.2 コロケーションの抽出方法

本研究では複合動詞と漢語動詞を除いた和語動詞に限定し、「名詞ヲ動詞」、「名詞ガ動詞」「名詞ニ動詞」の3つのパターンを抽出した。コロケーションの抽出手順として、まず形態素解析器（茶まめ Ver. 2.0 for Windows & unidic-mecab 2.1.2）を用いて、形態素解析を行い、作文データを Excel に出力した。その次に、手作業で「名詞+動詞」コロケーションを一つずつ抽出した。

コロケーションであるかを判断する際に、Nesselhauf (2005) 及び Laufer & Waldman (2011) と同様に、辞書を使用することにした。具体的にはコーパスでの産出文において、動詞が最も基本的な意味（以下、本義）として使われている場合は、自由結合とし、抽出しなかった。本義の判断は語の意味情報がより豊富な『スーパー大辞林 3.0』及び『新明解国語辞典 第六版』を使用した。一般的に本義は辞書の最初に挙げられるため、2つの辞書ともに一番目の意味として挙げられている場合、その動詞の意味が本義であると判断した。なお、2つの辞書で判断できない場合は、『大辞泉』（CD-ROM 版、第1版）を参考にした。また、「名詞+動詞」語結合は『日本語慣用句辞典』に載っており、且つ『スーパー大辞林 3.0』では「句」として挙げられている場合は慣用句に分類した。ただし、両方には載っているものの、コーパスでの産出文における意味は比喩的な意味ではなく、個々の語から予測できる場合はコロケーションとした。

4.3 コロケーションの許容度判定

コロケーションの誤用状況を明らかにするために、学習者が産出したコロケーションは日本語として許容できるかを、①辞書での掲載の有無、②大規模母語話者コーパス（NINJAL-LWP for BCCWJ）での用例の有無、③母語話者による判定という3段階によって判定した。③の母語話者は日本語教育を専攻とする日本人大学院生であり、全員日本語教育の経験者である。母語話者に判定してもらう際に、産出文における使用は、「適切であり、完全に許容できる（○）」、「少々問題がある、或いはもっと適切な表現があるかもしれないが許容できる（△）」、「完全に許容できない（×）」の3つの基準を設けた。最初は母語話者2人に判定してもらったが、2人の判定結果が異なる場合、第三者に判定してもらい、3人の判定結果で総合的に最終結果を出した。

4.4 母語の影響の判断

コロケーションの許容度判定で△や×と判定されたものが学習者の母語である中国語の影響を受けているかを見るために、Nesselhauf (2003, 2005), Laufer & Waldman (2011) の判断方法に倣い、日本の大学院に在籍し、日本語教育を専攻とする中国語母語話者2人に判断してもらった。2人の判断結果が一致しない場合はコロケーションの許容度判定と同様に、第三者に再判断してもらった。ここでの判断はあくまでも中国語母語話者の直感によるものであり、RQ3では、母語の影響を受けていると判断されたものを KJL の産出と照らし合わせ、本当に母語の影響を受けているかを検証する。

5. 結果及び考察

上級 CJL 26名の作文データより抽出されたコロケーションは延べ数 838 例あった。コロケーションの許容度判定の結果、誤用（△+×）は 126 例で、全体の 15%を占める（表 2）。先行研究で指摘された 1/3 より低いが、無視できない数字だと言える。以下、各 RQ につい

て見ていく。

表2 上級CJLが使用したコロケーションの許容度判定結果

	○	△	×	合計
コロケーション数	712	57	69	838
比率	85%	7%	8%	100%

5.1 JNS との比較 (RQ1)

上級CJLの作文データから抽出されたコロケーションには誤用もあるため、JNSと比較する際は、Laufer & Waldman (2011)と同様に、適切だと判定されたコロケーション（本研究では○と判定されたコロケーション）のみを扱うことにした。

5.1.1 全体的な使用数において (RQ1-1)

上級CJLが使用したコロケーションのうち、○と判定されたものを10万語あたりの相対頻度を基準に、JNSが使用したコロケーションと比較した結果、上級CJLは1%水準 ($\chi^2(1)=7.37, p<.01$) でコロケーションを多く使用していた。

また、本研究で扱う上級CJLとJNSのコーパスは同じタスクに基づいたものであるため、上級CJLとJNSがともに使用したコロケーションの割合が高いだろうと予測される。しかし、上級CJLの○と判定されたコロケーションとJNSが使用したコロケーションを、①上級CJLのみが使用したもの、②JNSのみが使用したもの、③両方が使用したものの、の3種類に分類した結果、③の両方が使用したものは20%に達していないことが分かった(表3)。

表3 コロケーションの使用分布 (異なり数)

	上級CJLのみ	上級CJLとJNS	JNSのみ
上級CJL	341 (83.8%)	66 (16.2%)	
JNS		66 (18.1%)	299 (81.9%)

このように、コロケーションの全体的な使用数においては、これまでの研究 (Granger, 1998; Laufer & Waldman, 2011) とは違い、上級CJLがJNSより多く使用している。これは調査する言語の違い及び学習者母語の特徴によるものだと考えられる。中国語と日本語の対照研究では、中国語が日本語に比べ、動詞を多用する傾向にあると指摘されている (加藤, 2014)。例えば、日本語では「おタバコはご遠慮ください」と言うのに対し、中国語では“禁止吸烟 (タバコを吸うのを禁止する)”と言う。また、同じタスクに基づいたコーパスにも関わらず、上級CJLとJNSがともに使用したコロケーションはそれぞれの20%に達していないことから、上級CJLが同じ事柄を説明するにはJNSと異なる表現を使用する傾向にあると言えよう。このことは、上級CJLが使用したコロケーションは母語話者に許容できるものの、まだPawley & Syder (1983)が述べた「母語話者並みの選択」になっていないことを示唆している。

5.1.2 具体的な共起語において (RQ1-2)

上級CJLがどのような動詞を過剰使用または過少使用しているかを両方ともに使用した動詞の粗頻度の差及び共起名詞の異なり数の差を基準に調べた。具体的な基準として、そ

の動詞は上級CJLとJNSにおける粗頻度の差及び共起名詞の異なり数の差が5以上で、上級CJL対JNSが1.5倍以上の場合を過剰使用、上級CJL対JNSが0.5倍以下の場合を過少使用とした。その結果、上級CJLが過剰使用の基準に達した動詞として、中国語にも対応語があり、よく使用される動詞「書く・持つ・見る・入る・感じる・読む」の6つ、過少使用の基準に達した動詞として、意味範囲が広く抽象度の高い多義動詞「かかる」の一つが観察された。このことから、学習者の母語と意味的に対応している動詞の過剰使用は起こりやすく、意味範囲が広く抽象度の高い多義動詞の習得は遅れていると言える。また、本研究で観察された過少使用語「かかる」に関しては、上級CJLでも習得が進んでいないと指摘した鷲見(2014)の結果と一致している。

5.2 コロケーションの誤用について (RQ2)

5.2.1 上級CJLのコロケーションの誤用タイプ (RQ2-1)

CJLが使用した「名詞+動詞」コロケーションのうち、△や×と判定された126例をNesselhauf(2003)の分類方法を参考に、①動詞、②名詞、③助詞、④用法1(コロケーション自体は適切であるが、動詞の変形、テンス、アスペクト、ヴォイスなどに問題があるもの)、⑤用法2(コロケーション自体は存在するが、違う文脈で使用されたもの)、⑥用法3(コロケーション自体は存在せず、しかも一つの要素だけで修正できないもの)、の6種類に分類した。そのうち、動詞による誤用が37例もあり、最も多かった。動詞の意味・用法をきちんと理解することの重要性が示された。それに続いたのは用法3(27例)、用法1(22例)である。

5.2.2 母語である中国語の影響について (RQ2-2)

誤用と判定された126例を中国語母語話者に判断してもらった結果、中国語の影響を受けていると判断されたものは34例(異なり数:29例)あり、誤用の全体の27%だった。Laufer & Waldman(2011)、Nesselhauf(2003, 2005)が指摘した半分にはなっていないが、上級CJLの誤用にも母語の影響が大きいと指摘できる。

5.3 母語の影響の検証 (RQ3)

中国語の影響を受けていると判断されたコロケーションをKJLの産出と照らし合わせた結果、KJLにも同じ産出があったものとして、「だしをつくる」、「(婦人)科が備える」、「声が出る(「意見が上がる」の意味で使用)」の3例のみで、10%以下であった。つまり、母語の影響だと判断された誤用の大部分は中国語の影響を受けている可能性が高いと言える。上級CJLが母語の影響を受けているコロケーションの代表例として、中国語の知識で動詞の意味を理解する「牛を放す」(産出文:その男の人は「彥星」と言って、牛を放す人です。)、日本語には存在しないが、中国語では典型的なコロケーションである「愛の河に陥る(「恋に落ちる」の意味で使用)」、中国語と日本語の両方に存在するが、使用文脈の異なるコロケーション「声が出る」(産出文:留学生奨学金を増やすべき声が出てきました。)、などがある。

このように、学習者の母語では典型的であるが目標言語では存在しないコロケーション及び、両方ともに存在するが使用文脈が異なるコロケーションは母語の影響を受けやすいと考えられる。前者に関しては、目標言語では存在しないことを明示的に教える必要がある。後者については適切な文脈の中で提示し、理解してもらうことが大切であろう。

6. まとめ

本研究では JNS の使用と比較することで、上級 CJL が母語話者より「名詞+動詞」コロケーションを多く使用していること、過剰使用語と過少使用語があることを明らかにした。母語の影響に関しては、母語話者による判断のみではなく、他言語話者の産出とも照らし合わせた。学習者コーパスと母語話者コーパスを比較する際に、同世代の人を対象にした、同じタスクに基づいたコーパスを選ぶようにしたが、タスクの種類が限られており、類似のタスクもあったため、結論は一般化し難いと考えられる。今後、コロケーションの抽出方法及び判断基準、母語の影響の検証方法を再検討し、より膨大な学習者コーパスを用いて調査する必要があると考えられる。しかし、本研究で上級 CJL のコロケーション使用及びその誤用の特徴について明らかにした点は、日本語教育の現場に多くの示唆を与えるものである。

参考文献

- (1) 大曾美恵子・滝沢直宏 (2003) 「コーパスによる日本語教育の研究—コロケーション及びその誤用を中心に—」『日本語学』第 22 卷(5), 234-244.
- (2) 金澤裕之 (2014) 『日本語教育のためのタスク別書き言葉コーパス』ひつじ書房
- (3) 加藤晴子 (2014) 「『中国語好み』の表現—名詞的表現と動詞的表現」『言葉とその周辺をきわめる 3』東京外国語大学オープンアカデミー教養講座
- (4) 小森早江子 (2003) 「英語を母語とする日本語学習者の語彙的コロケーションに関する研究」『第二言語としての日本語の習得研究』6 号, 33-51.
- (5) 鷲見幸美 (2014) 「中国語を母語とする日本語学習者の和語動詞の使用—KY コーパスの分析」『言語文化論集』第 36 卷(1), 65-79.
- (6) 曹紅荃・仁科喜久子 (2006) 「中国人学習者の作文誤用例から見る共起表現の習得及び教育への提言—名詞と形容詞及び形容動詞の共起表現について—」『日本語教育』130 号, 70-79.
- (7) 西川朋美 (2014) 『『母語話者レベル』の L2 能力に関する一考察—YNU 書き言葉コーパス』の『超』上級日本語学習者を対象に— 金澤裕之 (編) 『日本語教育のためのタスク別書き言葉コーパス』403-417, ひつじ書房
- (8) 村木新次郎 (1991) 『日本語動詞の諸相』ひつじ書房
- (9) Granger, S. (1998) Prefabricated patterns in advanced EFL writing: Collocations and lexical phrases. In A. P. Cowie (Ed), *Phraseology: Theory, analysis, and applications*. Oxford University Press. (南出康世・石川慎一郎監訳 (2009) 『慣用連語とコロケーション—コーパス・辞書・言語教育への応用』185-204. くろしお出版)
- (10) Laufer, B., & Waldman, T. (2011) Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647-672.
- (11) Nesselhauf, N. (2005) *Collocations in a learner corpus*. John Benjamins Publishing.
- (12) Nesselhauf, N. (2003) The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223-242.
- (13) Pawley, A., & Syder, F. H. (1983) Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. *Language and Communication*, 191-226.