

はじめに

前回は学習者コーパスを用いて、基本動詞の習得過程をネイティブ・スピーカーの使用法と対比してみるという試みを紹介した。今回は、学習者コーパスによって言語習得途上で犯す誤りのパターンをとらえ、その変化を追うことで中間言語の変容していくようすを捉えようとする試みを紹介しよう。

学習者コーパスによるエラー分析

学習者コーパスは、作文や会話などのプロダクション・データをコンピューターに大量に蓄積して、その特徴分析をコーパス・ツールを駆使して行なう。エラー分析は、その中でも現在最先端の研究分野である。従来の error analysis は日本語では「誤答分析」と訳されていたが、学習者コーパスを用いた分析では「誤答」という用語は避けた。コーパス・データの採取方法が、特別な「正解のある」テストに基づくものではなく、学習者の自然な発話データや作文データであるので、「誤答」という言い方は少し誤解を招くからである。

さて、エラーの分析方法であるが、コンピューターはデータを見て自分で誤りを認識して処理してくれたりはいしない。現在、一般の現代英語コーパスに関しては自動で品詞のタグや統語構造の情報を付与するプログラムが存在するが、エラー情報に関しては全部マニュアルである (Milton 1994; Meunier 1995)。

そこで、研究する我々の側でどのようなエラーに関してどうやってタグ付与を行ない、それをもとにどういった分析を施すかといったことを組織的に研究して行かなければならない。そのためには、エラータグの体系のようなもの(これを tagset という)を決めなければならないのだが、これもどいう視点でエラーを捉えるか、分類するか、でずいぶん異なる。コーパスを用いた分析では、汎用的なエラータグセットの開発はかなり難しいので、現状ではとりあえず自分の研究目的に即した

タグセットを作り、試験的にタグ付与をしてみて、その結果を他の研究者とシェアし合っているという形である。

将来的には、ある種の機械的なエラーについては自動的にエラー検知がされてタグ付与がなされることは大いにあり得るだろう。ちょうど現在ワープロソフトに添付され始めている grammar checker のようなものの精度を上げたものをイメージしてみればよい。

もしこのようなエラー情報がきちんと付与された学習者コーパスが整備されれば、我々は学習者の使う英語に学年を追ってどのような誤りがあり、それがどういう風に変化して行くかを捉えることが可能だ。もしこのような発達的な側面をエラーの変化、定着度の変化、という視点で捉えていけば、第2言語習得研究のみならずシラバス作成や指導書の内容にも大きな影響を及ぼすことになるだろう。

誤答分析の反省

誤答分析 (error analysis) は、S. P. Corder (1967) によってその重要性を指摘されてから 70 年代にかけて非常に活発に研究がなされた分野であるが、いくつかの問題点のゆえに 80 年代になって急速に興味を失われてしまった。1 つには、中間言語を誤答の観点からしか見なかったため、正しく習得されている部分の関係にはまったく焦点が当てられなかったという事実がある。これに光を当てたのが、ZISA project (L2 German の大規模な習得度調査: これにより Clahsen, Meisel といった研究者がドイツ語の語順の習得に普遍的な段階があることを主張した) などのデータをもとに multidimensional model (現在は processability theory という) を構築した Pienemann だった (Pienemann 1998)。

第2に、誤答の分類はかなり大きな議論の争点になったわけだが、ある誤りをどのような誤答に分類するかで、研究者の意見が割れるようになった。つまり誤答の表面的な分類のみでなく、その原因(たとえばその誤答が母語の干渉が原因で起き

学習者コーパスと

たのか、それとも習得中の第2言語内での規則性から出てきたのか、など)で分類を始めた途端に、コンセンサスを得られなくなってしまったのである。それほど誤答のパタンの分類はいろいろな角度から切ることが可能で、難しいのである。

経験的事実をきちんと示す

90年代に入って、コーパス言語学の急速な発達によって学習者の話す中間言語のコーパス化という新しい領域も脚光を浴びるようになった。しかし、エラー分析に関しては、1つ誤れば20年前と同じような結果になりかねない。この点については、ランカスター大学のTony McEnery氏も私の論文指導の際に再三警告していた。言語理論を背負ってそれにもとづいた解釈で分析をすることはコーパス言語学の宿命ではあるが、理論が変わっても言語事実がなるべくはっきりと提示できるように、「経験的事実」に重きを置くようになりサーチの方向性が必要である。ここでは、そのような具体例として、私のチームの研究成果を部分的に紹介しよう。

文法事項のエラーパタン別頻度の推移

英語学習者は習得途上の文法でも一定の規則に基づいて発話している。完全に目標言語と同じ使い方ができる場合とできない場合などの個人内での揺れや、書いているときちゃんと使えるが話すときだめというようなproductionのモードによる違いもある。しかし、学習者コーパスはタスクや環境の変数をコントロールして同一レベルの学習者データを大量に集めて見ることで、個人差で終わってしまいそうな習得の「大筋」をよりはっきりと捉えようとするのである。

その一例として、Aoki (1998)を見てみよう。Aoki (1998)は中学生・高校生が学習する主要な5種類の文法事項(現在時制、受動態、関係詞、仮定法)が導入されてから習得されるまでの定着度の推移を和文英訳の課題と自由作文をコーパス化して調査した。中学1年生から大学2年生までの各学

年延べ869人を被験者にしており、すべての被験者から作文データを取っている点で大規模で、学部卒業論文ではあるがデータそのものも非常に興味深く参考になる。

次ページの図1がAoki (1998)による、主要文法事項の定着度のグラフである。これを見る限り、一般的に難しいと考えられている関係詞や仮定法は定着が早いようであるが、これは実際にはコーパスに出てきた出現率が低いので、あまり短絡的に判断はできない。Aokiの研究で、興味深いのはむしろ出現率が多かった現在形の用法である。それについて少し詳しく見てみることにしよう。

現在形の導入から定着の時差

ここではすべてのデータを示すことはできないが一例として現在時制のエラーパタンを例に見てみよう。次ページの表1はAoki (1998)をもとに現在形の学年を追うごとの正答率とエラーパタンの推移を見たものである。中学1,2年は比較的現在形の誤りは少ないが、中学3年から未来形が導入されてくると現在形の過剰使用(overuse)が見られる。未来形を使うべきところを現在形で済ませてしまうというエラーだ。

現在形の場合は形そのものを間違えるというエラーは非常に少ない。むしろ、他の時制との関連での用法上の間違いを多く犯すということがわかる。そしてこのような傾向は、紛らわしい時制がある場合には導入されてから高校3年くらいまで続き、大学になると終息する。

もちろん、Aoki (1998)の実験結果は英作文のトピックや目標となる教科書などの文法事項の取り扱いの影響などを考慮して注意深く解釈されるべきであるが、それでも単純だと思われる現在時制がこのように多くのエラーパタンの変遷を経て、定着して行くというのは興味深い事実ではないだろうか。

corpus-based なシラバス構築の可能性

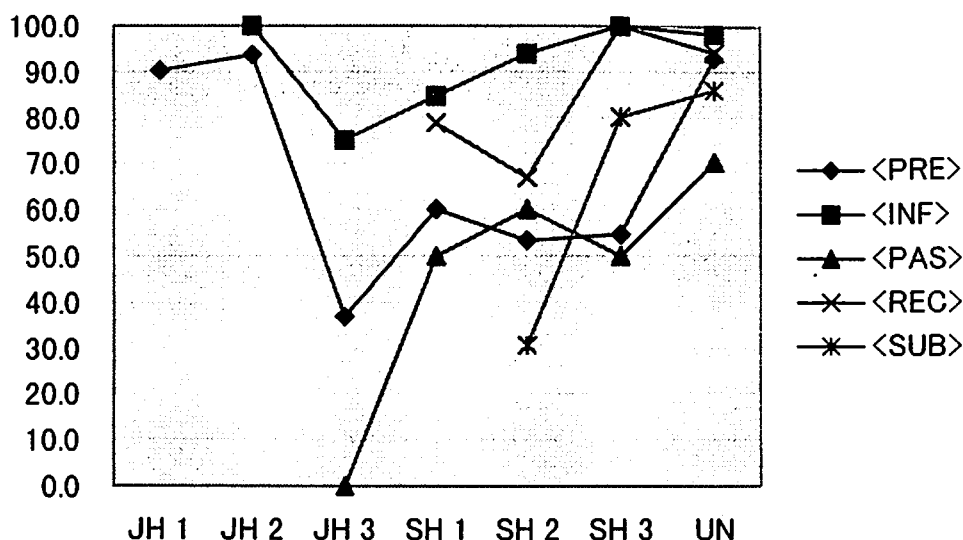
このような結果を受けて、我々はいくつかの仮



……(5) 学習者のエラー分析

words 投野由紀夫

図1 5種類の文法事項の定着度 (Aoki 1998 に基づく)



(注) <PRE>: 現在時制 / <INF>: 不定詞 / <PAS>: 受動態 / <REC>: 関係詞 / <SUB>: 仮定法

現在形	中1	中2	中3	高1	高2	高3	大学
エラー: 過剰使用 (overuse)	4%	2.2%	58%	27%	37%	32%	4.1%
エラー: 省略 (omission)	0%	0%	2.6%	6.7%	3.8%	4.5%	0.8%
エラー: 形の誤り (incorrect form)	3%	4.4%	2.6%	5.5%	6%	4.5%	2.4%
エラー: 過剰使用でかつ形の誤り	3%	0%	0%	0.8%	0%	4.5%	0%
正しく使われていた場合	90.2%	93.4%	36.8%	60%	53.4%	54.5%	92.7%

表1 現在形の正答率とエラーパタンの推移 (Aoki 1998 に基づく)

説を立てることができる。1つには、現在時制が未来時制と混同されるのは、時制の概念が時間を置いて1つ1つ導入されるからだとも考えられる。そこで、現在形と未来形を対比して一緒に教えてみるとか、日本語と比較して同時期に複数の時制を提示してみるなどの提示順序や方法に関する改善が必要ではないかと思われる。One at a time というのは学習理論上は効果的な面もあるが、言語というのはそのような1つ1つ確実に積み上げるというものではなく、ある総体が捉えられると細かい部分もすっと身についたりすることもある。そのような言語特有の学習過程に敏感なシラバスを今後も追求していく必要があるだろう。

もちろん、これは時制だけに限った問題ではないから短絡的な判断はできないが、少なくとも英語学習者の長期的な定着度の調査が学習者コーパスによって十分に行なわれてくれば、行き詰まっている日本の英語教育の改善に一石を投じられる

ことは間違いない。

【参考文献】

Aoki, M. (1998) *A Corpus-based Analysis of Japanese EFL Learners' Process of Interlanguage Construction*. Unpublished B.Ed. thesis. Tokyo Gakugei University.
 Corder, S. P. (1967) "The significance of learners' errors." *International Review of Applied Linguistics* 5: 161-69.
 Meunier, F. (1995) "Tagging and Parsing Interlanguage." In L. Beheydt (ed.) *Linguistique appliquee dans les annees nonante*, ABLA papers no. 16: 21-29.
 Milton, J. (1994) "Tagging the interlanguage of Chinese learners of English." In L. Flowerdew & A.K.K. Tong (eds.) *Entering Text*. Language Centre, Hong Kong University of Science and Technology: 127-43.
 Pienemann, M. (1998) *Language Processing and Second Language Development.: Processability Theory*. (Series: Studies in Bilingualism) J. Benjamins.
 (とうの・ゆきお / 元東京学芸大学講師
 ランカスター大学博士課程在籍)

