

Unit C3
L2 acquisition of
grammatical morphemes
CASE STUDY 3

富田 真由

C3.1 INTRODUCTION

- L2英語学習者のコーパスを用いる
- SLAのプロセスをより理解できれば、言語授業の実用的な面に応用することができる(シラバスデザイン、資料開発、タスクデザイン、言語テスト)
- 学習者言語(learner language)の説明への主なアプローチ(Ellis)
 1. The study of learners' errors
 2. The study of developmental patterns
 3. The study of variability
 4. The study of pragmatic features

C3.1 INTRODUCTION

- *Interlanguage* (Selinker) : 言語習得の過程で学習者が構築するとされる特別な心の中の言語体系
- コンピュータで学習者の言語運用データを大量に集めることで、つまり学習者コーパスを作って、学習者言語を調査する必要があるという意識の高まり
- Learner corpusという用語はEFL learnersに共通の間違いを提供する、Longman Learners' Corpusに基づいたLongman's learner dictionariesで初めて使われた
- 1990年、L2データを集めるためにICE projectの一部としてICLEが始まった

C3.1 INTRODUCTION

- Learner corpusがどのように形態素習得の研究に役立つのか
- 形態素習得研究はL2学習者言語の発達の説明を規定することを目指す意味では言語運用分析の一種
- 言語運用分析は次の重要な質問を調査するための基礎を提供
 - 教わる順と習得の順に違いはあるか
 - 教わることによって習得の自然な順を変えることは可能か
 - 教わった学習者は教わっていない学習者と同じ習得順か、違う習得順か

C3.2 Morpheme Studies: a short review

- Brown (1973) 英語話者の子供ははっきりした順で形態素を習得
- Dulay and Burt (1973) スペイン語話者英語学習者の子供で再現し
共通の習得順を発見
 - 形態素習得にはL1とは無関係の一樣の習得順がある
 - 学習者が特定の言語の構造を見つけることを可能にするuniversal cognitive mechanismの存在
- 形態素習得研究の様々なアプローチの出現
→ Larsen Freeman and Long (1991) それらの研究はL2形態素習得の
発展的な順序の証拠を示した
- 学習者コーパスを使って形態素習得を調査する

C3.3 The Longman Learners' Corpus

- LLCのサブコーパスを用いる
 - 研究のために公に利用できる数少ない学習者コーパスの1つ
 - L1 backgroundごとの構成要素が豊富で異なる習熟度の学習者の間違いやデータを得ることができる
- 日本人EFL学習者の作文データのサブコーパスを使って、先の研究と同じ形態素習得の順をとっているのかを調べる
- 研究を始める前に、品詞とerror typesのannotationが必要

C3.4 Problem-oriented Corpus Annotation

- 調査手順

1. 元のコーパスデータを予備処理する

→headerをWordSmithを用いるのに適したフォーマットに変える

2. POS tagger を使ってサブコーパスをタグ付けする

3. 手動で形態素の間違いをタグ付けする

4. Error tag と関連するPOS tagの数に基づいて形態素の用法の正確さの割合を得る

5. 結果を先の研究結果と比べる

C3.4 Problem-oriented Corpus Annotation

C3.4.1 Basic formatting of the text

- Header information はCOCOAに似たフォーマット
- WordSmith で使うためにSGMLフォーマットに変換する
→<RF>の前に<head>、<PA>の前に</head>
- 習熟度別に8つのファイルに分ける
(Beginning, Elementary, Pre-intermediate, Intermediate, Upper-intermediate, Advanced, Proficient, Academic studies, Degree-course)
- 8つのファイルを4つのグループに分ける
(Elementary, Intermediate, Advanced, Proficient)

C3.4 Problem-oriented Corpus Annotation

C3.4.2 Error tagging: a proper way and a dirty way

- Error informationを抽出する2つの方法
 1. コーパス全体をannotate→error tag の頻度を抽出
 - 再利用できるが時間がかかる
 2. WordSmithを使ってconcordance linesを検索し、計算のために一時的に間違いにしろしを付ける
 - 時間的な効率が良いが、データがコンピュータに保管されている間しか有効でなく、結果を二次的な目的で利用できない

C3.4 Problem-oriented Corpus Annotation

C3.4.2 Error tagging: a proper way and a dirty way

- POS tagging

(1) Be動詞 <AUXBE> <COP>

(2) Plural-s <PL>

(3) Possessive-s <POS>

(4) Progressive-s <PROG>

(5) 3rd person singular -s <3PS>

(6) Article <ART>

(7) Irregular past <IRPST>

(2)-(6)はPOS tagsを使い半自動でタグ付けできる

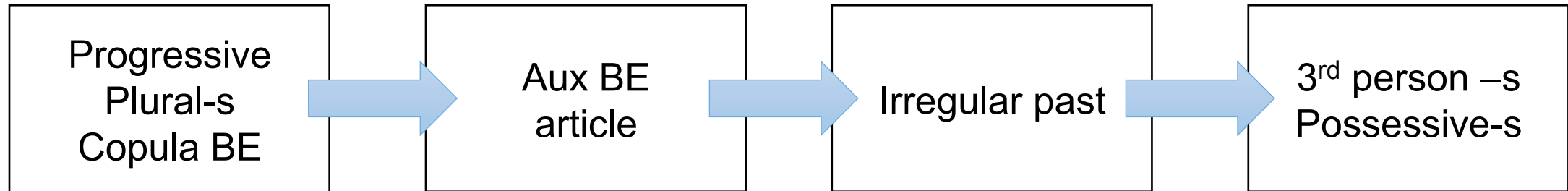
(1),(7)は手動でタグ付け

C3.4 Problem-oriented Corpus Annotation

C3.4.2 Error tagging: a proper way and a dirty way

- 手動でerrorのタグ付け
 - 1) Overuse error
 - 2) Omission error
 - 3) Misformation error
- Dulay and Burt の方法に従ってobligatory context のみを調査するのでoveruseの情報は使わない
- Dirty way: WordSmith version 4 のSET functionを使う
 - 不純だが早い

C3.5 DISCUSSION



Copula BE
Aux BE

Progressive -ing

Possessive -s

Plural-s

3rd person -s

Irregular past

article

- 日本人には冠詞が難しく、所有の-sは割と簡単
→ 日本語には冠詞がなく、所有の「の」がある

C3.5 DISCUSSION

- Dulay and Burt の研究と今回の研究の結果には違いがある

↓ implication

- 共通の習得順の概念は審議する必要がある
- LLCの習熟度のデータは主観的なことがある
デザインされていない目的でコーパスを使うと誤った結果になる

Summary

- 学習者コーパス Longman Learners' Corpus
- 日本人英語学習者がどのように形態素を習得するのか
- Dulay and Burt: L2英語学習者には共通の形態素習得順がある
- 今回の学習者コーパスを使った調査では共通の習得順とは異なる結果になった
- 手動のerror taggingの技術