

# 教材とコーパス

投 野 由紀夫

本日は「教材とコーパスのかかわり」、「コーパスが教材開発とどういう接点があるか」についてお話します。最近、教材を作るようになって、具体的にコーパスデータから教材まで下ろす作業をしていますが、研究分野として整理してみる視点から、その中でいろいろ考えていることを皆さんと分かち合いたいと思います。コーパスには色々種類がありますから、各種コーパスと教材開発の関係を考えなければなりません。まず、コーパスがどのように教材開発に利用されているかについては、大きく分けて二つの利用法があります。まず一つは教材になる前の基礎資料としてコーパスを利用することができます。教材を作るまでの具体的な製品にはならないが、製品になる一步手前の資料をコーパスで出すという発想です。具体的な学習目標の設定、教材の分析や評価、シラバス・デザイン、テスト開発などが考えられます。もう一つは、コーパスを教材の中身として利用するという発想があります。参考書や問題集とかではすでにこの方法が使われています。これらの視点から、「教材とコーパス」は今、こんなふうになっているということを紹介しながら、英語教育におけるコーパスの利用について考えると同時に、コーパス利用を実際にしてみると直面する問題点について整理し、今後の課題について話したいと思います。

## 高まるコーパスの教育利用への関心

広く言語に関わる研究全体で見えますと、今までもコーパス言語学という分野として言語学者がコーパスを利用する発想はあったのですが、言語教育への応用は、この5、6年、急速に進んできた新しい分野だと考えられます。海外のコーパスを利用した英語教育関係の論集では、古いのはランカスター大学で第1回が1994年に行われたTALC (Teaching and Language Corpora) から、2000年代まで数々の論文集が出ています (Wichmann et al. 1997; Kettemann & Marko 2000; Burnard & McEnery 2000; Aston 2001; Hunston 2002; Granger et al. 2002; Tan 2002; Sinclair 2004; Aston, Bernardini & Stewart 2004)。

日本でもこの種の集会が増えており、英語コーパス学会でも10周年の記念シンポジウムではさまざまな分野の教育にコーパスを如何に応用するかということが扱われています。昨年12月、昭和女子大学で学習者コーパスの構築と分析をテーマに、言語教育分野においてどのようなことができるかについての国際会議も行われました。

言語教育におけるコーパスの利用についての基本文献として、ランカスター大学のG. Leech博士が1991年に発表した論文があります。Leech先生は言語教育プロパーの研究者ではありません

せんが、コーパスの言語教育に果たす役割について、非常に深い洞察を示されています。ここでは次のように大きく分けて3つのことを述べています (Leech 1991)。

第一はコーパスを直接的に利用する方法 (direct use) です。この直接的利用法は、さらに、三つに細分できます。一つはコーパス言語学そのものを教える (Teaching about)。コーパスを使いながらコーパス言語学を教えるわけです。もう一つは、コーパス検索の方法を教えて、各自の目的にあわせてコーパスを利用させる方法です (Teaching to exploit)。三つ目として、あることを教えるためにコーパスを使ってみる方法が挙げられます (Exploit to teach)。コーパスを使わないでもいいが、使ったらどうなるかとか、直接コーパスデータに接するとどのような結果が生まれるのかなどを見ることが考えられます。コーパスの大きな利用法の2番目は間接的利用 (indirect use) です。各種辞典編纂、云々というコンテキストで紹介しています。最後に、教育志向のコーパス作成が取り上げられています。教育に特化した目的のコーパスをつくるという分野もあるのではないかという問題提起です。

### コーパスと教材開発の接点

ここまでは、広く言語教育とコーパスとの接点を考えてきましたが、教材開発との接点はどのようなところがあるかに移りたいと思います。私なりにまとめてみますと、教材開発にコーパスを使う利点として、まず大規模なコーパスデータを利用すれば典型的な言語使用例がわかること、が挙げられます。日本語には「すごく」「すごい」、「やっぱり」「やはり」などの似たような表現がありますが、大量の日本語データがあれば、それらが使用される際の傾向や頻度などがわかります。より基本的なものとか、使用頻度の高いものを選定することができ、優先順位を設定する資料になります。もう一つ便利なのは、単語の共起情報が出せますから、一緒に使う単語がわかるので、一個一個の単語を単体で覚えるのではなく、こういうフレーズで覚えよう、こういう組み合わせがいいという示唆ができます。これは、より英語らしい表現という指標に関係してくると思います。

コーパスは教材開発の視点からは大別して母語話者のコーパスと学習者コーパスの2種類に分けられます (表1参照)。

表1：コーパスの教材作成の観点から見た分類

母語話者のコーパス	
汎用メガコーパス	辞典類、一般教材
ESP コーパス	専門語辞書、ESP 教材
難易度調整コーパス	ライティング支援
学習者コーパス	
エラー分析結果	辞典・語法・文法書
習得データ分析結果	シラバスデザイン
	テスト開発

母語話者のコーパスはさらに大きく分けると3つあります。一つは汎用のメガコーパス。これは大量のテキストをコーパスにしたものですから、一般的な語彙関係の統計情報データベースとして、あるいは辞典の編纂によく使われます。当然、一般の教材にも使えます。2つ目は、もっと分野を狭めて、特殊な分野に特化したコーパスです。このようなESPコーパスからは専門語辞書とかESP教材がつくれますので、徐々にこのような試みもされ始めています。3つ目は難易度を調整したコーパスです。いきなりネイティブの生データに触れさせずに、易しい英文を含んだようなコーパスデータに触れることによってライティングの支援を考えています。これにより、英文を書く際に母語話者の自然な言語使用を記録したコーパスデータに触れる環境をまず整えることが可能になります。

学習者コーパスを編纂すれば、エラー分析ができ、それを用いて世界的に辞典や文法書の編纂に応用する動きが見られます。それだけではなく、第2言語習得の研究にとって重要な資料を提供してくれます。中1から高3までの中間言語のデータが分かってくると、日本人学習者がどのような習得のプロセスを経るかが記述できますから、それをもとにシラバス・デザインを組むことや、テスト開発に生かそうとする発想も出てくるわけです。

## コーパス・データの間接利用（Indirect use）

そこで、今日はこのようなことをもとに、教材開発へのコーパス利用の分野を概観してみたいと思います。一つは教材開発用の資料としてコーパスを使う場合、世界中でどんなことが行われているのか、実例をいくつかお見せしたいと思います。コーパスデータを直接教材に利用するというでもいくつか実例を示したいと思います。NHK「100語でスタート！英会話」もその一つだと思います。

### （1）学習語彙表

まず、学習語彙表作成ですが、電子化以前からコーパスをもとにした仕事は存在していました。アメリカではThorndikeという有名な実験心理学者が1920年代から語彙表作成をしていました。総計1810万語のコーパスのデータから3万語の基本語リストを作ったりしています。そういう時代が20年代から戦前まで続きました。戦後では、アメリカでJohn B. Carrollが挙げられますが、Noam Chomskyが50年後半から生成文法を提唱してから、教育分野では語彙に関する興味が失われてしまいました。言語学の興味の中心がシンタックスに移ってしまった結果、言語習得の論文もシンタックス中心になってしまった。というわけで、一時期、語彙表作りも姿を消してしまいました。90年代からボキャブラリーとか語彙中心の文法などの見直しが行われ始めまして、最近になって語彙がリバイバルしてきたという感じです。このへんの歴史的な変遷を示したのが図1です：



図 1：学習語彙表の歴史的な変遷

言語学者や自然言語処理の人達が、どういうことをしたかと言いますと、今日初めてブラウンコーパスの磁気テープを見ましたが、あのような時代から徐々にコーパスに基づいたさまざまな資料が出てくるようになります。ただ、教育に特化した語彙表はあまりありませんでした。あるコーパスデータを使ったコーパス全体の語彙頻度表、分野別あるいは品詞別の頻度表などは数多く作成されますが、コーパスに基づいた教育語彙表の作成は最近まであまり見られませんでした。最近出てきた主な教育語彙表の一つは数年前にできた北大の語彙表があります。アカデミックなものではJACETの8000が最近出ました。コマーシャルなものではアルクのSVL12000が挙げられますが、初期の頃に私もかかわっていたので、どのようなつくり方をしたかを紹介します。

まず、大きな1億語のイギリス英語コーパスであるBritish National Corpusを基に見出し語リストを作成します。その見出し語リストには既存の学習語彙表に出てくる語彙がたくさん出てきます。そこで、Palmerが作った昔のものとか、色々な基本語リストがありますので、そのようなリストをたくさん集め、それをデータベース的に扱って、何々という単語はどのリストに出ているかということ各単語に関してすべて列挙して重み付けをするわけです。最終的に30種類くらいの学習語彙リストの重みづけのデータベースを作って、それを係数として最初の見出し語リストにかける形で、どういう単語の重要度が高いかを計算してSVLを作りました。最後に付け加えたのは日本語のコーパスデータからの情報も取り込むことです。アルクの社内で作成した日本語コーパスをもとにして、日本語の用法を取り出し、どういうフレーズがよく出てくるかを検索し、英辞郎と照らし合わせた上で、日本人がよく使うような表現はリストに入れました。SVLはこのようにして出来上がりました(図2参照)。

## 教材とコーパス（投野）

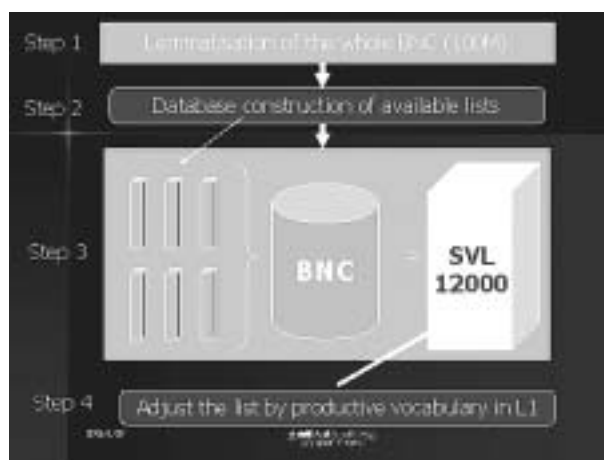


図 2：ALC SVL 12000 の作成過程

教育用語彙表の作成には、もう一つJACETの発想があって、これはもっとCorpus-Basedです。図3を見てください。

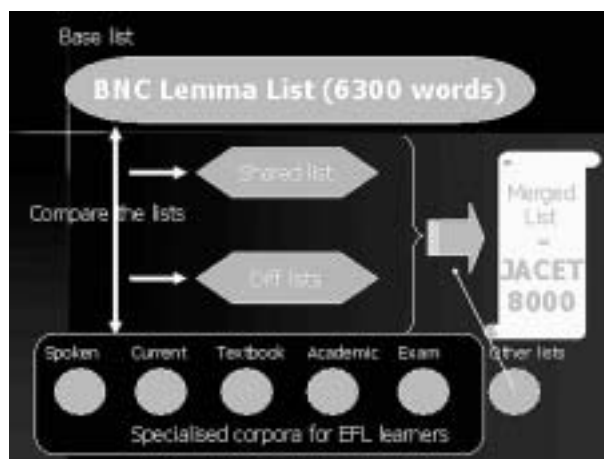


図 3：JACET 8000 の作成過程

既存の語彙表を利用するのではなく、まず、自分達でたくさんコーパスデータを作成しました。分野ごとに集められたコーパスデータ、例えば、さまざまな英語教科書のデータとか、話し言葉のデータ、時事英語などのコーパスデータを作成し、それぞれから抽出した頻度リストを、ベースになるBNCの見出し語リストと合成するような作業をしました。一致点と相違点を統計的に見て、どの語を採用し、どの語を落すかを我々で研究して、その一つの結論として現在JACET8000を提案しているわけです。このように大量のコーパスデータがあると、何が基本

データなのか、どのようなテキストが、どのような単語を使っているかを簡単に調べることができますから、おそらくこういう単語は学習する必要があるだろうということを推測することができます。コーパスは確実にその基礎資料を提供してくれます。我々のワーキンググループでは単語集を作成できないかという話が出ていまして、語彙表はその一步前の資料ですが、今後、色々な形で活用される可能性があります。単語帳だけではなく、教科書を作っても良いわけで、多くの可能性も秘めています。

## (2) 既存の教材の分析評価

コーパスデータは、もう少し違う方面へも応用されています。既存の教材の分析評価がそれに当たります。ドイツの Dieter Mindt がグラマーの教科書の分析をコーパスを利用して行っています。グラマーの教科書の英文には不自然なものが多いということはよく言われることですが、コーパスデータに照らし合わせることにより、どんなテキストを作ったら良いのかの提案が可能になります (Mindt 1995)。スウェーデンの Ljung も英語教科書の分析を行っています (Ljung 1990)。彼は、コーパスデータに比べて、教科書には具体的な表現ばかりが出てくることを指摘しています。抽象的な概念を表現したり、メンタルな心の状態を表す表現が非常に少ないことは私も感じています。「100語」をやっていた時に、教科書には具体的な表現ばかりがまず導入されていて、抽象的なものに移行するのが遅いことを実感しました。haveの動詞としての使い方を見ても、よく用いられる高頻度の表現は抽象的なもの (have a look とか have no time など) が多いわけです。教科書も “I have a pen.” で始まり、具体的な表現からより自然な用法にどう移行していくかということを経験教材の視点から考えないといけません。国内でもそういうことを気にしている人が出てきて、教科書分析をコーパスに基づいて行う動きが、最近 1 ~ 2 年、学会では目立つようになってきました。

## (3) シラバス・デザイン

シラバス・デザインでは、主にコーパスに基づいているものに Lexical Syllabus があります。J. Sinclair と A. Renouf の二人が最初に提案しました。その後、D. Willis によって Lexical Syllabus という本 (Willis 1990) が出て、最近の Susan Hunston の本 (Hunston 2002) でも「コーパス中心のシラバスは語彙が中核となる」と主張しています。私の「100語」も語彙ベースの 100 回のアプローチですので、これも Lexical Syllabus の一種だと言えましょう。

教授法の変化も海外ではかなり論議されています。

### 教授法の変化

Three P's (Presentation - Practice - Production)

Three I's (Illustration - Interaction - Induction)

DDL (Observation - Classification - Generalization)

従来のプレゼンテーションの仕方は演繹的 (deductive) で、いわゆる「3つのP (Three P's)」といわれます。まず先生が教えようとしている表現に関する文法的なルールを示し

(presentation), 学習者は練習を繰り返した後 (practice), 実際に使用する (production) というプロセスを取ります。それに対してコーパスデータに基づく教授法は帰納的 (inductive) で、これを「3つのI (Three I's)」と称します。先生は、まずある表現が実際にどのように使われているかをイラストレートします (illustration)。学習者はまず実際に使用されている例を見て、次に、データとインタラクトしつつ (interaction), そこから帰納的に結論を得る (induction) わけです。こういう学習活動の連鎖がコーパスを使うと可能になると言われています。Data-driven Learningも考え方は同じで、コーパスが利用できることにより、教授法の切り口にも変化が起こっているということです。

#### (4) 言語テスト開発

コーパスを利用した言語テスト開発は海外では徐々に出てきています。ランカスターの Charles Alderson はコーパスを利用する可能性について述べており、「コーパス準拠の Cloze Test をつくる」という論文を書いています (Alderson 1996) が、彼自身はコーパスを使ったテスト開発は行っていません。他に、ポーランドの研究者がコーパスから自動的に文法問題をつくるシステムの開発を行っています (Coniam 1997; Kaszubuski and Wojnowska 2003)。最近では、ケンブリッジ大学の UCLES というテストの部門で、大量のテストデータに基づいた学習者コーパスを編纂しています。学習者の作文データをさまざまな形で分析して、レベル毎にどんな誤りをするかというデータベースが出てきていますので、これをもとに語彙や口語のテスト開発に役立てていくことが可能になっています。国内での最近の動きですとアルクの SST (Standard Speaking Test) のデータが 1,280 人分ですが、コーパスとして公開され、この間、書籍の形で出版されました (和泉他 2004)。ただ、これはテストの結果をデータとしてコーパスにただけですので、テスト開発にフィードバックする段階までには至っておりません。これまで、間接的な利用のお話をしましたが、主要なコンセプトは教材開発に生かす一歩手前の、教材やテストを開発するための基礎資料を手に入れるのに色んな分野でコーパスが使われているということです。

### コーパス・データの直接利用 (Direct use)

#### (1) 辞典、文法書、語法書

英語教育におけるコーパス利用の第二の分野として、その直接利用を取り上げます。まず、コーパスは辞典の編纂に使用されます。辞典名とその依拠するコーパスを、出版年を追って見ると、どんどん辞典開発とコーパスの関連が深くなっていること、色んな形でコーパスと辞典が結びついていることが分かります。現在主要なものとしては海外では 5 点、Big 4 が Big 5 になっています。CALD, OALD, LDOCE, COBUILD, Macmillan は全てコーパス準拠を謳っています。国内では井上永幸先生と赤野一郎先生が編集された『ウィズダム英和辞典』が最初のコーパス準拠の辞書です。小学館の『小学館 - ケンブリッジ英和辞典』も私が日本語の監修をしました。ケンブリッジの 6 億語のデータを、日本人用に出し直して使っています。また、ケンブリッジの学習者データを我々のチームが使って、COMMON LEARNER ERROR というコ

ラムを書いたりしました。

文法書、語法書もコーパス準拠のものが数点、出始めています。文法書は、今まで人間に限られたデータの中で観察したり、直感に基づいて書かれていたものから、コーパスから得られるデータに基づいて文法書を書くという発想になっています。極端なのはロングマンのもので、D. Biber の *Longman Grammar of Spoken and Written English* ですが、こんなところまで書いても良いのかと思うところまで書いてあります。非常に、割り切って書かれた文法書です。

## (2) 英会話教材

私の英会話教材はコーパス準拠ですが、このアイデアは私がオリジナルではなく、10年くらい前にすでに、COBUILD シリーズで会話教材が出ています。しかし、ほとんど売れませんでした。マーケティングが下手というか、営業がだめでアツという間に廃本になりました。私のものはテレビが助けになって、「コーパス君」というキャラクターを中心に、年少者にも使ってもらえるようになりました。こういう形で人気が出てくるのも悪いことではないなと思っています。ただ、10分間だけ見ていると、有り難味が無いみたいですが、本当は物凄く手間隙がかかっているのです。図4をご覧ください。

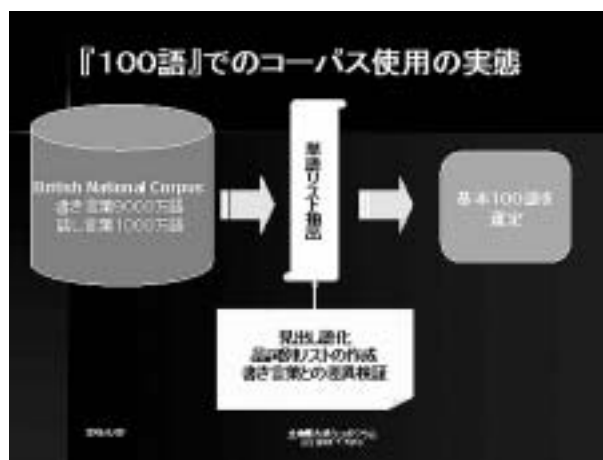


図4：「100語」でのコーパス使用の実態

まず、話し言葉1,000万語のコーパス・データから単語リストを作成しますが、この時点で総語数1,000万語のコーパスから約57,000語の見出し語リストが切り出せることが分かります。これから、最もよく用いられるトップ100語を抽出します。これは誰でもできることですが、この基本語100語を選定して面白かったのは、見出し語57,000語を頻度順に並べると最初の100語でコーパスの総語数の67%に達するという事実でした(図5参照)。



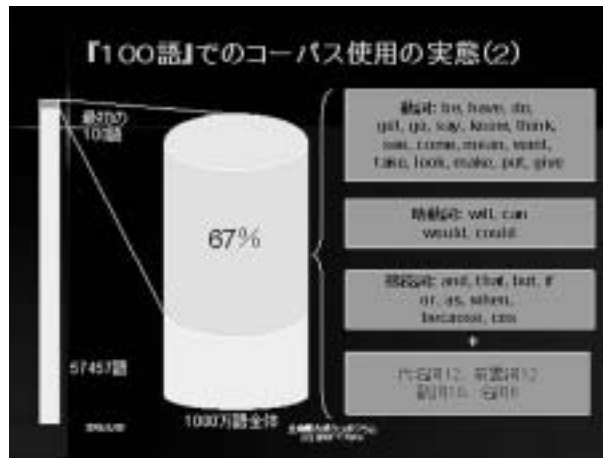


図5：話し言葉での100語が占める割合

100語知っている与会話の骨組みになるような言葉が7割近く言えることになります。もちろん句動詞とか、色んな複雑な用法がありますが、基本語を知っていれば、相当骨組みを作れるということなんです。あとの残りの30%は、そこに入れていくテーマ毎の、あるいはトピック毎の内容語なのです。骨組みをきちっと覚えて、しっかり使いこなすことができれば、後はトピック語彙をたくさん増やすことになります。これが語学学習のメリハリでしょうが、それができていなくて、平面的に文法書を頭からこなしていくからダメなのだと思います。

こういった100語の選定に加えて、1つ1つのキーワードの使いこなすを練習するために、コロケーションのテーブルを出したり、クラスターを出して、内容語に関する情報を分かりやすい形で提示することにしました。ただ、「100語」の場合には、それだけでは面白くないので、ランクされた高頻度フレーズが入ったスキットを作り、現地ロケをして、という大変な作業をしていることも知っていただきたいと思います。

### コーパスと教材：今後の課題と問題点

ここからはシリアスな問題で、今後の課題と問題点について述べます。皆さんの中にも「コーパスはいいな、すごいな」と思っている方もあるかと思いますが、色々問題点が無いわけではありません。一つは中身の問題です。2番目にインターフェイスの問題、3番目に教室内の言語活動としてのコーパスの直接利用の問題を取り上げます。

#### (1) コーパスの中身の問題

まず、中身の問題ですが、corpus-basedと言うからにはベースになる中身がどのようなものなのかを問う必要があります。コーパスの中身を検討せずにcorpus-basedの教材を作っても駄目だと思います。一般コーパスから取り出す例文は英語教育には不向きだと思われるからです。ヨーロッパ辞書学会でも、自然な言語使用の生データから例文をいきなりそっくり取った教材

は、一般の英語学習者には厳しいと言うような議論がたくさん出ています。何が教材としてオーセンティックなのかが問題になります。我々の問題としている日本の学習者のレベルでいくと、いきなり大規模一般コーパスから英語をすぐに抜いてきて、教材にできるかと言うと、非常に厳しいと思います。大規模コーパスを使う時には、語彙の頻度や連結パターンを考慮しながら、それをいかにうまく教材の中に仕込むかがカギだと思います。

もう一つのアプローチは、例文を直接利用できるような難易度調整済みのコーパスを作ることです。こういう試みがちゃんとされてくると、現場の中高の教室内で実際に使ってみて、よかったという報告がされる可能性が出てくると思います。香港や台湾では、大規模コーパスを使って、語彙レベルでフィルタリングする試みを行っています。難しい単語や固有名詞が入っているものを排除して、やさしい例文だけ出すという方式ですが、それでも内容については問題が残るのではないかと思います。

## (2) インターフェースの問題

2 番目はインターフェースの問題です。一般の要求に応じて誰もが利用できるコーパスは、その検索がやさしいことが条件です。検索が難しいと利用してもらえません。英語コーパス学会のようなおたくっぽい先生方ではなく、一般教員が何時でも使えるようなソフトでないといけないわけです。そこで、Web 上に大規模コーパスの BNC を載せて、ネットワーク上で利用できる簡単なソフトを小学館と開発し、ちょっとずつ立ち上がってきています (小学館コーパス・ネットワーク <http://www.corpora.jp>)

## (3) コーパスと言語教育の本質論

最後は、コーパスと言語教育の本質論との関連で 3 つほど問題が残っています。まず、コーパスに懐疑的な人がよく問題にすることで、私もよく考えなければならないことだと思っていることに、頻度偏重、オーセンティシティの偏重ということが挙げられます。1998 年に Guy Cook が Ron Carter の疑問に対するリプライで *ELT Journal* に書いて指摘した問題点です (Cook 1998)。H. Widdowson も 2000 年に「コーパスデータは文脈からとりだして教室内で使った途端に不自然になる」と *Applied Linguistics* に書いています (Widdowson 2000)。それがオーセンティックなものであればあるほど、一文だけを取り出すと、分からなくなるということです。どんなコーパスからとってきても、discourse から切り離されているものはだめではないかということが一つあります。

もう一つは私がやっていることに対する批判です。パターンをとってきても、そのパターンをもとにしてシンプルな例文をつくと不自然になることが起こりうるわけです。修飾語句がついているから自然なので、骸骨のようなパターンをとってきてシンプルな 10 語の例文を書いても何となくピンとこないことがあるわけですね。一番頻度が高いから、実は一番重要ではないのではないか。これはよく考えなければならないことです。

3 つ目に、DDL で観察、分類、一般化というようなインダクティブなアプローチを紹介しましたけれど、実はインダクティブなアプローチは大きな問題を抱えています。なぜかという、経験的にルールを取り出しても、その後、ルールをマスターして、どのようにプロダクション

に結び付けていくかについては何も触れられていないからです。ルール提示のところまでしか触れていない。その後、どういうふうに活動して習得するかについては全く提案がなされていません。コーパス言語学で英語教材を扱う時には、そのあたりをよく考えた議論をしていかないと、おそらく袋小路に入るという感じがします。

#### 参考文献

- Alderson, C. 1996. 'Do corpora have a role in language assessment?' in J. Thomas and M. Short (eds.) *Using Corpora for Language Research*, pp. 248-259. London: Longman.
- Aston, G. (ed.) 2001. *Learning with Corpora*. Houston, TX: Athelstan.
- Aston, G., Bernardini, S. and Stewart, D. (eds.) 2004. *Corpora and Language Learners*. Amsterdam: Benjamins.
- Burnard, L. and McEnery, A. (eds.) 2000. *Rethinking Language Pedagogy from a Corpus Perspective*. New York: Peter Lang.
- Coniam, D. 1997. 'A preliminary inquiry into using corpus word frequency data in the automatic generation of English language cloze tests'. *CALICO Journal* 16/2-4: 15-33.
- Cook, G. 1998. 'The uses of reality: a reply to Ronald Cater.' *ELT Journal* 52/1: 57-64.
- Granger, S., Hung, J. and Petch-Tyson, S. (eds.) 2002. *Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching*. Philadelphia: John Benjamins.
- Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- 和泉絵美, 内元清貴, 井佐原均 (編著) 2004. 「日本人1,200人の英語スピーキングコーパス」(アルク)
- Kaszubski, P. and Wojnowska, A. 2003. 'Corpus-informed exercises for learners of English: the TestBuilder program' in E. Oleksy and B. Lewandowska-Tomaszczyk (eds.) *Research and Scholarship in Integration Processes: Poland - USA? EU*, pp. 337-354. Łódź University Press.
- Kettemann, B. and Marko, G. 2002. *Teaching and Learning by Doing Corpus Analysis*. Amsterdam: Rodopi.
- Leech, G. 1997. 'Teaching and language corpora: a convergence' in A. Wichmann, S. Fligelstone, A. McEnery and G. Knowles (eds.) *Teaching and Language Corpora*, pp. 1-23. London: Longman.
- Ljung, M. 1990. *A Study of TEFL Vocabulary*. (Stockholm Studies in English 78.) Stockholm: Almqvist & Wiksell.
- Mindt, D. 1995. *An Empirical Grammar of the English Verb: Modal Verbs*. Berlin: Cornelsen.
- Sinclair, J. (ed.) 2004. *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins.
- Sinclair, J. and Renouf, A. 1988. 'A lexical syllabus for language learning' in R. Carter and M. McCarthy (eds.) *Vocabulary and Language Teaching*. London: Longman.
- Tan, M. 2002. *Corpus Studies in Language Education*. Bangkok: IELE Press.
- Wichmann, A. Fligelstone, S. McEnery A. and Knowles, G. (eds.) 1997. *Teaching and Language Corpora*. London: Longman.

Widdowson, H. 2000. 'The limitations of linguistics applied'. *Applied Linguistics* 21/1: 3-25.

Willis, D. 1990. *The Lexical Syllabus: A New Approach to Language Teaching*. London: HarperCollins.