

『複合動詞レキシコン』ver.1 —形態的・統語的・意味的情報付与—

神崎 享子

豊橋技術科学大学 情報メディア基盤センター

要 旨

本発表では、形態的・統語的・意味的情報を付与した「動詞+動詞」型の複合動詞レキシコンについて述べる。本レキシコンは、国立国語研究所の「日本語レキシコンの形態的・統語的・意味的特性」プロジェクト（リーダー：影山太郎）の研究成果の一部として作成されたものである。

和語の「動詞+動詞」型複合動詞は使用頻度が高く日本語表現に特徴的な語彙であるものの、量的・質的両面で十分に情報を付与したレキシコンは未だ見られない。本発表では、語彙的複合動詞を対象とした複合動詞レキシコンの設計や付与情報について述べる。そして本レキシコンが言語研究や日本語学習者の複合動詞学習に応用できる可能性を示す。

1. はじめに

これまで、単純語の動詞に比べて、「動詞+動詞」型の複合動詞について形態的・統語的・意味的情報を十分に付与し、かつ、ある程度の語数を扱ったレキシコンは見られなかった。また、複合動詞は、辞書などの登録語について登録基準が明確ではないという問題点がある。たとえば、いくつかの辞書では「なだれ込む」はあるが「練り込む」は登録されておらず、「運び出す」はあるが「運び入れる」は登録されていない。

このような状況は、たとえば日本語教育の分野でも問題となる。森田（1978）は、複合動詞の学習・習得は必要性があるにもかかわらず、体系的に教えることは難しいと指摘している。教育での必要性から日本語教育の側では、これまで辞書やデータベースがいくつか作られている。たとえば、日本語学習を目的とした複合動詞辞書には、Tagashira and Hoff(1986)がみられるが、詳細な情報は掲載されているものの、規模は小さい。規模の大きい日本語教育用のデ

ータベースとしては山下(2007)があるが、量的な面が充実している一方で、日本語学習で使用する場合には、質的な面でさらなる情報を追加する必要がある。また、野村・石井（1987）は計量的に語構成や用法を調査・分析しているが、量的な調査という目的から、逆に質的な面は手薄である。

複合動詞は東アジアに特徴的な語彙ということもあり、複合動詞辞書を充実させることは、欧米言語との翻訳などの際にも、柔軟な表現を訳出する一助になると考える。量だけではなく質的にも多様な情報が付与されている複合動詞辞書を作成することで、今まで扱えなかった統語情報や意味情報からの複合動詞検索や、動詞と動詞の組み合わせパターンからの検索、システムティックな複合動詞学習、柔軟な表現の翻訳など、幅広い研究・教育で役立つと考える。我々は、複合動詞に形態的・統語的・意味的情報を付与した『複合動詞レキシコン』¹を作り、2013年3月に公開する。本発表ではこの複合動詞レキシコンについて述べる。

2. 対象とする複合動詞

2.1 複合動詞のタイプ

本レキシコンでは、和語動詞からなる「動詞+動詞」型複合動詞を扱う。「動詞+動詞」型複合動詞は、前項と後項の動詞結合の緊

¹本レキシコンは、国立国語研究所の「日本語レキシコンの形態的・統語的・意味的特性」プロジェクト（リーダー：影山太郎）の研究成果の一部として作成されたものである。国立国語研究所ホームページより公開する。

密性から、語彙的複合動詞と統語的複合動詞の二つに大別できるとされている（影山 1993）。語彙的複合動詞の場合は、二語の結合に制限があり生産性が低い。たとえば「聞き漏らす」はよいが「見漏らす」「話し漏らす」などとは言わないことから、「漏らす」がどのような動詞にも結合するわけではないことがわかる。一方、統語的複合動詞の場合は統語的な派生から形成され、比較的制限を受けずに自由に結合することができる。たとえば「始める」などは、「歌い始める」「勉強し始める」「通い始める」「作り始める」など様々な動詞にでも結合することができ、補助動詞の役割をもつ。意味の面でも、二つの単純語の動詞の語義が変化を受けずに合成されている。また、統語的複合動詞は後項動詞の数が少なく、影山(1993)では 27 語、姫野（2001）では 3 語追加して 30 語としている。これらの動詞は、前項動詞と自由に結合することから、生産性が高い。本レキシコンでは、前者の語彙的複合動詞を対象とし、現在 2735 語を登録している。

2.2 収集対象の資料

データの収集は、複合動詞の辞書や研究書などを参考にした。

影山太郎（1993）『文法と語形成』

田中茂範・松本曜（1997）

『空間と移動の表現』

姫野昌子（1999）

『複合動詞の構造と意味用法』

姫野昌子（2004）『日本語表現活用辞典』

山下喜代（2007）

『平成 17 年度（2005）～平成 19 年度（2007）

科学研究費補助金 基盤研究（c） 研究成果報告書 日本語教育のための合成語のデー

タベース構築とその分析』

また、上記資料で収集したものだけではなく、不足分はその都度補充していくこととした。

3. 複合動詞レキシコンの設計

これまでの言語研究や動詞辞書などの知見を活かした形で複合動詞に情報を付与するために、事前調査を行った（神崎 2012）。その中で、第一版では最も基本的な情報を付与した。

① 語構成

複合動詞の語構成が、自立した動詞どうしの組み合わせ「動詞(V)+動詞(V)」か、あるいは「接頭語(p)+動詞(V)」か「動詞(V)+接尾語(s)」か、などの情報。

② 複合動詞の漢字表記、ひらがな表記、ローマ字表記

③ 前項動詞の漢字表記、ひらがな表記、ローマ字表記

④ 後項動詞の漢字表記、ひらがな表記、ローマ字表記

②から④については、ひらがな表記だけではなくローマ字表記も対応することで日本語学習者が検索しやすくする。

⑤ 自他と主語の意志性の有無

言語研究において、主語の意志性は複合動詞形成の制約にかかわる重要な要素となる（影山 1993）。また、他動詞・自動詞・自動詞（意志性のある主語をとるもの）の区別は複合動詞の統語的・意味的特性を考える際に重要な情報となる。

⑥ 複合動詞の格パターン

複合動詞の意味を形成するのに必須と思われる格を記載している。ただし、必ずしも実際の文にすべての格が出そろ

っているとは限らない。

- ⑦ 異表記や固定した用法などの記載
異表記での検索にも対応する。

- ⑧ 定義文
日本語の定義文を記載している。

- ⑨ 用例
現時点では一部の用例を記載しているが、将来的には、国立国語研究所で公開しているオンライン検索 NINJAL-LWP for BCCWJ (NLB) (国立国語研究所 2012) と連動して、BCCWJ コーパスの用例を検索できるようにする。

- ⑩ 頻度情報
上記項目以外に統計的情報もオンライン検索 NLB を利用して調べることができる。BCCWJ コーパス内での複合動詞

の頻度、前項動詞の頻度、後項動詞の頻度など様々な角度からの頻度情報が見られる。

以下にいくつかの使用例を示す (例 1～例 3)。BCCWJ コーパスは、書籍、雑誌、新聞、ブログなどいくつかのサブコーパスから構成されているが、例 1 では BCCWJ 全体での総出現頻度とサブコーパス別の出現頻度を示すとともに、その複合動詞がどのような活用形で使用されていることが多いか、受身形、使役形、否定形がそれぞれどのような割合で使用されているのかなどを示している。例 2 ではどの動詞が複合動詞の前項にくる動詞として多いかについて情報を示し、例 3 では前項と後項の動詞の組み合わせの情報を示している。

例 1 BCCWJ コーパスでの頻度情報の一例

語彙素：繰り返す

読み：クリカエス

活用：五段・サ行

総頻度：5574

サブコーパス頻度 (PMW) : NR: 4628 (104.61), SP: 240 (43), OM: 326 (65.66), OC: 273 (51.16), OY: 107 (44.6)

表記形頻度 (表記形サブコーパス頻度) : 繰り返す: 4450 (NR: 3576, SP: 188, OM: 325, OC: 260, OY: 101), くり返す: 578 (NR: 544, SP: 27, OM: 0, OC: 3, OY: 4), くりかえす: 394 (NR: 370, SP: 16, OM: 0, OC: 7, OY: 1), 繰返す: 94 (NR: 85, SP: 4, OM: 1, OC: 3, OY: 1), 繰りかえす: 49 (NR: 45, SP: 4, OM: 0, OC: 0, OY: 0), 繰り返えす: 7 (NR: 6, SP: 1, OM: 0, OC: 0, OY: 0), くり返えす: 1 (NR: 1, SP: 0, OM: 0, OC: 0, OY: 0), 繰り返へす: 1 (NR: 1, SP: 0, OM: 0, OC: 0, OY: 0)

活用形頻度：連用形: 3459, 基本形: 1276, 未然形: 780, 假定形: 49, 未然ウ接続: 10

れる・られる後続割合 (頻度) : 11.5% (繰り返される: 642)

せる・させる後続割合 (頻度) : 0.3% (繰り返させる: 17)

ない・ぬ・ません後続割合 (頻度) : 2.3% (繰り返さない: 107, 繰り返さぬ: 11, 繰り返しません: 8)

例 2 前項動詞からの検索

全体から

前項から

後項から

出現位置

クロス表

後項	読み	動詞数
取り～	トリ	71
引き～	ウチ	55
打ち～	ウチ	51
押し～	オシ	43
見～	ミ	43
突き～	ツキ	35
思い～	オモイ	30
切り～	キリ	30
立ち～	タチ	29
差し～	サシ	26
振り～	フリ	22
吹き～	フキ	22
言い～	イイ	21
書き～	カキ	21
掻き～	カキ	21
踏み～	フミ	21
飛び～	トビ	19
聞き～	キキ	18
追い～	オイ	16
買い～	カイ	16

<Prev1234>Next

取り～ 118

複合動詞	読み	自・他	BCCWJ頻度
取り上げる	トリアゲル	他	3,256
取り巻く	トリクム	自	2,782
取り出す	トリダス	他	2,722
取り入れる	トイルレル	他	1,491
取り戻す	トリモドス	他	1,406
取り込む	トリコム	他	919
取り除く	トリノク	他	919
取り扱う	トリアツカウ	他	912
取り巻く	トリマク	他	804
取り付ける	トリツケル	他	684
取り囲む	トリカコム	他	582
取り済す	トリクセス	他	466
取り替える	トリカエル	他	438
取り合う	トリアウ	他	421
取り残す	トリノコス	他	367
取り立てる	トリシタデル	他	324
取り締まる	トリシマル	他	305
取り寄せる	トリヨセル	他	250
取り外す	トリハズス	他	236

<Prev1234>Next

例 3 前項と後項動詞の組み合わせ検索

全体から

前項から

後項から

出現位置

クロス表

前項

後項

直す, 戻す, 返す

作成

前項	直す	戻す	返す
引き～	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
取り～	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
打ち～	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
思い～	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
煮～	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
繰り～	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
追い～	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
やり～	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
吹き～	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
掃り～	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
送り～	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
言い～	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
握り～	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
作り～	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
読み～	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
書き～	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
眺め～	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

<Prev

1

2

3

4

>Next

4. 今後の課題

複合動詞レキシコンは、2013 年 3 月に国立国語研究所のホームページから公開予定である。『複合動詞レキシコン』ver.1 では基本的な情報を付与しているが、今後、複合動詞の選択制限となる共起名詞や、複合動詞の上位語などいくつか主要な情報についても検討していく予定である。

謝辞

Lago 言語研究所の赤瀬川史朗氏には、システム開発や様々なアドバイスなど多くのご尽力を頂きました。感謝いたします。

参考文献

- 影山太郎 (1993) 『文法と語形成』 ひつじ書房.
- 神崎享子 (2012) 「複合動詞データベース構築のための付与情報」『国立国語研究所論集 3』 国立国語研究所, pp.1-18.
- 国立国語研究所 (2012) NINJAL-LWP for BCCWJ (NLB) <http://nlb.ninjal.ac.jp/>
- Tagashira, Yoshiko and Jean Hoff (1986)

Handbook of Japanese compound verbs. Hokuseido.

- 田中茂範・松本曜 (1997) 『空間と移動の表現』 研究社.
- 野村雅昭・石井正彦 (1987) 『文部省科学研究費補助金特定研究 (1) 言語データの収集と処理の研究 複合動詞資料』.
- 姫野昌子 (1999) 『複合動詞の構造と意味用法』 ひつじ書房.
- 姫野昌子 (2001) 「複合動詞の性質」『日本語学 20(8)』 明治書院, pp.6-15.
- 姫野昌子 (2004) 『日本語表現活用辞典』 研究社.
- 森田良行 (1978) 「日本語の複合動詞について」『講座日本語教育 14』 早稲田大学語学教育研究所, pp.69-86.
- 山下喜代 (2007) 『平成 17 年度 (2005) ～ 平成 19 年度 (2007) 科学研究費補助金 基盤研究 (c) 研究成果報告書 日本語教育のための合成語のデータベース構築とその分析』.