

2017年3月10日
第二回学習者コーパスワークショップ

学習者コーパスと 第二言語習得研究

コーパスを使ってできること・できないこと

麗澤大学 外国語学部
大関浩美

本日の話の内容

1. コーパスを使った研究方法をめぐって
2. コーパスを使った研究の貢献
英語L1関係節習得研究の事例から
3. コーパスを使ってどのようなSLA研究が可能か

コーパスを使ったSLA研究

- * 産出データをみることは、もともとSLAの中心的方法のひとつ
- * これまでは、データの収集に時間がかかり、収集・文字化に時間を費やした
- * コーパスにより、大量の産出データを利用・共有することが可能に
- * コーパス・ツールの利用で、自分の見たいものを簡単に取り出せるようになった

1. 学習者コーパスを使った研究方法 (Granger 2002, 2012など)

- * Contrastive Interlanguage Analysis (CIA)
対照中間言語分析
- * Computer-aided Error Analysis (CEA)
コンピュータ支援による誤用分析

Contrastive Interlanguage Analysis (CIA)

1. 学習者データを母語話者データと比較することにより、学習者の中間言語特有の特徴を明らかにする

- 頻度の比較⇒過剰使用・過少使用がわかる
- 学習者言語のnon-nativenessな特徴がわかる

Granger (2002)

Contrastive Interlanguage Analysis (CIA)

2. 母語の異なる学習者間、学習環境の異なる学習者間、年齢の異なる学習者間などの比較により、母語や習得環境その他の変数の影響を探る

Computer-aided Error Analysis (CEA)

トラディショナルな誤用分析に比べ

- * コーパスサイズが大きいため、信頼性がある
- * 誤用と正用の比較ができる

Granger (2002, 2012など)

CIA①の研究例 Altenberg and Granger (2001)

- * 作文コーパスで、makeの使用を比較
フランス語NS vs. スウェーデン語NS vs. 英語NS
- * フランス語NSはmakeを過少使用
- * makeの使役をスウェーデン語NSは過剰使用、フランス語NSは過少使用
- * makeの使役で使われる動詞が、学習者と英語NSでは異なる

CIA②の研究例

Hokari・Kumagami & Akimoto (2012)

- * KYコーパスで2種類の複合動詞を分析
- * 英語・中国語・韓国語母語話者を比較
 - * 統語的複合動詞(例:歩き始める)
 - * 語彙的複合動詞(例:取り締まる)

Hokari・Kumagami & Akimoto (2012)

- * 学習者の母語での複合動詞は
 - * 英語:無し
 - * 中国語:2種類の複合動詞どちらも有り
 - * 韓国語:語彙的複合動詞のみ有り

Hokari・Kumagami & Akimoto (2012)

- * 結果
 - * 語彙的複合動詞が早く使われる
 - * 産出量も語彙的複合動詞のほうが多い
 - * 母語別では、韓>中>英だが、
統語的複合動詞は、中>韓>英

ただし...

- * データ数が非常に少ない
- * 複合動詞のように学習者にあまり使われない形式は、コーパスでは分析にくい
- * より大規模なコーパスが必要
- * 実験研究での検証も必要

異なる母語話者コーパスの 比較は重要

- * 母語の転移をみるには、複数の言語の母語話者を比べる必要がある
- * しかし、2言語以上の母語を対象とした研究は少ない(奥野2003)
- * 「○○語母語話者の○○の習得」をみただけでは、言語転移は可能性しか言えない

しかし、CIA①とCEAは SLA研究の観点から見ると...

- * SLA研究者からは批判の声が多い
- * CIA①(母語話者との比較)・CEAは、Comparative fallacyを引き起こしてしまう研究方法

Comparative fallacy (比較の誤謬) Bley-Vroman (1983)

- * Comparative fallacyとは...
母語話者の言語と比較して分析することにより、中間言語体系を見誤ってしまうこと

Gass and Selinker (2008)

- * SLA研究のゴールは、「第二言語の根底にあるシステム」を明らかにすること
- * 学習者の第二言語を目標言語の「規範」と比べると、学習者の言語体系が持つ “systematic nature” を理解することが不可能になってしまう。

目標言語を規範とした見方は...

* target deviation perspective

(Klein 2011)

= 目標言語からの「逸脱」を見る見方

「学習者が何をしているか」という観点ではなく、「学習者が何ができないか」という観点から学習者言語を見る見方

中間言語分析の必要性

* 母語話者の言語を「規範」として「逸脱」や「誤り」のみに注目するのではなく、学習者の作り出す言語体系 (= 中間言語) を分析しなければ、第二言語の発達はわからない。

学習者の作る言語体系とは...

目標言語を「規範」とした見方では...

- * 今日はうちでご飯を食べます。 ○
- * パスポートはかばんの中にあります。 ○
- * その本は図書館であります。 ×
- * 駅の前に友だちと会います。 ×

(追田2001の結果から作例)

学習者の作る言語体系とは...

目標言語の「規範」をはずすと...

- * 今日はうちでご飯を食べます。
- * パスポートはかばんの中にあります。
- * その本は図書館であります。
- * 駅の前に友だちと会います。

迫田(2001)で見られた 学習者の「に」と「で」のシステム

- * 中級学習者を対象に、穴埋めテストで調査
「位置を示す名詞(中・前など)」+に
「地名や建物を示す名詞(東京・食堂など)」+で
- 前置される名詞の種類によって「に」と「で」
を使い分けるシステムを作っていた

母語話者データの使用

- * 母語話者のデータは必要
- * ただし、「規範」としてではない
- * 大関(2008)
 - * 進行中の動作を表す名詞修飾節が非常に少ない
(例)「あそこで本を読んでいる人」
 - * しかし、母語話者のコーパスでも少ない
⇒「学習者が使わない」「学習者にとって使いにくい」のではない

母語話者の使用を「規範」としない 比較を

- * 韓国語の、「に」と「で」に相当する助詞を研究する際、
日本語の「に」と「で」を「規範」として分析するか？
- 中間言語も自律した言語体系として分析を
- * まず母語話者を規範とせずに、言語体系を分析
 - * その上で、母語話者の使用を見る
 - * 母語話者のデータは、インプットを考える参考にも

2. コーパスを使った研究の貢献

英語L1関係節習得研究の例から学ぶ

The cat [that chased the mouse].

[ネズミを追いかけた]ネコ

(ネコは関係節の主語の関係) = SU

The mouse [that the cat chased].

[ネコが追いかけた]ネズミ

(ネズミは関係節の目的語の関係) = DO

SU > DO (SUのほうが易しい)と予想されてきた

従来の実験研究で使われてきた文は

- * **The horse [that bumped the rabbit]** kicked the sheep.
- * **[ウサギさんにぶつかった]ウマさん**がヒツジさんを蹴った

他動詞文に他動詞文の関係節が埋め込まれた構造

⇒一致した結果が出ない

ところが、コーパスを見てみると...

- * Diessel & Tomasello (2000)
CHILDESデータベースの4人の子どもの発話を分析
↓ 結果
- * 実際に子どもが習得の初期に使用するのは...
This is ... That is... Here is ... などの構造
(例) Here's a tiger [that's gonna scare him].

実験で使われるような構造は、ほとんど使われない

そこで、 コーパスからの結果を生かした実験

- * Diessel & Tomasello (2005)の実験研究
子どもが実際に使う構造を使用
This is/Here is + N + 関係節
↓
はっきりとした結果が得られた(英語L1・ドイツ語L1)
SU>DO

さらに、コーパスを見て初めて わかったこと

- * 英語L1児の初期の関係節構造は、
presentational relative clause
(例) Here's a tiger [that's gonna scare him].
虎があの子を脅かそうとしてるよ!
(Diessel & Tomasello 2000)
複文構造ではあるが、全体で示す命題は一つ

子どもの関係節は、命題が一つの(=単文構造に近い)
関係節から、徐々に本格的な複文構造へ発達

コーパスを見ることの必要性

- * 実験研究では、仮説の検証はできる
- * しかし、言語使用の「事実」はわからない
ある言語形式がどんな形で使われているのか
どんな機能で使われているのか
それがどのように変化(発達)していくのか
↑
コーパスを見ることで、明らかにできる

さらに、コーパスからわかったこと

- * 英語・独語L1 (Kidd, Brandt, Lieven & Tomasello 2007)
幼児の使うDO(直接目的語の関係節)の多くは、
無生の被修飾名詞+節内は有生の代名詞
例: the **book** [**I** read]
実験研究でも: このパターンは易しい
- * 成人NSのコーパスでも、このパターンは高頻度で、
実験(文処理)研究でも、プロセスが易しい
Reali & Christiansen (2007)

これらの研究から言えること

- * あるパターンの使用頻度が高い
↓
インプットに何度も接触し処理を繰り返す
↓
習得に影響
処理も早くなる

Reali & Christiansen (2007)

これらの研究では...

- * コーパスを使った研究(子ども・成人NS)、実験研究(子ども・成人文処理)など、様々な方法による研究が一体となって、一つの方向を示した例

重要な役割を果たしているのが、コーパス研究
どう使われているか、事実が明らかになる

これらの例のお手本になる点

- * インパクトのある「事実」を明らかにしている
- * 使用されるパターンを見つけ出している
- * 「発達」のメカニズムを議論している
- * コーパス研究の結果から実験を計画している
- * コーパスと実験を組み合わせた研究をしている
- * 様々な研究の結果を入れて大きな議論をしている

3. どのようなSLA研究が可能か

1. コーパスから得られた「事実」を理論的に説明する
2. SLA理論をコーパスを使って検証する

1のタイプの例 家村(2003):否定表現

- * KYコーパスで中国語母語の学習者の否定表現を分析
- * 得られた結果から、習得プロセスを考察

第一段階	多様な否定形が混在
第二段階	動詞・名詞・ナ形容詞否定形がほぼ習得
第三段階	イ形容詞否定形がほぼ習得され、 否定辞の活用形が出現
第四段階	否定辞が分析的に使用され、 多様な否定表現が出現
第五段階	否定疑問文等、多様な否定表現が習得

1のタイプの例 大関(2008):名詞修飾節

- * L1: CHILDESデータベースの5名の幼児のコーパス
L2: 3名の学習者の縦断データ(非公開)
5名の自然習得者の発話データ(非公開)
KYコーパス
- * 名詞修飾節を状態性で分類する枠組みを作り分析
 - ①「属性・状態」
 - ②「習慣」
 - ③「進行」
 - ④「過去・未来の出来事・状態」

結果

- * 日本語L1幼児、L2自然習得者
被修飾名詞の属性・状態を表す修飾節から発達
- (L2) 結婚してる人・日本に住んでる人
知らない人

(L1) 野球する物・きもの着たお化け
- 教室学習があれば、テンスのある修飾節も早くから使用

日本語の名詞修飾節の発達過程

日本語の名詞修飾節は、形容詞修飾から連続的に発達

面白い人・つまらない本
↓
結婚してる人・知らない人・日本に住んでる人
↓
昨日行ったレストラン

語彙から文法への発達の連続性
= usage-based modelと合致

形容詞による修飾
↓
属性・状態を表す修飾節 (完全な「節」ではなく、句に近い)
↓
テンスのある修飾節

(Ozeki, 2011)

さらに、英語L1の関係節習得と比較すると...

- * 英語では、新情報を表すものから発達
Here's a tiger that's gonna scare him.
- * 日本語では、属性・状態を表すものから発達

しかし、どちらも共通して

命題が一つの(=単文構造に近い)構造から、徐々に本格的な複文構造へ発達

2のタイプの例 (峯2007)

- * 接続辞の習得にPienemann(1998)の処理可能性理論を応用。KYコーパスを分析
- * 処理可能性理論
 - * 第二言語の発達は認知的な制約を受ける
 - * 徐々に複雑な処理が可能になる
 - * 語→句→単文→複文という発達段階に沿って言語処理が発達

峯(2007)

南(1993)の従属節の分類に発達段階を応用

- * A類(テ(付帯状況)、ナガラ) ⇒ 句処理
- * B類(ト、バ、タラ、ナラ、ノデ、テ(理由・継起)等) ⇒ 複文処理
- * C類(ケド、ガ、カラ、テ(並列)) ⇒ 文処理

予測 A類 ⇒ C類 ⇒ B類 の順に習得

結果

- * 予測された順に、言語処理能力が発達する
- * 接続辞表現の広がり方は

事実的	⇒	仮定的
順接 ⇒ 逆接		順接 ⇒ 逆接

- * 産出の際の思考的負担の少ない接続辞表現から使用が広がる

コーパスでわかること・わからないこと

- * 意味のある文脈の中で、学習者が目標言語を使って伝えたいことをどう伝えていくのか、が見られる
- * 学習者が何をどう使っているかという事実
- * どんなパターンで使っているか
- * どんな機能で使っているか
- * どんな変化のプロセスが見られるか

コーパスでわかること・わからないこと

- * 「起こらなかったこと」については解釈が困難
- * 使えないのか、使わなかったのかは、わからない
- * トピックが影響する可能性も
 - ⇒「できる」「できない」を見るのには向かない
 - ⇒「できる」「できない」は、実験研究で

ハイブリットな研究を

- * コーパスを使った研究と実験研究は相互補完的
- * 両方を組み合わせれば、より貢献度の高い研究に

(例) 研究1で、コーパスを使った研究
そこから仮説をたて、研究2で実験研究

まとめ

- * コーパスからは、学習者の言語使用の事実が得られる。
- * 何をどのように使っているかが、わかる。
- * 自律した言語体系として学習者言語の分析を。
- * 母語話者のコーパスは、「規範」としてではない使用を。
- * 母語の転移をみるには、母語の異なる学習者コーパスの比較を。
- * コーパスから得られた事実を理論的に説明したり、理論をコーパスで検証することで、貢献度の高い研究に。
- * 「起こらなかったこと」は、コーパスではわからない。
- * 実験研究との組合せも有効。