

理論言語学と応用言語学が会うコーパス研究

深谷輝彦*

Corpus Studies as a Meeting Point between Theoretical and Applied Linguistics

Teruhiko FUKAYA

1. はじめに

「理論言語学と応用言語学が会う」という問題設定をするときに、最初に「理論言語学と何か」「応用言語学とは何か」という問いかけに答える必要がある。例えば、小池(2003)は理論言語学と応用言語学を次のように定義する。

理論言語学：言語構造をコントロールする本質的な原理・規則を構築することを目的とする。

応用言語学：言語理論も応用して、幅広い意味でのさまざまな言語行動の分析を行い、人間の言語行動によって立つ原理・規則を発見しようとするものである。

以上の定義に従い、研究分野を振り分けると次のようになるであろう。

理論言語学：音韻論，統語論（形態論も含む），意味論

応用言語学：日本語・外国語教育学，言語習得，社会言語学，語用論，言語接触，言語コミュニケーション，応用認知言語学，心理言語学，言語と脳，コーパス言語学・辞書学，教育工学，研究と測定

言語の形式，構造に焦点を置く理論言語学に対し，言語を使用している人間・社会・文化に関心がある応用言語学とすることができる。

さらに，理論言語学と応用言語学の違いを明らかにするために，理論言語学の生成文法が取り上げる文法的话题と応用言語学の言語教育が取り上げる文法的话题を比較する。

理論言語学(生成文法)：句構造，句構造と移動，英語助動詞のシステムにおける主要部移動，動詞句 (渡辺(2009)目次第2章より)

* 国際コミュニケーション学部 国際言語コミュニケーション学科

応用言語学(言語教育):文法規則の種類,形式と機能,タイプ対トークン,談話文法,
話し言葉・書き言葉文法

(Larsen-Freeman and Decarrico (2010) より)

生成文法は,言語能力,すなわち人類共通の普遍文法と言語ごとに異なる選択を行うパラメータの解明をめざして,上述の話題を取り上げている。他方,応用言語学は,文法規則に加えてどのようなコンテキストでその規則が使用されているのか,という言語運用面を重視する。例えば,過去完了形について,過去の出来事の時間的順位を表すという文法記述に終わらずに,出来事の背景や理由づけをするというテキスト形成機能に興味関心がある。このように理論言語学と応用言語学が取り扱う文法問題について比較しても,なかなか両者が共同作業できるような論点を探るのが難しそうである。

しかしながら,本論では,理論言語学と応用言語学の接点の一つとして,コーパスに基づく「語彙-文法」(lexicogrammar)という考え方の有効性を論じる。理論言語学としてこれまで主に生成文法を想定してきたが,選択体系機能文法(systemic functional grammar)に目を転じると,語彙と文法が結びつきを深める。Halliday (2009: 73-74)は次のように述べ,語彙-文法を特徴づける。

... the boundary between grammar (or syntax) and lexicon is extremely fuzzy; the two are joined in a continuum, and they are of the same order of abstraction, so that while we do need to recognize the distinct categories of grammar and lexis we also need to model them as a unity on a single stratum.

文法と語彙を切り離す考え方を取らずに,両者は連続しているという意味で「語彙-文法」という。時々見かける,語彙という「れんが」が文法という「モルタル」でつながれているという比喩は妥当性を欠くという。むしろ語彙と文法は一体となって,意味を具現するために共同作業すると考える。一例をあげると,同様の意味を表現するときにも,語彙-文法は書き言葉が得意とする名詞型と話し言葉が得意とする節型の区別をする。

(1) a. **Fire intensity** has a **profound effect** on **smoke injection**. [名詞型]

b. The more **intense** the **fire** the more **smoke** it **injects**. [節型]

(Halliday 2009: 77)

(1a)は節が一つ,語が六語彙素という構成から成り立ち,名詞・形容詞+名詞という語彙の「結晶」を含むコンパクトな表現が特徴的である。逆に(1b)は,二つの節と四語彙素から構成された結果,語彙の密度が低く,障害なく流れる川のようなものである。このように,語彙と文法は一致協力しながら書き言葉,話し言葉それぞれにふさわしい文を実現する。

応用言語学の立場から「文法」を論じる Larsen-Freeman and Decarrico (2010) は,文法と語彙は相互依存関係にあるとし,文法的規則性がしばしば語彙的制約を被ることを例証する。英語動詞の進行形が取る意味は,語彙による左右される。Mary is taking a nap. では,進行形が一時的行為を, Mary is taking a class では,一定期間持続する行為をそれ

ぞれ表している。つまり目的語の名詞句が a nap か a class のいずれかにより、進行形の一時的の長さが変わってくる。語彙 - 文法 の存在を証拠づける応用言語学的現象をもう一つ上げたい。Francis, Hunston and Manning (1996: 174-176) は、Verb + by + V-ing というパターンを取る動詞をコーパスデータを基に三つの意味グループに分けている。

- ・ 開始または終了を意味する動詞 : begin, close, end, finish, open, start
- ・ 反応, 対応を意味する動詞 : atone, compensate, counter, react, reciprocate, reply, respond, retaliate
- ・ 資源を意味する動詞 : live, profit

ここで注目すべきは、意味を共有する動詞群が動詞パターンも共有するという事実である。動詞群, その意味, そのパターンの中で文法的・語彙的相関が生じているとも言い換えられる。

以上の議論をまとめると、理論言語学特に選択体系機能文法と応用言語学がめざす文法は「語彙 - 文法」という接点で出会う可能性を探る価値がありそうである。そしてこの両者を結びつけるのがコーパスに基づく言語研究である。語彙 - 文法の実例として紹介された話し言葉, 書き言葉の比較研究, 進行形とその意味, 動詞パターンのいずれも、コーパス言語学の得意とする分野であり、コーパス研究を利用してこそ有意義な成果が得られる研究題目である。2章では、現代コーパス研究につながる経験主義的言語研究の流れを言語学史で確認する。3章は、コーパス研究から応用言語学に生かせる語彙 - 文法的知見を考察する。理論言語学とコーパス研究の成果を取り上げる4章では、Hoey のコロケーション研究を概観する。最後に、結論としてコーパスに基づく語彙 - 文法という共通基盤に立つことにより、理論言語学と応用言語学が出会うと結論する。

2. 経験主義の流れ

言語学における経験主義とは、言語運用の結果である言語資料に基づき言語記述を行う研究手法をさす。これは、生成文法のように、母語話者の言語直観による文法性判断を拠り所に言語研究を行う合理主義と対峙する。言語研究史と経験主義のつながりを振り返る結果、言語データから言語の実態を探る言語研究法が歴史的裏付けを有していることを明らかにしたい。そしてことばの使用現場をしっかりと観察する経験主義にたつてこそ、理論と応用の接点が生まれる点も指摘したい。

言語研究の歴史を振り返るとき、経験主義が成果を上げるのはまず辞書編纂の分野である。Samuel Johnson が 1755 年に発刊した A Dictionary of the English Language では 15 万例に及ぶ用例を採取する手段として、1560 年から 1660 年にわたる作品から用例を拾いメモ用紙に書き取る作業を繰り返した。その際 “Words must be sought where they are used.” という実証主義の方針を貫いた。それから一世紀を経て語義の歴史主義に立つ Oxford English Dictionary の編纂が開始され、1933 年に 12 巻の本体と補遺 1 巻から成る第一版が発行された。この辞書編纂では、2000 人もの読者が約 400 万例の引用を提供したとされる。語の記述を行う辞書の編集は膨大なデータに支えられていることが読み取れ

る。

英文法研究に目を向けると、いわゆる伝統文法が経験主義の立場にたち、膨大な実例収集を手作業で行っている。Jespersen の A Modern English Grammar on Historical Principles (1909-49) は、英文法記述を支えるために、多くの文学作品に加えて歴史書、哲学書、伝記などほぼ 1000 冊の文献から書き言葉用例を集めている。他方、米国では構造言語学が盛んな 20 世紀中葉に、Charles Fries が 25 万語の会話コーパスに基づき The Structure of English (1952) を、書き言葉である手紙 3000 通を利用し American English Grammar (1940) を著している。太田朗の博士論文 Tense and Aspect of Present-day American English (1963) も Fries の研究に連なる実証データに基づく英文法研究である。10 時間を超える草稿なしのラジオ会話、テレビ劇の脚本 10 本、3 万語の書き言葉というデータを徹底的に分析することで、時制と相に関する論を展開する。

同じ 20 世紀中頃、University College London の Survey of English Usage では、Quirk が文法情報を付したカード方式コーパス (1955-1985) の構築を開始した。ジャンルを考慮し 200 種類のテキストから 5000 語ずつ採集している、その半分が話し言葉である、教育を受けた教養ある人が話す「標準英語」を対象としているという特徴を有する。このコーパスはその後、A Comprehensive Grammar of the English Language (1985) という英文法の大著に結実した。同じ頃米国では、いわゆる Brown Corpus という 1961 年のアメリカ英語書き言葉 100 万語からなるデータベースが、W. Nelson Francis and Henry Kučera の手により誕生した。コンピュータ利用が可能な初のコーパスとして、その後同じデザインで作られた一連のコーパスすなわち LOB (英国英語 1961 年版)、Frown (米国英語 1992 年版)、FLOB (英国英語 1991 年版) のモデルとなった。その後のコーパスの発展により、語彙検索、品詞検索、文法検索、テキスト別検索、話者検索等が広く行われるようになり、今までにない言語運用へのアプローチが可能になりつつある。

1960 年代に入り、合理主義に立脚する生成文法が言語学研究の主流となるなかで、1969 年に Labov がアメリカ言語学会誌 Language に “Contraction, Deletion, and Inherent Variability of the English Copula” という論文を発表した。この論文は

- (2) a. He is unfortunately here. [完全文]
- b. He's unfortunately here. [BE 動詞縮約]
- c. He unfortunately here. [BE 動詞省略]

という三種類の文が黒人の話し言葉コーパスでどのように実現しているか、を調査している。この研究の意義は、義務・随意という生成文法の書き換え規則に代わり、規則が適用される比率を指数によって表す可変規則を提案している点にある。話者の年齢幅を設定し、BE 動詞の縮約や削除が生じる文脈を丁寧に観察し、統計処理を施している。発表当時、生成文法側から、このような頻度に基づく研究は無意味だと切り捨てられたが、現在から振り返れば、経験主義の伝統にのった実証研究であり、その後の社会言語学での変異研究につながっている。

言語資料の形はさまざまであっても、資料収集を行い、その資料に基づいて辞書を作る、文法記述を行う、言語変異を調べるといった経験主義的言語学研究は、理論と応用という区

別を全くしない点が重要である。現代コーパス言語学の鳥瞰図を与えてくれる Lüdeling and Kytö (2008) によれば、コーパス言語学は、音声、文法、意味研究に加えて、関連する応用言語学分野として、Computational Linguistics, Discourse Analysis, Historical Linguistics, Language Acquisition, Language Teaching, Sociolinguistics, Stylistics, Translation Studies 等にも広がり、理論言語学だけでなく、応用言語学でも強力な研究方法として認知されつつある。

3. コーパスから言語教育（応用言語学）へ

O’Keeffe et al. (2007) は、文法と語彙の中間に位置する word chunks, idioms, semantic prosody を紹介し、語彙 - 文法の必要性を示している。500 万語の CANCODE 話し言葉コーパスに含まれる高頻度の二語 chunks のうち 5 例をカッコ内の頻度とともに列挙する。

(3) you know (28013), I mean (17158), I think (14096), sort of (9586), have to (5914)

これらの chunks が談話内で果たす機能を例示する。

・ 話題導入機能

(4) S1: **You know**, our Gregory he’s only fifteen but he wants to be a pilot.

S2: Does he?

・ 面子維持機能

(5) S1: Did you want to take out insurance?

S2: Erm I’d like to ask about it but **I don’t know if** I want to do that today.

いずれの談話機能も話し言葉を円滑に進行させるために不可欠であり、コーパスでの高頻度を考慮すると、母語話者の流暢さを目指す英語学習者にとって、chunks は是非とも習得したい言語項目であると言える。さらに、(4), (5)の会話例が示すように、chunks は、一連の会話テキストの中で使用されてこそ、その機能を発揮するので、その点を考慮した練習が求められる。

高頻度の chunks とは対照的に、低頻度のイディオムがある。(6)の a pain in the arse を CANCODE コーパスの 2 倍サイズある BNC の話し言葉コーパス (約 1000 万語) で検索しても 14 例しか見つからない。(3)の chunks の高頻度とぜひ比較したい。

そのイディオムの機能の一つが評価である。まずコーパスの会話を観察すると、イディオムは(6)のような exchange の S2 に生起することが多い。

(6) S1: Well I though you were gonna go on holiday.

S2: Yeah. The thing well I don’t think I’m gonna do that now cos none of us get together at the right time when we want to do it. Which **is a pain in the arse**.

太字のイディオムは、S2 という話者が休暇について観察を述べた後に、コメントをするという働きをしている。コーパス調査が言語教育に示唆しているのは、S1 ではなく S2 のターンでイディオムを使用する場面こそが、イディオム学習にふさわしいということである。そうすると、英語の授業では次のような教師と生徒の対話練習が想定できる。

- (7) Teacher: Josh adores cowboy films!
 Student: **There's no accounting for tastes!**

イディオムの仕事の評価であるとする、学習者には良い評価、悪い評価、中間の評価と選択肢を与え、学習者の一人一人の気持ちを表現させる練習が重要であることになる。

コーパスから得られ、外国語教育に有効な知見の第三として semantic prosody を取り上げる。border という語は名詞としては「境界」という文字通りの意味で使用されることが多いけれども、bordering や bordered という形で on とともに動詞用法に至ると、比喩的意味つまり「... と同然である」という意味が優勢になる。さらに on の直後に現れる名詞を集めると次のような語群が得られる。

arrogance, apathy, alcoholism, antagonism, bad taste, blackmail, carelessness, chaos

これらの語に共通しているのは something undesirable 特に a state of mind undesirable という意味である。つまり bordered/bordering on に後続する前置詞目的語は否定的な意味が込められているという semantic prosody を有する。その結果、border を動詞として使った瞬間から、話者は読者に目的語を望ましく思っていないという意味を伝えることになる。She felt overwhelmed by his personality, which sometimes **bordered on piety**. という例の場合、piety 自体の意味は別にしても、彼女にとって望ましいものとしてとらえていないことが読み取れる。実際に後続するテキストには “I think he is wrong to inflict his beliefs on everyone else.” という批判がなされる。このような話者の意味選択を英語学習でどのように扱うか、は検討課題の一つとなる。従来の提示、練習、産出という教授法でよいのかどうか、考え直す必要がある。学習者自身が談話内で semantic prosody を発見する言語教材を提供し、学習者が談話探検をするような教授法も考えたい。

外国語教育に対するコーパスの貢献、特に語彙 - 文法に属する言語事実として chunks, idioms, semantic prosody を順に論じた。従来の個別の語彙項目に加えて、複数語から成る固定表現の習得も大事であることがコーパス研究から判明してきた。

4. コーパスからコロケーション理論 (理論言語学) へ

文法と語彙は別々にあり、文法が生成する構造にふさわしい語彙を挿入するという伝統的文法観に従うと、膨大な語彙をリストとして脳内辞書に蓄える母語話者がなぜ流暢に発話し、文章を書けるのか、という問題が生じる。この間に答えてくれるのが Hoey (2005) である。彼は、母語話者は語句に出会うときにそのコンテキストとともに脳内辞書に登録すると想定する。そしてこの蓄積すなわち語及びそのコンテキストの無意識な格納を一種

の priming (刷り込み, 詰め込み) と考える。それでは語とともにどのようなコンテキスト情報を知っているのだろうか。Hoey は次のリストを掲げる。

- ・ 語彙的連語つまりコロケーションやイディオム
- ・ 共起する文法パターンつまりコリゲーション (文法的連語)
- ・ 連想される意味
- ・ 語句がどのようなジャンルで使用されるか
- ・ 文頭と文末のどちらで生起しやすいのか

理想的には個人が遭遇するすべての話し言葉・書き言葉テキストを収集できればいいが、それは非現実的なので、ある言語変種を代表している大規模コーパスを利用しながら、特定の語句とそのコンテキスト情報を調査するという研究方法をとる。

(8)では、*the time has come* というフレーズを使い、コーパスから得られるコンテキスト情報を頻度ともに例示することにする。

- (8) a. **I feel** the time has come for our team to move in a new direction.
- ・ *the time has come* は三例につき一例の割合で思考感情動詞 (e. g. *feel*) と共起する。
- b. The time has come to **declare** failure.
- ・ *the time has come* は五例につき一例の割合で *to* + 発言動詞 (e. g. *declare*) をとる。
 - ・ *the time has come* は二例につき一例の割合で文頭を占める。
- c. **Mark said:** "The time has come to **change roles** ..."
- ・ *the time has come* は直接被引用節となることが多い。
 - ・ *the time has come* は大きな変化とともに使われる。これまでの例文でも a new direction, failure, changing roles などはいずれも大変化の例である。

母語話者は意識せずに、自分が以前聞いたり読んだりしたことばを再生産していることになる。その際には刷り込まれたコンテキスト情報を活用し、その語句の意味を思い出し、さらには文法パターンを再現する。このように考えてくると、母語話者の流暢さも納得がいく。語句とともに、意味、文法、コロケーション、テキストのジャンル、テキスト中の位置を自分が以前聞いたり読んだりしたようにリサイクルしているので、流暢でかつ自然な言葉遣いができる。

Hoey が(8)で提供するコンテキスト情報をもう一度振り返り、文法情報と語彙情報が一体となっている事実を確認したい。(8a) では思考感情動詞 (例 *feel*) という語彙情報が *the time has come* に先行するという文法情報と結びつけられている。(8b) でも同様に、*the time has come* に *to* 不定詞が後続する確率が高く、かつその *to* 不定詞の動詞は発言するという意味と取ることが多いと、語彙の意味を指定している。(8c) の場合には、引用節 + その補文という文法形式に加えて、引用節には引用を合図する動詞が予想されるという。英語母語話者は以上のように *the time has come* にまつわるコンテキスト情報を十分に

蓄える。そしてその言語知識を活用しながら、例えば英国園芸雑誌 *The Gardener's World* の記者は、**The time has come** to think about putting something back into the soil. と自然な英語を書くことができる。

5. 要約

従来、理論言語学と応用言語学は前者（特に生成文法）が言語能力に、後者が言語運用に関心を持つから、共通点を持ちお互いに貢献するという関係にないとされてきた。しかしながら、本論は、コーパスに基づく語彙-文法 (lexicogrammar) という視点をとるとき、両者は対話を開始することができる、と主張する。つまり語彙と文法を別々の組織として切り離さずに、むしろその境界はファジーなものであると理解する。理論言語学（選択体系文法）と応用言語学（言語教育）がコーパスという共通の道具立てを用い、語彙-文法という同じ土俵に上がってくる。

言語学における経験主義の流れを、辞書編纂、英文法研究、コーパス構築、Labov の可変規則という順番で追跡した。いずれもデータに基づくという研究方法を共有し、そこに理論、応用という区別はない。中央にコーパスを含むデータがあり、それを活用するときには様々な広がりを見せ、コーパスを通じて共同作業することもしばしばである。

次に、コーパスに基づく言語教育研究を行うとき、語彙-文法という視点が有効な論点として、chunks, idioms, semantic prosody を議論した。chunks はコーパスにおいて高頻度で使用されるフレーズであり、その習得は外国語の流暢さにつながるため、豊富なコンテキストの中で学習することが期待される。文法と語彙の中間に位置するイディオムは一つ一つ意味を覚えるだけで終わっていたが、その評価という談話機能を生かし、話者の評価を重視した学習が必要である。三番目に「bordered/bordering on + 望ましくないもの」という semantic prosody をあげ、教えるというよりも学習者に発見させる教授法の開発が課題であるとした。

コーパスに基づく語彙-文法として Hoey のコロケーション論を最後にあげ、コーパスという道具を使うことで新たな言語観を紹介した。母語話者は豊富な言語資料に触れる中で、語句自体の情報だけでなく、そのコンテキスト情報も同時に脳内辞書に蓄積する、あるいは無意識な刷り込みを行う。その結果、一つの語句を選択した瞬間にそのコンテキストも同時に選択することができ、流暢な言語運用を実現している。そしてそのコンテキスト情報の解明にコーパスが大きな貢献をする。

以上で論じてきたように、コーパスに基づく語彙-文法という接点を設定すること初めて、理論言語学と応用言語学が文法について同じテーブルにつくことができる。そして対話を開始すると、次の課題としてどのような共同研究ができるのか、話し合う段階に入れる。実際に共同研究を行う時にはコーパスという共通の研究ツールを用いることで、お互いの研究を検証できる。コーパスと語彙-文法という立場を認め合えば、Hoey や Halliday という理論言語学者と O'Keeffe や McCarthy という言語教育研究者が積極的にお互いの発展に向けて意見を戦わせると大いに期待できる。

*本稿は、2009年7月18日上智大学言語学会代24回大会において「理論言語学と応用言語学が会うコーパス研究」というタイトルで口頭発表した原稿に加筆修正を加えたものである。

参考文献

- Francis, Gill, Susan Hunston and Elizabeth Manning. 1996. *Grammar Patterns 1: Verbs*. Glasgow: HarperCollins.
- Fries, Charles C. 1940. *American English Grammar*. Tokyo: Maruzen.
- Fries, Charles C. 1952. *The Structure of English*. New York: Harcourt, Brace and World.
- Halliday, M. A. K. 2009. "Methods—Techniques—Problems." In Halliday, M. A. K and Jonathan J. Webster (eds.) 2009. *Continuum Companion to Systemic Functional Linguistics*. London: Continuum, 59–86.
- Hoey, Michael. 2005. *Lexical Priming*. London: Routledge.
- Jespersen, Otto. 1909–1949. *A Modern English Grammar on Historical Principles*. London: Allen and Unwin.
- Johnson, Samuel. 1755. *A Dictionary of the English Language*. London.
- 小池生夫. 2003. 「応用言語学研究：その過去、現在、未来」In 小池生夫（編）『応用言語学事典』東京：研究社, xxxiii–xxxviii.
- Labov, William. 1969. "Contraction, Deletion, and Inherent Variability of the English Corpula." *Language* 45: 4, 715–762.
- Larsen-Freeman, Diana and Jeanette DeCarrico. 2010. "Grammar" In Norbert Schmitt (ed.) *An Introduction to Applied Linguistics*. 2nd edition. London: Hodder Education, 18–33.
- Lüdeling, Anke and Merja Kytö (eds.) 2008. *Corpus Linguistics: An International Handbook*. 2 vols. Berlin: Mouton de Gruyter.
- O’Keeffe, Anne, Michael McCarthy and Ronald Carter. 2007. *From Corpus to Classroom*. Cambridge: Cambridge University Press.
- Ota, Akira. 1963. *Tense and Aspect of Present-day American English*. Tokyo: Kenkyusha.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvick. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Schmitt, Norbert. 2000. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- 渡辺明. 2009. 『生成文法』東京：東京大学出版会