



東京大学大学院
総合文化研究科提出修士論文

A Corpus-based Study of Lexical and Grammatical Features of Written Business English

(コーパス資料に基づく現代ビジネス英文の語彙および文法的特性の分析)

染谷 泰正

1999年

指導教官

鈴木英夫教授 (主査)

岡秀夫教授 (副査)、松野和彦教授 (副査)

[和 文 要 旨]

1. 研究の背景

ビジネスの国際化の一層の進展とともに EBP (English for Business Purposes) への関心はますます高まってきている。しかし、この分野における学問的・科学的研究は、他の ESP 分野に比べて大幅に立ち遅れており (Dudley-Evans & St John, 1996:1-13)、その関心の高さにもかかわらず、われわれは「ビジネス英語」の実態について「ほとんど何も知らないに等しい」(Holden, 1989:43) という状態にある。その結果、現在、市中に出回っている EBP 関連の各種参考書およびテキストブックの多くは、いずれも執筆者の個人的な体験や主観的な判断にのみ基づいて記述されており、理論的な基盤やデータによる裏づけを欠いたものとなっている。教育的観点から見ても、現在、EBP について、何を、なぜ、どのように教えればよいかについての信頼すべき指針は存在しないと言ってよい。

2. 本研究の目的と意義

本研究は、上記のような問題意識に立ち、国際ビジネスの主要な手段である「文書によ

るコミュニケーション」に焦点を当て、そのうちの最も基本的な分野である「語彙（および語彙文法）部門」の実態と特徴を、おもに言語学的な視点から明らかにすることを目的とする。また、この目的を達成するために、本研究では、大規模なコーパスデータをさまざまな角度から分析するための各種コンピュータ・プログラムを開発し、その有効性を実証するとともに、他の ESP 分野への応用可能性を探る。本研究によって得られる知見は、今後の研究のための基礎データとなるばかりでなく、教育現場での指導や教材作成のための客観的な指針を提供するという意味で大きな意義を有するものであると考える。

3. 研究の方法および手順

本研究は、コーパス言語学的手法に基づく研究である。したがって、研究の開始に当たって、まず基礎データとなる「ビジネスレターコーパス」（以下 BLC [または Native BLC]）を作成した。収録語数は後述の対照コーパスの規模に合わせておよそ 100 万語とし、英米で発行されているモデル文例集を中心にデータを収集した。また、ビジネス英語を学習している日本人ビジネスマンの作成したビジネスレターもあわせてコーパス化し BLC との比較材料とした（以下「学習者コーパス = Learner BLC」）。このほか、一般英語を代表する英語コーパスとして *Brown Corpus* および *LOB Corpus* を、また特定ジャンルコーパスとして *TIME* 誌 100 万語分の記事を収録した *TIME Corpus* の 3 つを対照コーパスとして使用した。データ解析のためのソフトウェアとしては *WordSmith* および *TXTANA* を使ったが、これらの既成ソフトではできない細かな分析作業については *AWK* を使って独自のコンピュータプログラムを自作した。

4. 結果と考察

本研究の結果、EBP の特徴についてさまざまな事実が明らかになった。まず、EBP は予想どおり高い「語彙閉鎖性」(lexical closure) を示し、その語彙成長曲線はタイプベースでおよそ 1 万 5000 語から 2 万語で頭打ちとなる。レンマ化した場合、およそ 100 万語のコーパスのうちの 90.8 パーセントはわずか 1500 語でカバーされる。これは、他の一般英語コーパスが示す数字のおよそ 1/2 から 1/3 にすぎない。品詞別の分布では、BLC に現れる全動詞の 90.83 パーセントが 350 語のレンマタイプでカバーされ、副詞については 100 語で全体の 89.4 パーセントがカバーされる。この語彙閉鎖性の傾向は *Learner BLC* においてはいっそう顕著であり、動詞と副詞について言えば、それぞれ全体の 90 パーセント超のカバレッジを得るための数字は、前者が 109 語、後者が 43 語にすぎない。これらの数字が顕著に示すことは、EBP という特定ジャンルにおけるディスコースは、ほぼ一定の語彙項目群によって構成されており、これらを "an established common core of business language" (Dudley-Evans & St John, 1996:1-13) として認定することが可能だということである。本研究では、統計的手法を駆使してこれらのコア語彙 (core lexis) を形成する具体的な語彙項目のリストを作成した。

また、コーパス分析の結果、EBP における語彙は全体に難度レベルが低く、上位頻度語

彙 1500 語のうちのおよそ 90 パーセントが JACET (大学英語教育学会) の基礎 4000 語以内の語彙となっている。残りの 10 パーセントの大半はビジネスに特有の各種専門 (および準専門) 用語を構成している名詞および形容詞である (いわゆるビジネスジャーゴンの類を含む)。後者の語彙群 (本論文における分類では「レベル 6」以上の語彙) は、全体の数こそ少ないが、いわば EBP という特定ジャンルのディスコースに特有の雰囲気や「らしさ」を与えている要素であり、さらに詳しい分析を要する分野である。

Native BLC における高頻度語彙は、いくつかの特殊な例外を除いて、いずれも Learner BLC においても同様に高い頻度で使用されている語彙である。この意味で、上述の「コア語彙」は Native および Learner の双方に共通するものと考えてよい。ただし、後者の場合、これらのコア語彙におけるパフォーマンスエラーの比率が高く、特に高頻度・低難度動詞に関するエラーが多いのは特筆すべき事態であると考えられる。今回の分析では Learner BLC における上位動詞 100 語 (その 91 パーセントは JACET の基礎 2000 語以内の語彙) の平均エラー率は 15.57 パーセント (SD =11.87) となっており、基本語の習得がごく不十分であることを示している。これらのエラー例を詳細に検討した結果、日本人 EBP 学習者のエラーには一定のシステムチックな傾向があることが明らかになった。動詞について言えば、日・英で意味的に等価の単語があり、これがそれぞれ異なった「項構造」や「選択制限」(またはその両方) を持っている場合に特にエラーが起りやすい。例えば discuss については 67.69 パーセント、require では 20.51 パーセントという高いエラー率が記録されている。これらの数字は、わが国の英語教育における致命的な欠陥のひとつを象徴的に示していると言ってよい。

統語レベルでは、BLC の文章は比較的単純な構文で書かれており、対照コーパスと比較して文体的バリエーションが少ないのが特徴である。その中でも特に好んで使用されているのは関係詞構文と分詞構文であり、これに後述の条件法構文を合わせて、いわば「ビジネス文書の 3 大構文」を構成していることが明らかになった。ただし、Learner BLC についてはこの「3 大構文」の使用が不十分または不適切であり、全体として「プロフェッショナルな文章」に期待される文体的バリエーションが実現されているとはいいがたい。教育現場での指導に当たっては、今後、このような点にも配慮が必要であろう。

統語的なバリエーションという観点から言えば、おそらく最も重要な課題は条件法構文 (仮定法を含む) の正確な習得であると考えられる。言うまでもなく、ビジネスディスコースは、通例、過去と現在だけでは完結せず、何らかの形で将来のことについて言及しなければならないことが多い。将来のことは、典型的には条件文ないし仮定法を使って表現されるのであって、これらを適切に使いこなすことができるかどうかは、ビジネスディスコースの構成にあたって決定的に重要な事柄であると言ってよい。EBP における条件法・仮定法構文の重要性は、法助動詞および条件法標識の if の使用頻度の高さにも如実に現れている。POS 分析の結果によれば、対照コーパスにおける法助動詞の全語彙に占める割合は平均 1.34 パーセントであるのに対し、BLC では 2.43 パーセントであり、その差は他のカテゴリーにおける差に比べて最も有意に高くなっている ($Z = 9.25, p < 0.001$)。また、各

語彙項目のキーワード性を示す指標 K-score は、will = 9200, would = 1468, can = 1036, shall = 385, should = 206, may = 180 となっており、法助動詞の中でもとりわけ will と would の重要性が突出していることが伺える。この傾向は if についても同様である。

なお、本研究で同定された「コア語彙」の多くは特定の統語パターンと密接に結びついていることが観察されている。例えば important という形容詞は BLC 全体の中で 367 回出現し、そのうちの 57 ケース(15.53%) が It is ADJ (for NP) to 補文 または It is ADJ that 補文 (ないし ZERO-that) という統語フレームで現れている。また、この統語フレームの ADJ のスロットに現れる語彙もきわめて限定されていることがわかっている。本研究では、EBP の語彙を論ずるに当たって最低限必要な範囲で統語的な問題 -- 例えば前述の条件法・仮定法構文や名詞化構文、受動態構文、関係詞構文など -- についても議論したが、今後は、この方向での議論をさらに進め、語彙レベルと統語レベルのインターフェイス、さらに談話レベルの分析へと研究を進めていきたいと考えている。

チョムスキーが述べたように、「言語の習得とは詰まるところそれぞれの個別言語の語彙の習得に尽きる」(Chomsky, 1989:44) のであり、生成文法における投射原理が示唆するように、EBP を含む外国語習得においても、その「焦点はますます語彙習得に置かれることになる」(Cook, 1993:204 [傍点筆者]) ものと予想される。ただし、この場合の語彙習得とは、それぞれの語彙項目の統語的振る舞いや、意味論的・語用論的制約などを包括したものであり、今後は、語彙の指導に当たって、このような観点を取り入れた新たな指導法を工夫し、実践していくことが望まれる。

5. 本研究の貢献

前述のとおり、本研究は EBP の全体像を把握するための第 1 歩として、まず「語彙」に焦点を当ててさまざまな観点から分析を行った。本研究の貢献は、1) EBP 研究におけるコーパス言語学的手法の有効性と可能性を実証したこと、および 2) その手法に基づく各種の分析の結果、従来、直観的にしかとらえられていなかった EBP の実態 -- とりわけその語彙的諸側面について客観的な議論が可能な定量的方法で明らかにしたこと、さらに 3) EBP の「コア語彙」を各品詞別に同定したこと、などが挙げられる。このうち第 1 点の点については、本研究において開発・作成された各種のコンピュータ・プログラムを広く研究者に公開し、共有することで、研究活動の進展にいつそう寄与することができるものと考えられる(これらのコンピュータ・プログラムについてはすでに筆者のホームページ上で順次公開している)。また、第 3 点については、ごく小規模かつ実験的な試みを除けば、本研究によって始めて行われたものである。ただし、本論文で提示した「コア語彙リスト」は、このままの形ではさほど意味のあるものではない。今後は、このリストに基づいて各単語のひとつひとつについて詳細な「語彙プロファイリング」(lexical profiling) を行っていく必要がある。これは次の課題のひとつということになるが、その概要を本論文の Appendix C3 に示した。

6. 今後の課題

今後の課題についてはすでに述べたとおりである。以下、おもなものについて列記すると、a) 本論文第4章で議論した "metadiscourse markers" についてのより詳細な分析、およびこれを契機とする EBP におけるモダリティー全般の研究、b) 「コア語彙」とその統語パターンおよび連語関係の分析（語彙レベルと統語レベルのインターフェイス研究）、c) 「コア語彙リスト」に基づく詳細な語彙プロファイリング、さらに d) EBP のディスコースレベルでの分析、ということになる。もとより、いずれのテーマも一朝一夕でできるたぐいのものではないが、本研究の成果をベースにして、今後、自らのライフワークとしてこれらの課題にひとつずつ取り組んでいきたいと考えている。

研究の概要 (Phase 1)

