

修辞学的手法による英語コロケーションの抽出

佐良木 昌

純心女子短大英米文化科/Saraki MT R&D
saraki@st.rim.or.jp

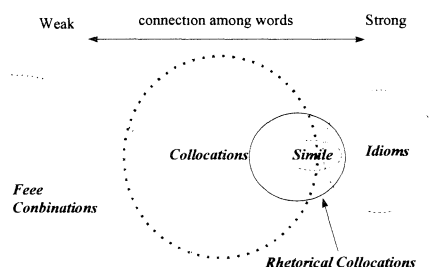
第一章 修辞学の導入

1. collocation の定義

BBI¹, COUBILD², CIDE³, OALD⁴などにおいて、collocation definition が様々になされているが、ここでは BBI 辞書の定義 [Benson et al. 1984] について考察する。

BBI 辞書では、Grammatical Collocations(GC) と Lexical Collocations(LC)とを区分する。さらに、GC の下位区分として G1~G8 (細区分を含む 25 パターン) を、LC の下位区分として L1~L7 を、それぞれ設け、全部で 33 パターンに分類している。*idioms* については、句全体の意味に、個々の語の意味が反映していないとの見解から、*Collocations* には含まれないとする。一方 *simile* については、collocations と idioms との中間にあって過渡的なもの (transitional) で、句全体としての意味に、個々の語の意味が部分的に反映しているとの見解から、これは collocations として採用している。

しかし BBI 辞書では、*simile* を、"miscellaneous"として扱っている。GC にも LC にも分類できない雑多なものとしている。本報告では、この BBI 見解を修辞学の観点から見直すことにより、*Rhetorical Collocations* を新たに構想した。



2. Rhetorical Collocations 修辞的連語

文法的・意味的に、語と語とが結合可能であると共に、単なる結合を越えて強固であることが collocations の成立条件である。free combinations—collocations—similes—idioms、これらにおける結合度に関して、結合強度の測定基準として、「頻度」「共起制限」「意味の特殊化」[中本 1997] が提案されている。また、free combinations まで含めて収集すると、網羅すべき collocations の数が爆発的に増大するが、このような副作用を招かぬためには、collocations の文法的な結合モデルや意味的な結合モデルの検討と共に、形態的な結合モデルの検討が必要である。

ここでは、修辞技法的な要素を、Rhetorical Collocations の基準として設定することにより、collocations の形態的な規定を見出しうることを、以下に示す。

まず、修辞学の観点を導入することにより、修辞的連語 *Rhetorical Collocations* の諸範疇を新たに構想した。

[A] 等位接続型連語

- α) 反対語の並列による連語
- β) 同義語⁵、類義語の並列による連語
- γ) 関連語の並列による連語
- δ) 同じ語の反復

[B] 接辞型連語

- α) 頭韻一致 (Alliteration) の連語
- β) 脚韻一致 (End Rhyme) の連語

[C] 直喩型連語 (強意の直喩 Intensifying Simile)

これら修辞的連語の諸範疇は、以下の五つの修辞技法[池田1992][市川1940][鍋島1962][Aronstein 1924][Suartengren1918][Deutschbein1932]に基づいて区分された。

- [1]対照 "Antithesis" [2]二詞一意 "Hendiadys"⁶
- [3]列举 "Catalog" [4]直喩 "Simile"
- [5]韻律 "Rhyme"

¹ "The Dictionary of English Word Combinations"1986

² "Collins Cobuild English Language Dictionary"1995

³ "Cambridge International Dictionary of English"1995

⁴ "Oxford Advanced Learner's Dictionary of Current English"1995

⁵ 同義語 Synonyms、類義語 Analogous words

⁶ 二詞一意 《『形容詞+名詞』などの意を「名詞+and+名詞」などの形で表わす法、例: bread and butter (=battered bread) / death and honor (=honorable death) / nice and cool (=nicely cool)》。【研究社新英和中辞典】

頭韻および脚韻は、形態的には接頭辞・接尾辞の一致である。修辞技法[1]～[4]による連語表現が、同時に[5]を兼ね備えると、すなわち韻を踏むと、その文体的効果が高まり、句のまとまりが強められる。範疇[A]が韻を踏む場合には、範疇[B]と重なり collocation 形成に対して共起制限を強める。また Rhyme によって直喩は、強意の直喩 *Intensifying Simile* となる(範疇[C])。

3. Rhetorical Collocations の抽出方法

抽出の方法としては、例えば

等位接続型連語 " N_1 and/or/but N_2 "

直喩型連語 "(as) ADJ as NP"

という対句の一般形式を記述することができる。この一般式について、頭韻句・脚韻句の形式を条件として設定することで、等位接続型連語で且つ接辞同一の修辭的連語の候補を大量に抽出することができる。例えば、次の条件を設定することで、頭韻句・脚韻句の形式を採る等位接続型連語を抽出する。

α 接頭辞同一(頭韻句 Alliteration)

$Prefix_1 = Prefix_2$

β 接尾辞同一(脚韻句 End Rhyme)

$Suffix_1 = Suffix_2$

γ 三つ組み(Triad) " N_1, N_2 and/or N_3 "

$Prefix_1 = Prefix_2 = Prefix_3$

$Suffix_1 = Suffix_2 = Suffix_3$

[例1]⁷例えば、等位接続型連語で且つ接尾辞同一(脚韻句)についての検索式は、脚韻データ-ation を例に採るならば、正規表現では以下のように記述できる。

[a-z]+tion and [a-z]+tion

この検索式で、例えばコーパスを検索すると、*democratization and decentralization* 民主化と分権
desertification and deforestation 砂漠化と森林破壊
など 80 以上の脚韻句候補が抽出できた。

[例2]また、トリアーデ形式は、等位接続型連語で且つ接尾辞同一の特殊例であるが、この修辭形式の連語を抽出するための検索式は、例えば、接尾辞-tion の一致するを抽出するための検索式は、正規表現では以下のように記述できる。

[a-z]+tion, [a-z]+tion, and [a-z]+tion

この検索式でコーパスを検索すると、*education, urbanization and professionalization*
高学歴、都市化および専門化
sanitation, nutrition and immunization
衛生設備、栄養摂取および予防注射
など接尾一致の連語候補多数を抽出できた。

⁷ 抽象名詞の場合、無冠詞名詞の対が、修辭型連語の一般的特徴の一つである。

第二章 抽出事例

1. 等位接続型連語／接辞型連語

α) 反対語の並列

・相反する言葉をつなぐ"X and Y" の連語の訳は、「NP₁ や NP₂」「(～の)V₁ や V₂ を行う」を基本とする。

| | |
|-------------------------------------|------------|
| <i>addition and deletion</i> | 追加や削除 |
| <i>connection and disconnection</i> | 接続や切断(分離) |
| <i>creation and destruction</i> | 創造や破壊 |
| <i>add and delete</i> | ～の追加や削除を行う |
| ・ 二字・四字の漢語に対応させて訳語を表記するものがある。 | |
| <i>increase and decrease</i> | 増減 |
| <i>positive and negative</i> | 正負 |
| <i>present or absent</i> | 出欠 |
| <i>directly and immediately</i> | 直接間接 |

β) 同義語・類義語の並列

[構造・階層・階級の上下]「XおよびY」と訳す。

section and subsection セクションおよびサブセクション
[行為や処理などの連鎖・運動・継続共起・同時進行など] 中点「・」を使って表記する。

| |
|---|
| <i>addition, subtraction and multiplication</i> |
| 加算・減算・乗算 |
| <i>stability, maneuverability and followability</i> |
| 安定性・機動性・追従性 |

γ) 関連語の並列

中点「・」を使って表記する。

| | |
|---------------------------------------|----------|
| <i>activation and initialization</i> | 起動・初期化 |
| <i>disconnection and reconnection</i> | 切断・再接続 |
| <i>extraction and purification</i> | 抽出・精製 |
| <i>plug and play</i> | 差込・動作させる |
| <i>select and activate</i> | 選択・起動する |

δ) 同じ語の反復

「XおよびX」と訳す。

| | |
|---|--------------|
| <i>compound devices and devices drivers</i> | |
| 複合デバイスおよびデバイス・ドライバ | |
| <i>key and key cylinder</i> | キーおよびキー・シリンダ |

2. 直喩型連語(強意の直喩)⁸

as good as gold

"ADJ as N" の形態のとき、形容詞 ADJ と名詞 N とが同じ頭文字を有しているときも、一まとまりの本範疇に入れる。 *as big as bushel baskets/as dangerous as drug firms/as high as Heaven/*

⁸ 強調の直喩が慣用表現になっているもの *busy as bees* 蜜蜂のように忙しい (頭韻) / *cold as ice* 氷のように冷たい / *still as a stone* 石のように動かない / *cool as cucumber* 落ち着き払って (頭韻)。頭文字一致の慣用表現 *dry as dust* 埃のように乾いた

第三章 修辞型連語辞書の効果

1. 効果事例

以下に修辞的連語辞書を用いた機械翻訳の具体例を従来の処理と比較して示す。

In one preferred embodiment of the present invention, five icons, each representing one of the subject matter categories, are displayed and may **be selected and activated** by the viewer manipulating the remote control.

従来の機械翻訳では、

「1つの本発明の好適な実施例において、5つのアイコン(内容カテゴリーの各々を表している一つ)は、表示されて、選んでもよくて、通って遠隔制御を操作している視聴者をアクティブにした。」と翻訳される。ここでは、

- 1) “**be selected and activated**” の and の解析に失敗し、助動詞が **activated** を支配せずと解析。
- 2) さらに、by the viewer manipulating the remote control が、受動態の意味上の主語であることを解析できず、the viewer を **activated** の目的語とし、能動態で過去形として解析してしまった。
- 3) したがって、by を副詞として解析することになり「通って」と訳出した。

これに対して修辞的連語辞書を使用した場合、同じシステムでの訳文出力は

「1つの本発明の好適な実施例において、5つのアイコン(内容カテゴリーの各々の表している一つ)は、表示されて、遠隔制御を操作している視聴者によって選択・起動されてもよい。」であった。

- 1) **selected and activated** が受動態で、may に支配されていると解析した。may の訳出も正しい。
- 2) by が受動態の動作主を目的語にとる前置詞として正確に解析がなされている。

[考察]

・修辞的連語辞書を参照することで、等位接続の句形態を一つの形態素として認識し、分離することなく構文解析を行うことができる。その結果、統語解析における誤りの発生を抑えることができる。

・さらに、辞書に格納されていない表現形態であっても、当該表現形態の接頭辞・接尾辞が同一であるという条件を加えることにより、その表現形態を一つの句形態として認識することが可能になる。

・上記の二つの方法により、分解的に解析すべきではない句形態について、その統語解析を回避することができる。解析を回避することにより、文全体の統語解析の精度を向上させることができる。

2. 副作用の検査

名詞句の一部と修辞的連語とが一致する場合

[事例1] 辞書の登録内容: **identification and synchronization**「識別および同期」

One of the information fields is typically designated for storing data, while other fields contain sector **identification and synchronization** information, for example.

「情報フィールドの一つはデータを格納するための典型的に指名される、例えば、他のフィールドはセクター識別および同期情報を含む。」

“sector **identification**”と“**synchronization information**”と解析すべきところを、“sector”+“**identification and synchronization**”+“information”と誤って解析した。

[事例2] 連語辞書 **addition and deletion**「追加や削除」

In complex operations, window **addition and deletion** may be performed multiple times before the display is refreshed.

「複雑な動作において、ディスプレイが活力を与えられる前に、ウィンドウ追加や削除は実行された複数回であってもよい。」

“window **addition and deletion**”とすべきところを、“window”+“**addition and deletion**”と解析した。

3. 今後の課題

副作用の抑制のためには、直前に冠詞あるいは名詞がある時など、無冠詞を特徴とする修辞的連語の参照条件を確定する必要がある。

英語・ドイツ語の修辞学分野では、修辞的連語を大量収集した過去の遺産が埋もれている[10]。これらを発掘・収集することも大切である。

【参考文献】

- [1] Benson et al. 1984 “The Dictionary of English Word Combinations” pp. xv-xxxv
- [2] 中本恭平1997「コロケーション再考 頻度、共起制限、意味の特殊化」英語表現研究第14号 pp.58-66
- [3] 池田拓朗1992『英語文体論』研究社 pp.150-223
- [4] 市川三喜1940『英語学辞典』Hendiadsの項 研究社
- [5] 鍋島能弘『文体美学』1962 御崎書林 pp. 148-181
- [6] Aronstein, P1924アーロンシュタイン『英語文体論』研究社英語学ライブラリ65 pp.10-14木原・土屋抄訳
- [7] Suartengren, Torsten H.1918“Intensifying Similes in English” AMS Press 1975
- [8] Deutschbein, Max1932 “Neuenglische Stilistik” 『名詞構文と英語』研究社英語学ライブラリー24 東田部分訳
- [9] Hegel,G.W.F. “Vorlesungen ueber die Aesthetik” 『美学』第3巻の下 pp.2184-2200 1981竹内敏雄訳 岩波書店
- [10] Willert,Hans 1911“Die alliterierenden Formeln der englischen Sprache” 『英米故事伝説辞典』1972 alliterationの項<参考>斎藤静の引用