

コーパスを利用した現代英語における 二要素複合形容詞の分析

A Corpus-based Analysis of Simple Compound Adjectives in Present-day English

西部 真由美

NISHIBU Mayumi

愛知大学国際コミュニケーション学部

Faculty of International Communication, Aichi University

E-mail: mnishibu@vega.aichi-u.ac.jp

Abstract

This study tries to explore the characteristics of attributive simple compound adjectives, which consist of two elements of words or roots and function as adjectives modifying nouns. The analysis, based on the qualitative and quantitative data retrieved from British National Corpus, has been conducted to examine the following questions: 1) what is the frequency of simple compound adjectives in present-day British English, 2) what are the most frequent ones, 3) in what text genres do they frequently appear, 4) what sort of formal combinations are typical ones, and 5) what elements are productively combined to various types of elements? The results show that simple compound adjectives account for approximately 93% of all compound adjectives, and that the most frequent part-of-speech combinations are “noun-noun” and “adjective-noun.” Answers to the rest of the questions are also provided and discussed in the following main sections.

1 複合形容詞とは

複合形容詞 (compound adjectives) とは、広義では2つ以上の語 (または語根) で構成されている形容詞の働きをする複合語のことである。複合形容詞は、構造的には色々な組み合わせが可能で、形容詞-名詞 (e.g., *long-term*) ・ 名詞-名詞 (e.g., *part-time*) ・ 名詞-過去分詞 (e.g., *state-owned*) ・ 形容詞-過去分詞 (e.g., *ready-made*) ・ 副詞-現在分詞 (e.g., *ever-*

changing)・副詞-過去分詞 (e.g., *well-known*) などがその例である。さらに、複合形容詞には「形容詞-名詞 -ed」という構造を持つもの (*dark-haired, giant-sized, open-minded, warm-hearted, etc.*) もある。これらの語尾は屈折しないはずの名詞に脱名詞辞 -ed を加えて形成されるので、動詞の過去分詞形とは区別される。

複合語と混同し易いのが、接辞が語根に接合した語である。接頭辞や接尾辞は定義上、単独では語とならず、他の語根に接合して意味を成す拘束形態素 (bound morpheme) である。接頭辞や接尾辞が語根に接合した語は複合語ではなく、派生語 (derivatives) と認定される。したがって、以下の例で下線部の接頭辞が付いた語 (e.g., *non-governmental, anti-social, pre-school*) や接尾辞が付いた語 (e.g., *schoolboy-ish, media-wise, left-ward*) は、通常複合語には含めない。

さらに、接辞に似た紛らわしい語根も存在し、例えば *Anglo-, bio-, neuro-, mono-, multi-, socio-, super-, proto-, pseudo-, -graphic, -phobia* のような、ギリシャ語やラテン語といった古典的な言語を語源とする語根は、連結形 (combining form) と呼ばれる。連結形は、拘束語根 (bound root) で通常は複合語の最初か最後の位置を占めるので接辞に似ているが、連結形は 1) 単独の語と同様にはっきりとした意味を持ち、2) 自由語根 (free root) がなくても 2 つ以上の連結形同士で単語を形成することができ (e.g., *Anglo-phobia*)、3) 接辞と接合して単語を形成できる (e.g., *non-graphic*)、という点で接辞と異なっている。このような連結形を含む複合語は、言語学的には他と区別するために新古典複合語 (neo-classical compounds) と称される。

要するに、複合語の構成要素として接辞は認めず、連結形は認めるということになるのだが、当の接辞と見なされる形態素について必ずしも見解が一致している訳ではない。

複合形容詞は、限定用法で用いられることが圧倒的に多い (Biber et al. 1999: 535)。さらに、限定用法ではほとんど全てがハイフンでつながれており (Sinclair et al. 1990: 83)、同じ語でもイギリス英語の方がアメリカ英語よりもハイフンで繋がれる場合が多いことが指摘されている (Quirk, et al. 1985: 1569)。

本稿では、2 つの語 (および連結形) がハイフンで結ばれてできた限定用法の複合形容詞について、その全体像を探る。まず、名詞を限定修飾する複合形容詞はどのくらいの頻度でどのようなテキストのジャンルに現れるのか。どのような構成要素の組み合わせがあり、いかなる意味的・構造的特徴があるのか。典型的あるいは生産的な構成要素は何か。また、分析結果から、複合形容詞の語形成過程に対してどのような示唆が得られるのか。これらの点について述べて行く。

2 コーパスと複合語

一般的に、大規模コーパスでは複合語という品詞タグは設けられていないので、複合語に該当するものを瞬時に検索して網羅的に拾い上げることはできない。しかし、ハイフンで結ばれた限定用法の複合形容詞は、ハイフンを手掛かりにすればコーパスから該当例を容易に抽出することが可能である。また、限定用法では名詞句の前に置かれる点を利用して、ハイフンで結ばれた任意の形容詞とその後に任意の名詞句が来るもので検索すれば、高い精度で抽出することができる。

しかしながら、データの抽出や分析を進める上では不都合も生じる。まず、ハイフンで結ばれた複合語で形容詞の品詞タグが付いたものを検索すると、接辞を含むものも抽出されてしまう。実証的な量的分析を行うとなると、接辞と見なす形態素を定めて別途除外する必要がある。

そこで、本稿ではできる限り少なく接辞を設定した。その理由は、広義に接辞を設定すると、語形成論に関する先行研究で議論されてきた項目（例えば語源がギリシャ・ラテン語の形態素や、判断が分かれる *-like* など）の例が分析対象から除外されてしまうためである。代表的な英語辞典を参照し¹⁾、その全ての辞典において接頭辞あるいは接尾辞という品詞の記載がある形態素は接辞として扱い、それらが構成要素となっている検索例は除外した。接辞が連結形か議論が分かれているものについては、連結形の可能性がある形態素であれば連結形に含めることとし、連結形としての記載が最も多く接辞の定義が最も狭い *Oxford Advanced Learner's Dictionary (8th ed.)* に準拠して、検索例の取舍選択をした。ハイフンでつながれた構成要素で、接辞として扱い削除の対象としたものを、表 1 にまとめた。

表 1 接辞と見なして除外した要素

左側構成要素	右側構成要素
接頭辞	接尾辞
ante-, ant(i)-, be-, co-, de-, demi-, dis-, en-, ex-, extra-, hyper-, hypo-, infra-, inter-, intra-, mis-, non-, out-, over-, post-, pre-, pro-, re-, semi-, sub-, trans-, ultra-, un-, under-	-able, -esque, -ful, -ing, -ed, -est, -ic, -ish, -ist, -less, -ly, -ous, -ship, -some, -ward, -wise, -y

本稿で分析した英語の電子化コーパスは、イギリス英語の大規模コーパスである British National Corpus（以降は BNC と表記する）の内の書き言葉（written）の部分（1 億 35 万 1400 語）である。データの抽出には、Bringham Young University の Mark Davies の作成による BYU-BNC を使用した。

データ処理の手順については、BYU-BNC において、ハイフンで結合している 2 つ以上

の構成要素を持つ形容詞でかつ直後に名詞が来る例、および直後に形容詞が来る例（その後名詞が置かれている場合が殆ど）に該当する例を抽出し、表計算ソフトのエクセルに移してから、バグや該当しない例を手作業で取り除き、検索例を解析した。

3 コーパスで見た限定用法の複合形容詞

3.1 検索例総計

BNC コーパスの中に、ハイフンで結ばれた 2 つ以上の要素で構成される形容詞は約 32 万件あり、その内の約 85%（約 27 万件）が限定用法で使用されていた。限定用法の例では、被修飾語の名詞が直後に来る例が約 90% で、残りの 10% には複合形容詞の後に「形容詞＋名詞」の構造を持つ名詞句が続いていた。

コーパスに現れた回数は粗頻度（raw frequency）と呼ばれるが、コーパスのサイズによって異なる値が出るため、その語の一般的な頻度を知るためには不都合である。別のコーパスで得られた頻度であっても同じ尺度で測れるように百万語あたりの出現頻度（PMW）に調整する方法が広く使用されている。今回の分析で得られた限定用法の複合形容詞は 2735 語（PMW）である。これと同頻度で現れる単語の例を挙げると、*we, their, been, has, would* など、これらは英語の単語全体において頻度順で上位 37 位から 40 位になる（Leech, et al. 2001: 181）。イギリス英語を使用する人々は、これらの単語に遭遇するのと同程度の頻度で何れかの限定用法の複合形容詞に遭遇するということになり、全体で見ると複合形容詞は高頻度で現れる現象であると言えよう。

構成要素の数別に見ると、2 要素で結合するものが最も多く（PMW 2549.8）で全体の 93% 以上を占めている。次いで 3 要素が 6.2%（PMW 169.5）、4 要素以上になると 0.6%（PMW 15.6）と格段に少なくなっている。複合形容詞は 2 要素から構成されるものが殆どであり、3 要素以上から成るものは稀である現象であると言えよう。

以降の項では、特に 2 要素で構成される複合形容詞について、テキストジャンル別に見た分布（3.2）、トークン・タイプ頻度（3.3）、高頻度語（3.4-5）、品詞の組み合わせ（3.6）、多種類の要素と結合する要素（3.7）について詳細に調べることにする。

3.2 テキストジャンル別頻度

複合形容詞は、多くの情報をより少ない数の語で簡潔に表現するのに有効な言語形式である。例えば以下の（1a）のような関係代名詞の節で表す内容を、（1b）では 1 つの複合形容詞で表すことができる。

- (1) a. This is a task that will consume time.
 b. This is a time-consuming task.

語数を節約できる言語形式である複合形容詞は、当然ながら活字数に制限のあるテキストジャンルで多用されるのではないかと予測ができる。そこで、テキストのジャンル別に複合形容詞の分布について調べてみた。

BNC-BYU の書き言葉のコーパスは、学術文献 (academic) ・ 非学術文献 (non-academic) ・ フィクション (fiction) ・ 雑誌 (magazines) ・ その他 (miscellaneous) (広告、パンフレットなど) ・ 新聞 (newspapers) の 6 つのジャンルに分類されている。各ジャンルのコーパスのサイズが異なり、各分野の租頻度では比較できないため、百万語あたりに換算した頻度 (PMW) で比較してみる。次の図 1 は、各テキストジャンルにおける 2 要素複合形容詞の頻度を示している。

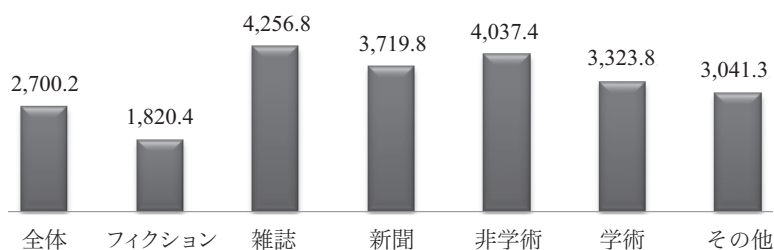


図 1. 二要素の複合形容詞——テキストジャンル別頻度

図 1 が示す通り、2 要素複合形容詞は、「雑誌」に現れる頻度が最も高く (約 PMW 4260)、次いで「非学術」(約 PMW 4040)、「新聞」(約 PMW 3720) の順に高い頻度で現れている。一方、「フィクション」では最も低い頻度 (約 PMW 1820) になっている。この数値は、概ね予想と合致し、活字数に制限がある分野ほど複合形容詞の頻度が高くなることを実証している。

3.3 トークン・タイプ頻度

コーパスから得られた該当例の頻度をまとめたものが次の表 2 である。表 2 では、検索例の租頻度 (トークン数) とタイプ数の両方の値が示してある。タイプ数とは異なり語の数のことであり、何種類の複合形容詞が出現しているかを表している。またタイプ・トークン割合 (type-token ratio (TTR)) は、数値が 1 に接近するほど多様な語が現れたことを意味し、逆に数値は低くなると特定の語が繰り返し現れていることになる。

表 2 複合形容詞の頻度

	租頻度	PMW	トークン割合	タイプ語数	タイプ割合	TTR
2要素	255872	2549.8	93.2%	57031	93.0%	0.22
全体	274449	2734.9	100.0%	61350	100.0%	0.22

注：トークン割合 = 2要素複合形容詞の租頻度 ÷ 複合形容詞租頻度総数

タイプ割合 = 2要素複合形容詞タイプ数 ÷ 複合形容詞のタイプ総数

TTR = タイプ数 ÷ トークン数

表 2 の通り、トークン数、タイプ数ともに 2 要素複合形容詞は複合形容詞全体の約 93% を占め、圧倒的な多数となっている。また、トークン数が約 25 万 6 千に対してタイプ数は 5 万 7 千となっており、タイプ・トークン割合が 0.22 であることから、繰り返し出現するタイプも存在することが分かる。

更に詳細に分析するために、コーパスの中で 1 回しか現れなかった例 (hapax-legomena) について見てみる。大規模コーパスを使用して語を分析する場合、頻度が 1 であるものはバグや一個人しか使わない奇異な語と見なされて除外されることが多い。しかし、複合語の分析の場合には、このような語は一個人だけが使った語、新たに生み出された語であると考えられるので、その中に何らかの語形成の規則が見出される可能性があると考えられている。次の表 3 に、コーパスに現れた租頻度 1 の語の総数、複合形容詞の総数に占める租頻度 1 の語の比率を示した。

表 3 租頻度 1 (Hapax-legomena) の複合形容詞の頻度

	租頻度 = タイプ数	PMW	割合	トークン比率	タイプ比率
2要素	34972	348.5	92.2%	13.7%	61.3%
全体	37940	378.1	100.0%	13.8%	61.8%

注：割合 = Hapax-legomena の租頻度 ÷ Hapax-legomena 租頻度総数

トークン比率 = Hapax-legomena の租頻度 ÷ 複合形容詞の租頻度総数

タイプ比率 = Hapax-legomena のタイプ数 ÷ 複合形容詞のタイプ総数

表 3 のタイプ比率が示す通り、複合形容詞のタイプの内の 6 割以上が、1 億語のコーパスの中で実に 1 回しか出現しないのである。しかしその一方で、トークン比率を見ると、租頻度 1 の語の頻度は総数の 14% 以下になっている。このことから、租頻度 2 以上の 4 割弱のタイプの複合形容詞の中に、極めて高頻度で現れるタイプがあると推測される。

大量の租頻度 1 の語がある一方で高頻度語も存在するという現状から複合形容詞の語形成過程について考えてみると、高頻度なものは個別に語彙として脳内に登録されているが、6 割以上のタイプは一語彙として人々の脳内に記憶されているとは考え難く、その多くは意味的・形態的規則や手がかりによって随時造られ、そして理解されると考えるのが

妥当であろう。

3.4 複合形容詞全体での高頻度語

それでは、高頻度で出現した限定用法の複合形容詞について見てみよう。検索例中の高頻度語を次の表4にまとめた。

表4 高頻度の複合形容詞

上位1-16	PMW	上位17-32	PMW	上位33-48	PMW
*-year-old ³	56.9	would-be	7.4	left-hand	4.3
*-year	48.0	first-class	6.7	up-to-date ³	4.0
*-century	45.5	left-wing	6.6	inner-city	3.9
long-term	36.9	small-scale	6.2	built-in	3.8
short-term	26.1	state-owned	6.0	present-day	3.7
so-called	24.8	one-day	5.7	full-scale	3.6
*-day	23.7	right-hand	5.6	one-off	3.4
working-class	16.9	middle-aged	5.4	high-speed	3.4
full-time	16.1	long-standing	5.3	wide-ranging	3.4
part-time	15.0	second-hand	5.3	last-minute	3.3
middle-class	11.3	decision-making	5.0	medium-sized	3.2
well-known	11.3	in-house	4.9	far-reaching	3.2
socio-economic ^準	10.9	Anglo-Saxon ^準	4.9	in-depth	3.1
large-scale	10.3	three-dimensional	4.5	two-way	3.1
day-to-day ³	8.8	object-oriented	4.4	*-month-old ³	3.1
right-wing	8.2	short-lived	4.4	home-made	3.1
old-fashioned	8.1	in-service	4.4	California-based	3.0

注：*は数・序数を表す。；³は3要素から成るものを表す。；^準は連結形で構成される新古典複合形容詞を表す。；頻度は百万語あたり（PMW）の数値である。

表4において、頻度はコーパス中に現れた租頻度ではなく、百万語あたりに換算した数値（PMW）で示してある。最高頻度の複合形容詞（数・序数に *-year-old* が付いたもの）は約 PMW 57 で、この値は英語全体では高頻度語上位 2000 語程度に相当する（Leech, et al. 2001: 195）。頻度が PMW 20 で上位 5000 語程度であり、表4の最下位に挙がっている語の PMW 3 では、英語全体で見れば決して高い頻度とは言えない。

構成要素の数で見ると、高頻度語の殆どが2要素であるが、3要素も4つ含まれている。3要素の複合形容詞の内の2つは数と組み合わせる年齢や月齢を表す *-year-old* と *-month-old* で、残りは慣用句の *day-to-day*（その日暮らしの）と *up-to-date*（最新の）である。前者の2つは各構成要素の表す意味を総計すれば全体の表す意味が分かり、意味的に「透明」（transparent）であるが、後者2つは各構成要素の意味の総計とは異なる比喩的な意味あるいは特殊化した意味を持ち、意味的に「不透明」（opaque）である。この視点で表4にある2要素の高頻度語を見てみると、意味的に「透明」なものが高頻度語の大部分を占めていると言える。

次に、高頻度語の構成要素の品詞別組み合わせを見てみよう。まず、高頻度語48項目の内の19が「形容詞-名詞」の組み合わせで、最も高い割合を占めている。この形式は要素自体では名詞句を形成している。この組み合わせの例は *long/short-term*, *right/left-wing*, *large/small/full-scale*, *present-day*, *high-speed* などである。名詞句が後続する名詞を限定修飾する現象 (adjectival nouns) は複合語に限らず英語では広く見られる現象である。これを複合形容詞とする見解 (Bauer: 1983) と複合名詞と考えて複合形容詞に認定しない見解 (Conti: 2010) があるが、判断は困難で現在でも議論が分かれている。

別の種類として、*old-fashioned*, *middle-aged*, *medium-sized* のような組み合わせが挙げられる。これらはそれぞれ *old fashion*, *middle age*, *medium size* という日常的に頻用される名詞句が存在するので、その句全体に脱名詞辞 *-ed* が付与されて形容詞化したと考えれば [*old-fashion*]-ed, [*middle-age*]-ed, [*medium-size*]-ed と分析され、要素が組み合わせられて新しくできた複合語というよりはむしろ名詞句から派生してできた形容詞であると解釈できる。

別の組み合わせとして、*in-house*, *in-service*, *in-depth* が挙げられる。これらはハイフンを取り除いた形では、「前置詞+名詞」で構成される前置詞句であり、形容詞や副詞の働きをする。このような複合形容詞は前置詞句の場合と意味的变化はなく、限定修飾するためにハイフンが挿入されて複合形容詞として定着したものと考えられる。また同様に、「動詞+副詞的小詞」の構造を持つ *built-in* もハイフンがない状態では句動詞として定着している。限定修飾のために句動詞から派生して形容詞になったと考えられる。

さらに、右側要素が動詞の分詞形 (または過去形) になっているものがある。その例は、過去分詞で終わる *state-owned*, *California-based* あるいは現在分詞で終わる *long-standing*, *wide-ranging* などである。この組み合わせを持つ複合形容詞は、ハイフンを取り除いた場合、要素をそのまま羅列しても統語的・意味的に成立しない場合があり、パラフレーズをして書き換える必要がある。例えば *state-owned* は、(something which is) owned by the state と解釈され左側要素の名詞は行為者であるのに対して、*California-based* は (something which) is based in California と解釈されカリフォルニアは場所を表す前置詞句の目的語となる。現在分詞の例 *long-standing* は (something which is) standing for a long time と解釈され、*wide-ranging* は (something which) ranges widely と解釈されるので、左側要素の *long* や *wide* は副詞的な意味を表すことになる。

このように複合形容詞が表す意味について考えて行くと、今見た例では右側要素が意味上の主要語 (head) で、左側要素がそれを修飾するという関係で解釈できた。このように解釈できるものは、表4の項目の内では約4分の3に及ぶ。複合形容詞は「基本的に内心構造 (endocentric) である」(大石1988: 100) とする見解もこの結果から妥当だと言える。

最後に、意味的に年齢や時間経過を表すもの (*-year*, *-century*, *-day*, *-year-old*, *month-old*) が最も多いことと、すでに慣用句として定着している *so-called* (いわゆる), *well-known*

(有名な), *would-be* (自称の・～志願の) や新古典複合語である *socio-economic*, *Anglo-Saxon* も高頻度語に挙がっていることに言及しておく。

3.5 2要素複合形容詞の高頻度語と修飾される名詞

表4で見た通り、高頻度で現れる複合形容詞は殆どが2要素で構成されていた。そこで、2要素複合形容詞だけの高頻度20位までを取り上げ、修飾を受けた名詞の例も見て行くことにする。高頻度20位までの2要素複合形容詞には意味的・構造的に共通点が見出せるので、次の表5にグループにまとめて示した。

高頻度語は、新古典複合語を除くと、意味的に「時間」・「規模」・「社会層」・「様態」に分けることができ、表す意味が同じであれば同様の構造を持つ傾向が見られる。また、高頻度語には対義語のペア (*long/short-term*, *full/part-time*, *large/small-scale*, *right/left-wing*) が含まれていることも特徴的である。また、社会層 (class) を表す複合形容詞が高頻度であることは英国の社会事情を反映しているものと考えられ、米語コーパスでは異なる結果が出る可能性もある。

表5 2要素の複合形容詞——高頻度語20

	構成	複合形容詞	PMW	%	修飾される名詞
時間	数字-名詞	*-year	48.0	4.6	girl, son, daughter, boy, 人名
		*-century	45.5		England, French, house
		*-day	23.7		visit, week, seminar, tour
	形容詞-名詞	long-term	36.9	3.7	unemployment, future, plan
		short-term	26.1		interest, contract, benefit
		full-time	16.1		employment, job, education
		part-time	15.0		farmer, employment, job
規模	形容詞-名詞	large-scale	10.3	0.6	production, industry, map
		small-scale	6.2		study, producer, industry
社会層	数/形容詞/名詞-名詞	working-class	16.9	1.9	people, family, culture
		middle-class	11.3		woman, people, family
		first-class	6.7		cricket, honour, match
		right-wing	8.2		party, conservative
		left-wing	6.6		politics, labour
様態	副詞-過去分詞	so-called	24.8	1.4	expert, friend
		well-known	11.3		fact, example, name, company
	形容詞-名詞 ed	old-fashioned	8.1	0.3	rose, style, clothes
	名詞-過去分詞	state-owned	6.0	0.2	company, enterprise
	助動詞 -be	would-be	7.4	0.3	emigrant, buyer, assassin
新古典	連結形-連結形	socio-economic	10.9	0.4	group, background, status

注：* は数・序数を表す。；% は2要素複合形容詞トークン総数に占める割合。

また、2要素複合形容詞の頻度総数に占める割合を見ると、高頻度語で左側要素が数値になっているものが全体の4.6%であり、右側要素では名詞であるものを合計すると最も大きな割合（約1割）を占めている。

なお、表には示されていないが、*so-called* は直後に名詞が来る割合が少なく、形容詞＋名詞が後続する例が顕著であった。

3.6 構成要素の品詞の組み合わせ

次に、2要素複合形容詞の右側要素と左側要素はどのような品詞の組み合わせになっているのかを調べた。手順は、品詞解析ツールである Penn Treebank を使用して自動で各要素別にタグ付した後に、エクセルを使って頻度の高い順にソートした。なお、約2%の要素は正しくタグ付けされていない可能性があり、特に動詞の ed 形と ing 形が形容詞に分類されるものがあるので、得られた結果は厳密な数値ではないが、大方の傾向を見るには適切であると考えられる。

品詞の組み合わせで高頻度のものを挙げ、2要素から成る複合形容詞の全体に占めるトークン割合およびタイプ割合を次の表6に示した。なお、複合形容詞の主要語 (head) は右側構成要素であるものが多いため、右側構成要素の品詞でグループ分けをした値と、左右両要素の組み合わせでグループ分けした値を示した。

表6 構成要素の品詞の組み合わせ割合

右側要素	トークン	タイプ	組み合わせ	トークン	タイプ
-名詞	43.4	38.9	名詞-名詞	14.6	18.0
-形容詞	20.9	21.6	形容詞-名詞	12.9	8.3
-動詞 ed	20.8	24.3	名詞-動詞 ed	10.7	13.2
-動詞 ing	8.0	8.5	名詞-形容詞	10.4	10.3
-前置詞	2.0	3.2	数-名詞	8.7	7.6
-名詞 ed	1.6	0.1	形容詞-動詞 ed	5.7	5.1
-副詞	1.0	1.1	名詞-動詞 ing	4.8	5.6
			形容詞-形容詞	3.9	4.6
			副詞-名詞	3.4	1.3
			副詞-動詞 ed	3.0	3.5
			副詞-形容詞	2.3	3.1
			名詞-前置詞	1.1	2.2

注：トークン＝トークン数÷2要素複合形容詞総トークン数（%）

タイプ＝タイプ数÷2要素複合形容詞総トークン数（%）

動詞 ed は規則動詞・不規則動詞の過去形・過去分詞形を示す。

右側要素では「名詞」がトークンとタイプの両方で40%前後を占めて最多で、次いで、「形容詞」（トークン、タイプ：20.9%、21.6%）と「動詞 ed 形」（20.8%、24.3%）の順に頻出する品詞になっている。反対に、「名詞 ed」（1.6%、0.1%）や「副詞」（1.0%、1.1%）が占める割合は小さい。

左右両要素の組み合わせでは、「名詞-名詞」が最も多く（トークン、タイプ：14.6%、18.0%）、次いで「形容詞-名詞」（12.9%、8.3%）、「名詞-動詞 ed」（10.7%、13.2%）、「名詞-形容詞」（10.4%、10.3%）の順で、大きな割合を占めている。各組み合わせの具体例は、「名詞-名詞」では *client-side, steam-age, cost-effectiveness, coal-gas* など、「形容詞-名詞」では *short-termism, large-screen, old-flame, present-generation* など、「名詞-動詞 ed」では *part-used, California-invented, castle-adorned, self-centred* など、「名詞-形容詞」では *soccer-illiterate, self-contradictory, tax-friendlier, space-visual* などである。

高頻度語では *well-known*（副詞-動詞 ed）や *in-house/depth/service*（前置詞-名詞）が挙がっていたものの、品詞の組み合わせとして全体で見ると少ない組み合わせであることが分かる。

3.7 多種類の要素と結合する要素

構成要素の中には、他種類の要素と結びついて複合形容詞を形成するものがある。このような要素は一般的には強い造語力がある、あるいは生産性（productivity）が高いものの候補に挙げられてきた。その特質を測る指標として、結合する要素のタイプ数の多さ、租頻度1の複合形容詞の中だけで見た結合するタイプ数の多さを見てみよう。次の表7に多くのタイプと結合した構成要素をまとめた。要素を右側と左側で分け、結合した要素のタイプ数、総タイプ数に占める割合、そのうち租頻度1であったものの割合が示してある。

まず、表7の左右の要素全体を比較すると、右側要素の方が結合するタイプ数が大きくなっていることが分かる。これは、右側要素を基軸に多様な左側要素が結合することを表し、複合形容詞では基本的には右側要素が主要語（head）である特徴を反映していると言える。

それでは左側の要素から詳しく見て行こう。まず、連結形（*self-, multi-, quasi-*）が総タイプ数の合計7%強を占めているが、右側要素には連結形は挙がっていないことから、連結形が語頭に来てその後多種類の要素が来るパターンが多いことが分かる。また、数値と数に関わる語が合計約17%で、最も大きな値となっている。その他には形容詞・副詞・前置詞に多タイプと結合する要素が含まれているが、前置詞は他と比べるとタイプ数は小さい。そして名詞は該当するものがない。個別に見ると、*self-, well-*（ともに総タイプ数の3.5%）が最も多種類の要素と結合し、次いで *half-, multi-*（2.8%）、*high-*（2.5%）となっている。

右側要素を見てみると、まず総タイプ数の約28%が動詞 ed と結合しており、次いで名詞で約19%となっている。特に名詞は数を左側に伴う時間を表す名詞（*-year, -day, -page, -minute*）が多い。名詞 ed、動詞 ing、形容詞、小詞（*-up, -off*）にも多タイプと結合する要素がある。個別に見ると、総タイプ数の10%以上を占める要素が2つ（*-based, -like*）

表7 多種類の要素と結合する要素

左側		タイプ	%	HL%	右側		タイプ	%	HL%	
連結形	self-	473	3.5	33.2	動詞 ed	-based	1554	14.1	49.9	
	multi-	381	2.8	53.0		-related	359	3.2	53.8	
	quasi-	148	1.0	66.2		-born	242	2.2	57.4	
数	one-	1278	9.4	50.6		-oriented	195	1.8	63.6	
	--five-					-backed	159	1.4	53.5	
	all-	238	1.8	55.0		-controlled	150	1.4	59.3	
	half-	386	2.8	62.4		-led	195	1.8	65.6	
	single-	262	1.9	13.4		-covered	170	1.5	55.9	
	double-	184	1.4	66.8		名詞 ed	-shaped	301	2.7	56.1
形容詞	high-	346	2.5	46.0			-sized	246	2.2	64.6
	low-	247	1.8	39.3	-coloured		230	2.1	49.6	
	long-	289	2.1	58.8	-faced		186	1.7	55.4	
	near-	190	1.4	66.8	形容詞	-like	1326	12.0	60.2	
	副詞	red-	166	1.2		53.6	-free	319	2.9	54.2
		white-	152	1.1	54.6	動詞 ing	-looking	503	4.6	60.8
black-		151	1.1	58.9	-making		167	1.5	52.1	
前置詞		well-	470	3.5	41.5	名詞	-style	567	5.1	72.7
	ever-	184	1.4	61.4	-type		560	5.1	78.4	
	newly-	188	1.4	62.2	-member		198	1.8	38.4	
	much-	144	1.1	52.1	-level		176	1.6	50.6	
	小詞	in-	112	0.8	58.0		-year	175	1.6	30.9
		off-	70	0.5	47.1		-day	152	1.4	41.4
小詞					-page		148	1.3	54.7	
					-minute		148	1.3	38.5	
小詞	in-	112	0.8	58.0	小詞	-up	256	2.3	48.0	
	off-	70	0.5	47.1		-off	157	1.4	48.4	

注：% = タイプ数 ÷ 2要素複合形容詞の総タイプ数

HL% = hapax legomena のタイプ数 ÷ タイプ数 (%)

存在する。これらは実に1300種類を超える要素と結合しているが、*-based* は地名や位置を表す語が左側に結合するので意味的には限られた範囲の要素と結合する。多タイプと結合する要素として他には、*-type*, *-style*, *-looking* (約5%) が挙げられる。

次に、表7の租頻度1のタイプ数とその要素の総タイプ数に占める割合を見てみよう。要素の殆どが50-70%に収まっているのに対して、例外的に割合が高いもの (*-style*, *-type*) と低いもの (*self*, *single-*) が見つけられる。租頻度1の割合が高い程、多種の要素との結合が可能であることを表し、この割合が低い程固定的で語彙として定着している組み合わせが多いことを表している。したがって、*-style*, *-type* は多種多様な要素を左側に結合させ固定的な組み合わせは少ないが、*self*, *single-* は比較的定着した組み合わせが多いと推測できる。また、数と結びつく右側要素 (*-minute*, *-member*, *-year*) も租頻度1の割合が30%台で低くなっているが、これは別途で解釈する必要がある。数は無限なので数のタイプ母数自体が大きく、また2回以上使用される数も多いので、租頻度1の割合は結果的

には低い値になっていると考えられる。

また別の観点から見れば、「接辞」として扱われている場合も少なくない要素 (-like, -style, -type, -looking) は、タイプ総数の約4.5%以上を占める多タイプと結合する要素であることが分かり、タイプ数が多いけれども -based や -minute, -member, -year は場所や数といった意味的な制限があるために「接辞」として一般的に受け入れられていないのではないかと考えられる。

4 おわりに

大規模英語コーパスを利用して2要素複合形容詞の特徴を質的・量的の両側面で分析することが可能となった。検索例の分析から、複合形容詞の93%以上は2つの要素から構成されており、「名詞(数詞)+名詞」の組み合わせや「形容詞+名詞」が最多であることが分かった。また、特定の高頻度語が極端に多く現れる一方で、6割以上のタイプの複合形容詞は1回しか出現していないという事実も明らかになった。複合形容詞の構成要素を単独で見た場合には、多種類の語(語根)と結合する要素は、辞書などで接辞として扱われる傾向が認められた。本稿の様な量的分析はこれまでになされておらず、分析から得られる数値が接辞の規定や複合語の定義づけに役立つ指針となることが期待される。

注

- 1) 代表的な英語辞書として参考文献 [Dictionaries] に挙げた辞書を参照した。

参考文献

- Bauer, Laurie. (1983) *English Word-formation*, Cambridge: Cambridge University Press: Cambridge.
- Biber, Douglas., Stig Johansson, S., Geoffrey Leech, Susan Conrad & Edward Finegan. (1999) *Longman Grammar of Spoken and Written English*. Harlow, England: Pearson Education.
- Conti, Sara. (2006) *Compound Adjectives in English: A Descriptive Approach to Their Morphology and Functions*, Doctoral Dissertation, University of Pisa.
- Jackson, Howard. (2002) *Lexicography: An Introduction*. London: Routledge.
- Leech, Geoffrey, Paul Rayson, and Andrew Wilson. (2001) *Word Frequencies in Written and Spoken English: Based in the British National Corpus*. Harlow, England: Pearson Education Limited.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. (1985) *A Comprehensive Grammar of the English Language*. Harlow, England: Longman Group.
- Sinclair, John.(ed.) (1990) *COBUILD English Grammar*. Glasgow, UK: Harper Collins.
- 大石強 (1988) 『形態論—現代の英語学シリーズ第4巻』 開拓社

Dictionaries

Cambridge Advanced Learner's Dictionary. (2003) Cambridge University Press.

Longman Dictionary of Contemporary English, 5th Edition. (2010) Pearson Education.

Oxford Advanced Learners' Dictionary, 8th Edition. (2010) Oxford University Press.

Oxford English Dictionary, 2nd Edition, Version 4 (CD-ROM) (2009) Oxford University Press.

『ランダムハウス英語辞典 CD-ROM 版 ver. 1.50』(2002) 小学館.

小西友七・南出康世編 (2002) 『ジーニアス英和大辞典 CD-ROM 版』 大修館.

井上永幸・赤野一郎編 (2007) 『ウィズダム英和辞典第 2 版』 三省堂.

小西友七・南出康世編 (2006) 『ジーニアス英和辞典第 4 版』 大修館.

竹林 滋・小島義郎・東 信行・赤須 薫編 (2005) 『ルミナス英和辞典第 2 版』 研究社.

Corpus

British National Corpus(BYU-BNC) Mark Davies at Brigham Young University: <http://corpus.byu.edu/bnc/>