

## 英語学習者の未習得言語知識習得のための冠詞誤り原因の付与 Annotating Causes of Errors for Learning Unlearned Linguistic Knowledge

小谷 克則<sup>†</sup> 吉見 毅彦<sup>‡</sup>  
Katsunori Kotani Takehiko Yoshimi

### 1. はじめに

英語学習者の英作文の支援を目的として、作文中の誤用を検出し、検出された誤用に対する正例を提示するツール(誤用検出ツール)が開発されてきた[1, 2など]。その作文支援効果として、訂正作業を通じて誤用の原因であった未習得の言語知識の習得を学習者に促し、同種の誤用の再発を防止する効果が挙げられる。しかし、一部の学習者、特に、習熟度の低い学習者にとって、提示された誤用とその正例から文法書などを用いても習得すべき言語知識を特定することは困難である。そこで、本研究は、更なる支援策として、誤用の原因となった未習得の言語知識(誤用原因)の提示に着目した。

誤用原因の対象となる文法項目は多岐にわたるが、本研究は冠詞の使用漏れによる誤用に着目した。その理由は、使用頻度の高い冠詞は誤用の生起頻度も高いため、本研究が冠詞を有しない日本語を母語とする学習者を対象とするため、冠詞の使用漏れが英語話者にとって許容度が低い誤用である[6]ためである。(以下、特に明示されない限り誤用は冠詞の使用漏れを指示する。)

まず、本研究は誤用原因を提示するツール(誤用原因提示ツール)の開発に必要な学習者の作文データ(学習者コーパス)を検討した。学習者コーパスには、作文に関する品詞や文法解析結果などの情報、および誤用に関する情報(誤用タグ)が含まれる。誤用原因提示ツール開発に必要な誤用タグは、未習得の言語知識の情報からなる誤用原因タグである。誤用タグの研究を調査した結果、第二言語習得研究[3, 4, 5]において誤用原因タグの有用性は指摘されているが、実際に作成された誤用原因タグは未習得の言語知識を提示する情報ではなく[4, 5]、未習得の言語知識を提示する場合[1]でも、冠詞の誤用は対象でなかった。そこで、本研究は誤用に対して学習者にとって未習得の言語知識を提示する誤用原因タグを作成し、誤用原因タグの分布を既存の学習者コーパスにおいて調査した。

### 2. 関連研究

まず、既存の学習者コーパスにおける誤用タグ関連研究を紹介する。多くの誤用タグは誤用の種類やその正例を対象とし、誤用原因を対象とする誤用タグは少数[1, 4, 5]であった。未習得の言語知識を提示する誤用原因タグ[1]は前置詞の誤用を対象とし、冠詞ではなかった。その他の誤用原因タグ[4, 5]は、1節でも言及したように、未習得の言語知識といった情報ではなく、ある言語知識が未習得となっている原因、例えば、学習者の母語からの影響などの情報であり、言語知識そのものではなかった。

次に、誤用検出ツール関連研究を紹介する。前置詞の誤用検出ツール[1]は「名詞以外と使用された前置詞句」や

「to が抜けている不定詞」といった誤用原因を提示する。冠詞の誤用検出ツール[2]は、誤用原因は提示しないが、正例を選択する統計的分析に寄与した素性をその正例の根拠として提示する。

### 3. 誤用原因タグ付き学習者コーパス

#### 3.1 目的

第一の目的は、誤用原因タグとして提示すべき情報を定めることである。誤用タグの情報を定める際、同種の誤用の再発防止に貢献すること、タグ付与において評価者間の揺れを最小限にすることを条件として課した。

第二の目的は、学習者の作文における誤用原因タグの分布を確認することである。仮に、ある誤用原因タグの頻度が皆無とすると、誤用原因提示ツール開発に向けてその誤用原因タグの頻度を増加させるような作文課題を設定する必要がある。

#### 3.2 手法

誤用原因タグを既存の学習者コーパス[7]に対して実際に付与し、その作業を通じて誤用原因タグとして提示すべき情報を検討した。

まず、誤用原因タグの付与対象となる学習者コーパスを選択した。その際、未習得の言語知識を作文の用法以外にも分析可能とするため多技能を対象とする学習者コーパスに着目した。そして、作文以外に読解、聴解、発音のデータを有する学習者コーパス[7]を選択した。但し、本研究では作文データ(90名の学習者による4,007文)のみを使用した。作文の課題は、四枚の絵からなる一連の出来事の描写、および英語学習やコンピュータ利用に関する20の質問への回答であった。学習者コーパス[7]には文法全般に対して誤用タグ(6,128件)が付与されていた。そのうち、冠詞の使用漏れによる誤用は439件あった。なお、冠詞の使用漏れが冠詞の誤用全体(639件)に対して7割程度と高い生起頻度であったことから、冠詞の使用漏れが誤用原因タグの課題として適当と考えられる。

次に、439件の冠詞の使用漏れをもとに、これらの誤用を防ぐために習得すべき冠詞の用法を文法書[8, 9, 10]を用いて調査した。その結果、8種類の冠詞の用法にまとめられた。各用法の説明と実例(太字の冠詞が正例)は以下の通りである。

- (1) 形容詞・副詞 *little/few* に対して不定冠詞を用いる。  
I speak **a** little English and French.
- (2) 慣用的に定冠詞を用いる。  
**The** Internet is very useful.
- (3) 慣用的に不定冠詞を用いる。  
When people have dreams, people make **an** effort.
- (4) 照応表現に対して定冠詞を用いる。  
He stopped **the** man's running with his stick.
- (5) 唯一の事物に対して定冠詞を用いる。

<sup>†</sup>関西外国語大学

<sup>‡</sup>龍谷大学

I think reading English is **the** most simple and useful way for learning languages.

(6) 単数名詞に対して不定冠詞を用いる。

I attended **a** German class last year.

(7) 修飾語句を伴う表現に対して不定冠詞を用いる。

A woman said in **a** loud voice, "I love you!" on the street.

(8) 修飾語句を伴う表現に対して定冠詞を用いる。

I do not know **the** meaning of the class.

この 8 種類の用法を、冠詞を含む (名詞) 表現の特徴、例えば、数量、照応など、により分類し、その特徴を節点とする決定木 (図 1) を作成した。例えば、節点 4 における定冠詞の誤用は、節点 1 から 4 のいずれかにおいて学習者が誤った選択を行ったため生じる。つまり、各節点を正しく判定する言語知識が未習得と判定できる (冠詞の使用漏れ以外の誤用も考慮し、無冠詞 (zero) の用法も加えた)。このように未習得の言語知識の判定だけでなく、決定木にもとづき誤用を評価することでタグ付け作業において評価者間の揺れを抑える効果も期待される。

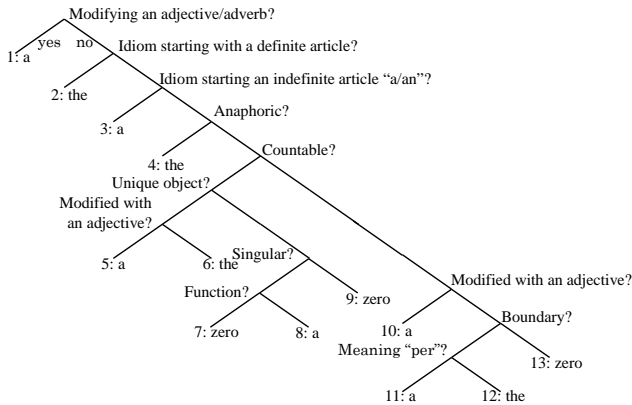


図 1 決定木化した冠詞用法

### 3.3 冠詞原因タグの付与

3.2 節の冠詞用法にもとづき誤用原因タグを作成した。学習者コーパス [7] における誤用タグの表記法 [11] に従って誤用原因タグを付与した。例文: A woman said in <at:4 cr="a"></at> loud voice, "I love you!"において、開始タグと終了タグ間に誤用に関する情報であるが、冠詞の使用漏れのため何ら情報は提示されていない。誤用の品詞 at (article の略)、正例 (不定冠詞は a、定冠詞は the) cr="a" が示されている。誤用原因タグを開始タグ中の品詞情報に後続する番号 (決定木上の節点) として付与した。

### 3.4 誤用原因タグの分布

表 1 に学習者コーパス [7] における誤用原因タグの分布を示す。誤用原因提示ツール開発用の学習者コーパスには全ての誤用原因タグが一定量以上含まれるべきであるが、実際には頻度の差が大きかった。

誤用原因タグ 4 (照応表現への定冠詞の用法) と誤用原因タグ 8 (非照応、単数表現への不定冠詞の用法) は高頻度であったが、誤用原因タグ 3 (慣用的な不定冠詞の用法) と誤用原因タグ 5 (修飾表現を伴う唯一事物に対する不定冠詞の用法) は皆無であった。

低頻度の誤用原因タグを増加させるために、作文課題の検討が必要である。対策としては、ある冠詞の用法の使用

を明示的に条件化する対策や、絵や質問を変更あるいは追加を通じてある冠詞の用法の増加を促進するといった対策が挙げられる。

誤用原因タグ	冠詞	頻度
1	不定冠詞	1
2	定冠詞	24
3	不定冠詞	0
4	定冠詞	160
5	不定冠詞	0
6	定冠詞	8
7	---	---
8	不定冠詞	235
9	---	---
10	不定冠詞	7
11	不定冠詞	3
12	定冠詞	1
13	---	---
合計		439

表 1 誤用原因タグの分布

## 4. おわりに

英語学習者の作文支援として同種の誤用の再発防止の観点から、誤用の原因となった未習得の言語知識を提示する必要性を論じた。そして、誤用原因提示ツールの開発に先立ち、学習者コーパスの誤用タグとして未習得の言語知識を提示する誤用原因タグを提案した。また、誤用原因タグを既存の学習者コーパスに付与し、その分布を調査した結果を報告した。その結果、分布は一様でなくデータの追加の必要性が判明した。

謝辞

本研究の一部は JSPS 科研費 22300299、15H02940 の助成を受けたものである。

### 参考文献

- [1] 永田亮, “文法誤り訂正における問題点の考察と新タスクの提案”, 言語処理学会第 22 回年次大会発表論文集, pp.681--684, (2016).
- [2] 梅澤次郎, 水野淳太, 岡崎直観, 乾健太郎, “冠詞誤り訂正時における訂正根拠の提示”, 言語処理学会第 18 回年次大会発表論文集, pp.353--356, (2012).
- [3] Ellis, R., *The Study of Second Language Acquisition* (2nd. Ed.), Oxford: Oxford University Press, (1996).
- [4] Taylor, G. “Errors and explanations”, *Applied Linguistics* 7, pp.144--166, (1986).
- [5] Richards, J. “A non-contrastive approach to error analysis”, *ELT Journal* 25, pp.204--219, (1971).
- [6] Knight, K. and Chander, I., “Automated postediting of documents”, *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pp. 779--784, (1994).
- [7] Kotani, K., Yoshimi, T., “Analysis of the association between reading proficiency and frequency of writing errors using an error-tagged learner corpus for multiple linguistic skills”, *ICIC Express Letters: An International Journal of Research and Surveys*, pp.569--573, (2016).
- [8] Thomson, A.J., Martinet, A.V. *A Practical English Grammar* (3rd. Ed.), Oxford: Oxford University Press, (1985).
- [9] Quirk, R. Greenbaum, S., Leech, G., Svartvik, J., *A Comprehensive Grammar of the English Language*, London: Longman, (1985).
- [10] 石黒昭博 (監), 総合英語 Forest (第 6 版), 東京: 桐原書店.
- [11] 和泉絵美, 井佐原均, 内元清貴, 日本人 1200 人の英語スピーキングコーパス, アルク (2004).