

複合動詞の生産性といわゆる「統語的／語彙的」の区別 —コーパスにもとづく考察—

松村 一登

東京大学

kazuto@tooyoo.L.u-tokyo.ac.jp

本発表で考察の対象とするのは「踏み外す」のような「動詞(V₁)＋動詞(V₂)」型の複合動詞である。V₁を前項動詞、V₂を後項動詞と呼び、前項動詞は一般に、いわゆる連用形で現れる。

本研究でコーパスと呼ぶ日本語資料は、文庫本400冊のテキスト・ファイル版から構成される。形態解析プログラム「茶筌」(2.3.3)によって分割された単位(ただし記号・句読点を除く)を語とみなすとすれば、約3070万語、新聞1年分の記事にほぼ相当するテキスト量がある。この言語資料は109人の著者のテキストを含み、テキスト量が最多の著者は112.68万語、最少の著者は4.57万語、中央値は19.99万語である。複合動詞の用例の検索、KWIC索引化、集計にはPerl、表とグラフの作成には、OpenOffice.org 2.1の表計算ソフトCalcを用いた。

日本語の複合動詞に関してはいくつかの理論的立場からの考察がなされているが、生産性(produktivity)に関しては、「食べ始める」のような生産的タイプがある一方で、「踏み外す」のような非生産的タイプもあるとする直観的な提示がされることはあっても、複合動詞の生産性の現象そのものを対象とした研究は見うけられないようである。

複合動詞の生産性の議論に理論的な拠りどころをおくのは、影山(1993)などに代表される一連の研究である。この理論的枠組みでは、複合動詞を2つのグループ(統語的／語彙的)に分け、その際の基準の1つに、生産性の違いを指摘する。本発表では、影山(1993)などの背景にある理論的な枠組みに立ち入るのを避けるために、論点を次のように定式化する。

[1] 複合動詞は、2つのグループに峻別され、その区別は、生産性の度合いに直接反映する。

複合動詞の生産性に関するこの主張(仮説)は「複合動詞の生産性が測定でき、比較できる」ことを前提としており、複合動詞の2つのグループ(統語的複合動詞／語彙的複合動詞)の間の区別は、その生産性の度合いの違いにおいて明確な形で観察されることを予測していると解釈される。複合動詞に関する仮説[1]の経験的妥当性をコーパスに基づいて検証するのが本発表の主たる目的である。

後項動詞(V₂)を共有する複合動詞は、1つの複合動詞の型(クラス)を作っていると考え、V₂にしたがって「～まくる」型複合動詞などと呼ぶ。特定の複合動詞(タイプとしての複合動詞)のテキスト中の具体的な現れを、その複合動詞の用例と呼ぶ。たとえば、「やがて思い直したように」には、「～直す」型の複合動詞「思い直す」の用例「思い直した」が現れている。

複合動詞の型は、「～はじめる」「～あるく」などのように、V₂と1対1に対応するとみなせる場合も多いが、「～かける」「～のこす」「～すぎる」「～きる」「～とおす」「～ぬく」「～だす」など、対応が一義的でないことがある(例文[2][3])。例文[4]のように、V₁とV₂がともに同じで、型が異なるとみなすべき場合がある。V₁自身が複合動詞であることも少なくない。例文[5]の「取り戻しかけた」「くみとりかねた」は、それぞれ「取り戻す」「くみとる」をV₁とする複合動詞であって、「戻す」「とる」をV₁とする複合動詞「戻しかける」「とりかねる」とは異なる複合動詞の用例である。V₂が受動形や使役形で現れている用例で、V₂が基本形(能動形)の複合動詞が用いられない場合は、独立の複合動詞の型とみなす。たとえば、例文[6a]の「焼け出された」は、「焼け出す」が用いられないので、「～だす」型の複合動詞の用例には含めず、独立した複合動詞「焼け出される」の用例と考える。他方、例文[6b]の「撮り直させて」は、複合動詞「撮り直す」の用例に含める。

- [2]a. 雪はほとんどやみかけていた。(半村良) [～かける (1): 統語的]
 b. 二階のテラスから呼びかけても、返事はなかった。(井原まなみ) [～かける (2): 語彙的]
- [3]a. 浦上はビールを飲み残したままレストランを出た。(津村秀介)
 b. 広瀬は、捨て台詞のように言い残して、ベンツから降りた。(門田泰明)
- [4]a. [...] 転がり出してしまった雪ダルマのように、もう止められない [...] (西村京太郎)
 b. ガラガラッとハンドルを回すとコロンと玉が転がりだす福引き機 [...] (金春智子)
- [5]a. 矢桐が持ち前の強気を取り戻しかけたとき、一個の事件が発生した。(森村誠一)
 b. 秋元も意味をくみとりかねた。(長井彬)
- [6]a. 五日前に火事があって、皆焼け出されたんですよ (小杉健治)
 b. [...] すいません、もう一度撮り直させて下さい (つかこうへい)

あるテキスト中における特定の複合動詞の用例の総数を、本発表では、その複合動詞の**使用頻度**と呼ぶ。特定の型に属するすべての複合動詞の総体としての使用頻度を、本発表では、その複合動詞の型としての**総頻度**と呼び、Nで表す。テキスト中に用例がただ1つ、すなわち使用頻度1の語は**一回語 (hapax legomena)**と呼ぶ。本発表では、ある型に属する複合動詞の総数をM、その型の一回語の複合動詞の数をHで表す。

ある型の複合動詞の総頻度に対する一回語の複合動詞の数の割合 $P (= H/N)$ は語形成における生産性の指標になる (Bauer 2001)。Pの値は一般に非常に小さい小数值になるので、本発表では、この値を 10^3 倍した数字を生産性の指標として用いる。この数字は、その型の複合動詞の用例1000例の中に含まれる一回語の数の期待値に相当する。本発表では、Pの他に、2つの指標を用いる。ある型に属する複合動詞について、ひとつひとつの複合動詞の当該コーパスにおける平均使用頻度 $X (= N/M)$ と、ある型に属する複合動詞の用例1000あたりの異なる複合動詞の数 $T (= M \times 1000/N)$ である。本発表では、Tをその複合動詞の型の**語彙的多様性**と呼ぶ。Tは、当該コーパスにおいてすべての複合動詞の型が1000例使用されていると想定した場合に、それぞれの型に属する(タイプとしての)複合動詞の異なり語数である。XとTの間には、 $X \times T = 1000$ の関係がある。

複合動詞の生産性に関して、コーパスから得たデータから、次のことが明らかになる。

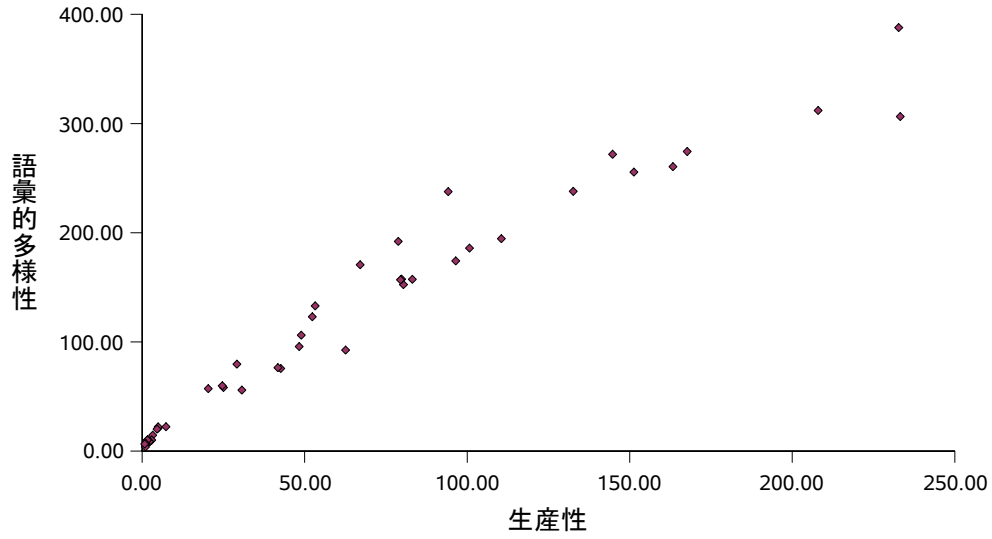
- [7] 統語的複合動詞のPの値は、一般に語彙的複合動詞のPの値より高いが、Pの値の変化は連続的であり、統語的／語彙的の区別に対応するような生産性におけるはっきりとした境界点は見いだせない。(表1)
- [8] Pの値とTの値の間には強い正の相関関係がある。すなわち、生産性が高い複合動詞の型ほど、属する複合動詞の数が相対的に多くなる。(図1)
- [9] Pの値とXの値の間には強い負の相関関係がある。すなわち、生産性の低い複合動詞の型ほど、ひとつひとつの複合動詞の使用頻度が、一般に高くなる。
- [10] Pの値とNの値の間にはある程度の負の相関関係が見られる。すなわち、生産性の低い複合動詞の型ほど、テキスト中での総頻度(総体としての使用頻度)が高くなる傾向がある。

生産性の高い複合動詞の型と生産性の低い複合動詞の型との間には、テキスト中での使用に非常にはっきりとした違いが見られる(図2～図7)。この意味で、複合動詞の生産性とその振る舞い方の違いの研究をする意義は大きいと考えられる。他方、統語的複合動詞とされているが、たとえば、生産性がさほど高くない「～なれる」と、語彙的複合動詞とされているが、生産性の比較的高い「～あるく」と比べた場合、その振る舞い方が際だって異なっているようには見えない(図8～図13)。複合動詞の生産性の違いは段階的なものであり、複合動詞の統語的／語彙的の区別の決定的な根拠となる指標の1つとは見なせないことを示唆していると考えられる。

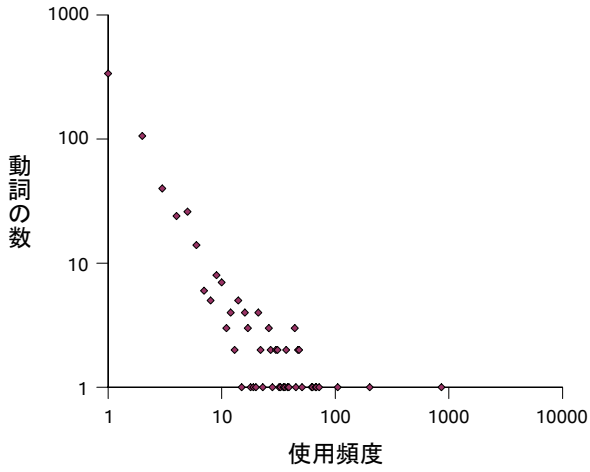
参考文献: : 影山太郎 1993: 『文法と語形成』 (ひつじ書房)
 Laurie Bauer 2001: *Morphological Productivity*. Cambridge University Press.

型	M 動詞数	N 総頻度	H 一回語	P 生産性	T 語彙的多様性	X 平均頻度
～かねない	247	806	188	233.25	306.45	3.26
～そこねる	45	116	27	232.76	387.93	2.58
～そこなう	78	250	52	208.00	312.00	3.21
～やすい	504	1837	308	167.66	274.36	3.64
～がち	142	545	89	163.30	260.55	3.84
～たがる	299	1170	177	151.28	255.56	3.91
～づらい	62	228	33	144.74	271.93	3.68
～まくる	113	475	63	132.63	237.89	4.20
～にくい	294	1511	167	110.52	194.57	5.14
～おえる	155	834	84	100.72	185.85	5.38
～すぎる(1)	464	2663	257	96.51	174.24	5.74
～たて	53	223	21	94.17	237.67	4.21
～つづける	783	4979	414	83.15	157.26	6.36
～かける(1)	641	4204	338	80.40	152.47	6.56
～おわる	142	902	72	79.82	157.43	6.35
～はじめる	1585	10106	803	79.46	156.84	6.38
～わすれる(1)	39	203	16	78.82	192.12	5.21
～のこす(1)	28	164	11	67.07	170.73	5.86
～得る	176	1901	119	62.60	92.58	10.80
～がたい	120	902	48	53.22	133.04	7.52
～ぬく(1)	47	382	20	52.36	123.04	8.13
～つくす	89	838	41	48.93	106.21	9.42
～かねる	109	1138	55	48.33	95.78	10.44
～とおす(1)	55	727	31	42.64	75.65	13.22
～たい	2432	31842	1330	41.77	76.38	13.09
～だす(1)	372	6648	204	30.69	55.96	17.87
～なれる	41	515	15	29.13	79.61	12.56
～なおす	193	3318	83	25.02	58.17	17.19
～きる(1)	461	7705	190	24.66	59.83	16.71
～あるく	45	787	16	20.33	57.18	17.49
～まわる	61	2737	20	7.31	22.29	44.87
～さる	51	2433	12	4.93	20.96	47.71
形+がる	63	2851	14	4.91	22.10	45.25
～たてる	66	3284	15	4.57	20.10	49.76
～変える	62	4281	14	3.27	14.48	69.05
～かえす	93	9152	27	2.95	10.16	98.41
～あがる	89	10447	25	2.39	8.52	117.38
～返る	39	4401	10	2.27	8.86	112.85
～とる	81	7765	13	1.67	10.43	95.86
～あげる	150	14728	24	1.63	10.18	98.19
～つける(2)	109	16061	18	1.12	6.79	147.35
～かける(2)	60	15807	16	1.01	3.80	263.45
～こむ	258	39750	31	0.78	6.49	154.07

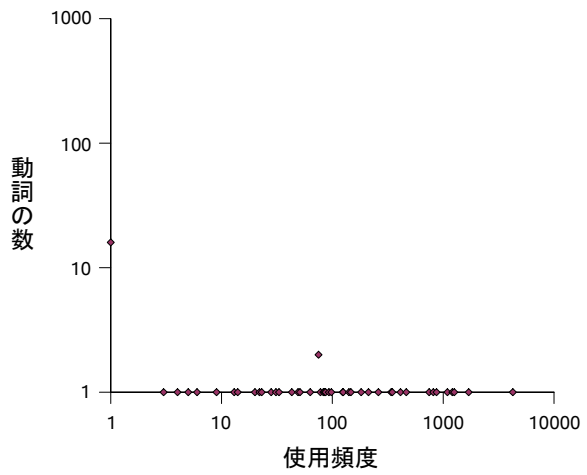
複合動詞の生産性と語彙的多様性



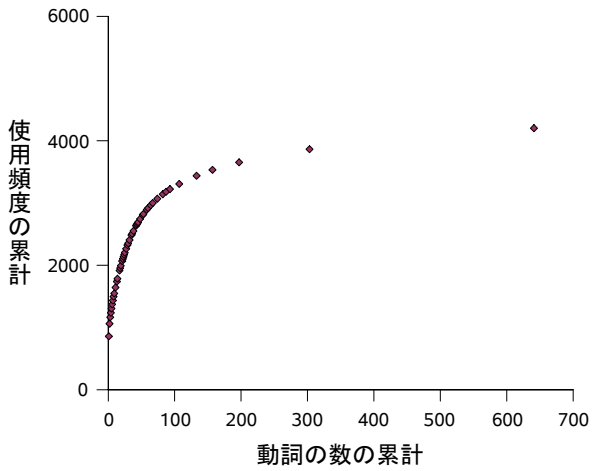
～かける (1)(641,4204,338)



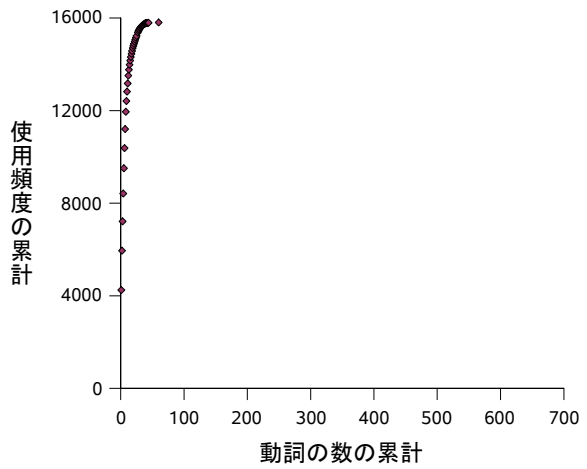
～かける (2)(60,15807,16)



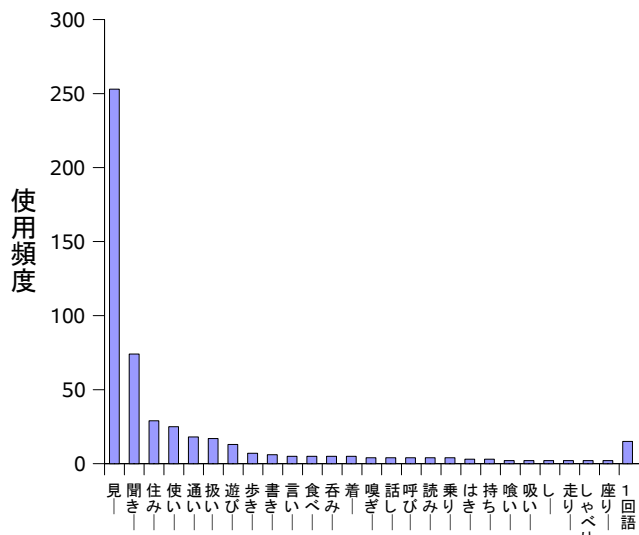
～かける (1)(641,4204,338)



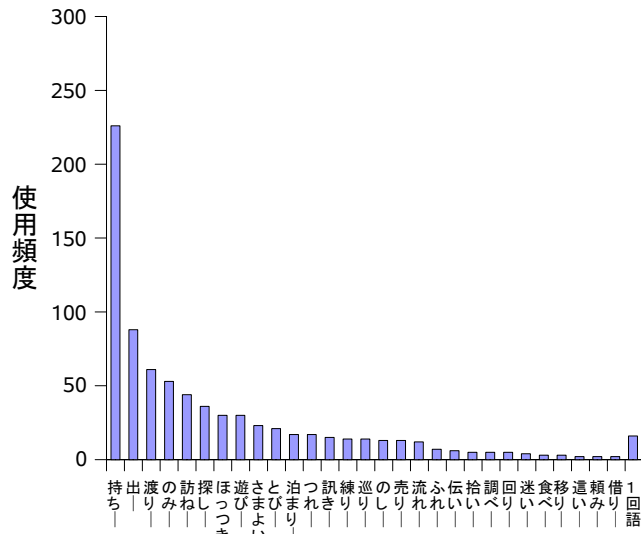
～かける (2)(60,15807,16)



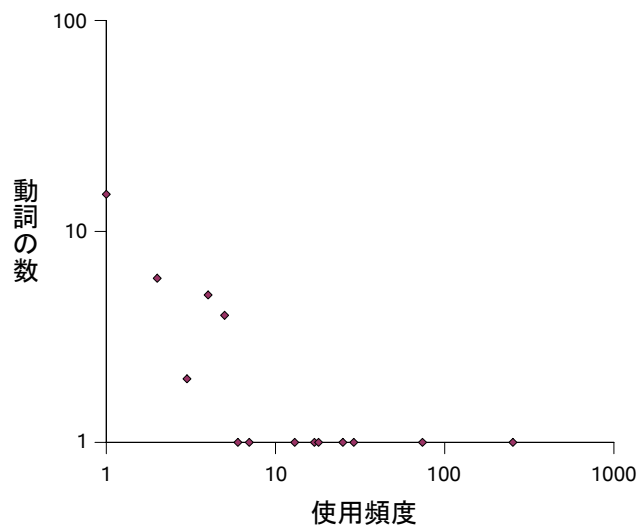
～なれる (41,515,15)



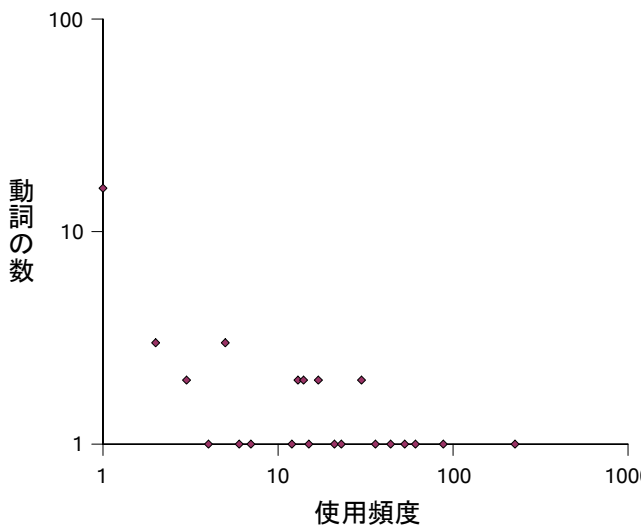
～あるく (45,787,16)



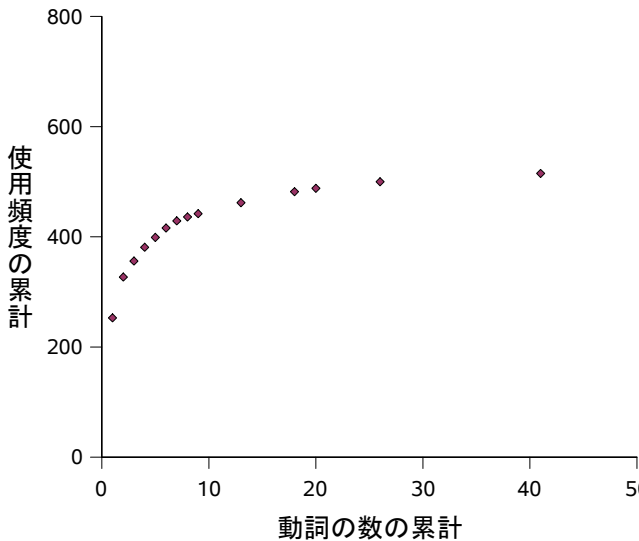
～なれる (41,515,15)



～あるく (45,787,16)



～なれる (41,515,15)



～あるく (45,787,16)

