

平成 30 年 6 月 12 日現在

機関番号：17102

研究種目：若手研究(B)

研究期間：2015～2017

課題番号：15K16798

研究課題名(和文)CEFRレベルと意味内容の対応付け：フレーム意味論の観点から

研究課題名(英文)Assigning CEFR levels to semantic frames: A frame semantics approach

研究代表者

内田 諭 (Uchida, Satoru)

九州大学・言語文化研究院・准教授

研究者番号：20589254

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：本研究はCEFRレベルと意味フレームの対応付けを目的とし、CEFR-J Wordlist、FrameNet、教材コーパス等を用いて研究を実施した。その結果、下位レベルでは日常生活に關与するフレームが観察され、レベルが上がるにつれて社会に關わるフレームや副詞的表現に關わるフレームなど、意味内容的により複雑と思われるフレームが特徴的に出現することが明らかとなった。また、コアフレーム要素の平均値等から、フレームの構造がレベルの上昇とともに複雑になることも観察された。これらの結果は、CEFRレベルごとの特徴フレームのリストとしてウェブページで公開されている。

研究成果の概要(英文)：The purpose of this study was to make a list of semantic frames for each CEFR levels. Using CEFR-J Wordlist, FrameNet and a coursebook corpus, it was shown that frames which are related to everyday life distinctively appear at lower CEFR levels, while more semantically complex frames that are related to social activities and adverbial expressions are observed at higher levels. It was also shown that the structures of frames become more complex with the increase in the CEFR level, particularly in terms of the average number of core frame elements. The list of feature frames are public on a website.

研究分野：英語学

キーワード：CEFR 意味フレーム FrameNet フレーム意味論

1. 研究開始当初の背景

ヨーロッパの複言語主義・複文化主義に根ざす CEFR (Common European Framework of Reference for Languages) は、「言語を使って何ができるか」という CAN-DO リストの形で規定され、国際的な言語学習の評価指標として広く用いられるようになってきている。日本でも文部科学省から「各中・高等学校の外国語教育における『CAN-DO リスト』の形で学習到達目標設定のための手引き」が平成 25 年 3 月に公表されたり、日本の英語教育での利用を目的に CEFR-J (cf. 投野(編)2013) が開発されるなど、その利用が活発になってきている。CEFR を使った評価のメリットは、「目標の具体化」、「(生徒・教師間での) 目標の共有」、「4 技能の統合」などが挙げられ、その重要性は今後さらに増すと考えられる。

この指標を現場レベルで活用するための 1 つの方法は、旧来のシラバスとの対応付けである。English Profile (<http://www.englishprofile.org/>) のプロジェクトや基盤研究 A (課題番号: 24242017[研究分担者]) では、文法形式と CEFR レベルの対応付けの研究が進められ、例えば「現在完了形は B レベルから顕著に見られる」といった形で、各レベルを特徴づける言語特徴の抽出が行われている。しかしながら、言語が表す意味内容と CEFR レベルの関係については、研究が進められていない。English Profile や CEFR-J の研究 (cf. 投野(編)2013) では、レベルごとの語彙リストは提供されているが、それらの語彙が表す意味内容を分析し、意味内容に関する言語特徴を抽出する試みは未着手である。

2. 研究の目的

本研究の目的は、CEFR レベルと意味フレームの対応付けを行うことである。フレーム意味論に基づいた分析を行うことで、(1) 意味をフレームという形で具体化できる、(2) フレームを構成する要素の観点から意味を構造化できる、(3) フレーム間関係により意味を体系化できる、というメリットが得られる。それぞれの CEFR レベルに特徴的な「フレーム」を明らかにすることができれば、レベルに応じた効果的な教育を可能にするものであり、英語教育に貢献できるものである。

3. 研究の方法

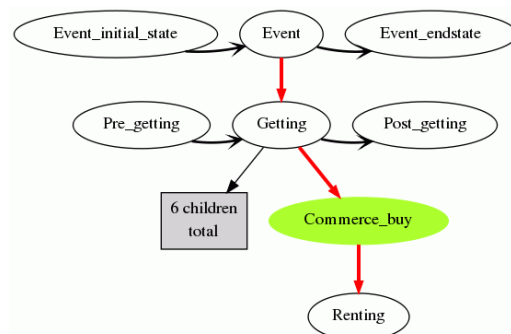
CEFR レベル別の特徴フレームを抽出するにあたり、ベースリストとして CEFR-J Wordlist (東京外国語大学投野由紀夫研究室作成) を利用した。CEFR-J Wordlist は日本の英語教育の文脈で使用することを目的に開発された語彙表で、一般英語のコーパスだけではなく、教科書のコーパスなど利用して作成されたものである。A1、A2、B1、B2 の 4 つのレベルに分類され、7814 語が含まれている。また、教材でのレベル感を確認するため、海外の出版社から出版されている

CEFR ベースの教材を中心とした教材コーパスを検証用に作成した。

一方、意味フレームの情報を抽出するため、FrameNet (<https://framenet.icsi.berkeley.edu/fndrupal/>) の情報を参照し、それぞれの対応表を人手で作成し、一貫性等のチェックを行った。FrameNet はフレーム意味論 (Fillmore 1982, 1985 など) を基盤として開発が進められているオンラインの「フレームの辞書」である (詳細は Ruppenhofer et al. 2016 を参照)。2018 年 3 月現在、約 13,000 の単語の語義に対して「どのフレームが喚起されるか」という情報が示されている。例えば、動詞の buy は Commerce_buy フレーム、sell は Commerce_sell フレームをそれぞれ喚起すると規定されている。フレームとは単語の理解の基礎となる背景知識を指し、前述の buy では参与者 (buyer) がお金 (money) と引き換えに売り手 (seller) から商品 (goods) を得る、という一連の商取引に関わる知識がフレームであり、その知識なくして単語の意味理解は達成されない。フレームには意味的に必須なコアフレーム要素とそうではない要素があり、フレームを定義する上ではコアフレーム要素が重要な役割を果たす。

一つのフレームは複数の単語によって喚起される場合がある (buy, purchase はどちらも Commerce_buy フレームを喚起する)。一方、一つの単語は複数のフレームを喚起する場合がある (cook(v) は Apply_heat フレーム、Cooking_creation フレームなどを喚起する)。

FrameNet を利用することで、CEFR-J Wordlist の単語をフレームごとにまとめることができる。さらに、FrameNet ではフレーム間関係によってフレーム同士の関係が記述されている。例えば、Commerce_buy は上位に Inheritance の関係で Getting フレームを持つ (下図参照)。この情報を利用することで、さらに一般的な単位で単語を集約することが可能となる。本研究では CEFR-J Wordlist にフレームを付与し、さらにその上位フレームを特定し、集計表に対して多変量解析の手法の一つである対応分析を実施することで一般的なレベルでの意味特徴を明らかにした。



4. 研究成果

CEFR-J の 7814 語に対して、FrameNet をベースに意味フレームを付与できたのは 3843 件 (49.2%) であった。フレームの種類は 761 で、平均して 5 語ごとに意味フレームをベースにグルーピングできたことになる。さらに、フレームが付与できた単語についてもっとも厳密な上下関係である Inheritance の関係を用いて上位フレームを付与した (2429 件に対して一般化できた)。

各レベルに見られる意味特徴を抽出するため、本研究では CEFR-J Wordlist において「出現頻度 5 以上のフレームでいずれかのレベルで占有率が 50% を超えるもの」を特徴フレームとして定義した。特徴フレームは 111 件抽出され、その内訳は A1:16 件、A2:16 件、B1:36 件、B2:43 件となった。以下、部分的にはあるが具体的な内容を内田 (2017) に基づいて報告する。

A1 レベルの特徴フレームで最も頻度が高いものは Food フレームであった。このフレームには banana, bread, cake, chicken, coffee などの食品に関する一連の語が含まれている。次に頻度が高かったものは Calendric_unit で、January, February, morning, night などが含まれる。さらに brother, child, dad, family などの語が喚起する Kinship フレームが続く。以上から、A1 レベルは食事、時間、家族など日常生活に関するフレームが特徴的であるということができらる。

A2 レベルで最も頻度が高い特徴フレームは Direction で east, left, forward, north, up などの語がこのフレームを喚起する。また、Organization フレームも A2 レベルの特徴フレームの一つで、agency, committee, government などが含まれる。さらに Posture フレーム (lie, position, stand など)、Suitability フレーム (appropriate, proper, suitable など) などこのレベルの特徴フレームである。A1 レベルと比較すると、A2 レベルの特徴フレームは「方向」、「組織」など英語使用者の外部にあるものと関連していることがわかる。また、「判断」を表す単語も多くなり、英語を通してより高度な知覚活動を行うことになることがわかる。

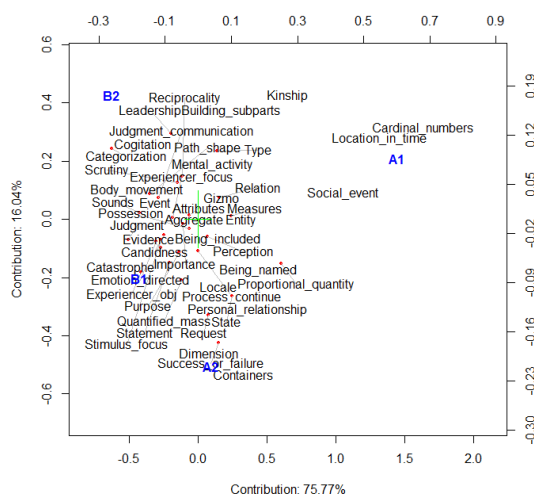
B1 では Frequency フレームが最も頻度が高かった。このフレームを喚起する語には constantly, frequently, normally, occasionally などの副詞が多く含まれ、副詞表現を使ったより詳細な事態の描写を行うための語彙がこのレベルの特徴だとわかる。また、Change_position_on_a_scale フレームが次に頻度の多いフレームである。このフレーム (decrease, fall, increase, rise などが喚起語) は複雑な変化の事態を表す。コア要素として変化の割合 (例: by 8.4%) や変化の始点 (例: from 20 million) 変化の終点 (to 168 million) などを含み、統語的にも複雑なものとなることが多い。この傾向は、レベル

ごとのフレームのコアフレーム要素の平均を見ても明らかである (下表)。特に A2 と B1 レベルの間には大きな隔りがあることが読み取れ、A レベルと B レベルでは意味フレームの構造の複雑性が大きく異なることが示唆される。

Level	コアFEの平均
A1	2.71
A2	2.79
B1	3.06
B2	3.22

B2 レベルでは Cause_harm フレームが最頻であった。chop, crack, slap, smash などの文脈が限定される多くの単語と関連する。また、Expertise フレーム (competence, expert, mastery など)、document フレーム (certificate, contract, testimony など)、Categorization (categorization, classification, interpretation など) など専門性の高い文書などで使われる意味内容が特徴的に見られる。

以上の傾向は、上位フレームでの分析においても同様に観察された。下図は各フレームを Inheritance の関係で一般化してレベルごとに上位フレームの出現を集計した表に対して対応分析を行った結果である。



対応分析の結果では関連する項目は 2 次元の平面上で近接する。例えば、A1 は Cardinal_numbers、Location_in_time フレームと近接していることから、これらのフレームと関連性が高いといえ、それぞれ「数字」および「時間」を表すことから、前述の観察と同様の傾向であることが読み取れる。

以上の結果は、各レベルの特徴フレームのリストとしてウェブページ (後述) で公開している。このリストは、報告者が知る限り CEFR レベルと意味内容の対応付けを行ったはじめてのリストであり、CEFR レベル別に教材のテーマの選定 (例えば A1 レベルでは「食事」に関するテーマを扱う、等) や第二言語習得の意味的に最適な順序を明らかにするなどの点で有益なものであるといえる。今後、本リストを用いて高校英語教科書等の教材の分析を進めたいと考えている。

[参照文献]

Fillmore, C. J. (1982) Frame semantics. In Yang, I. (ed.), *Linguistics in the morning calm: Selected papers from SICOL-1981*. pp.111-137. Seoul: Hanshin.

Fillmore, C. J. (1985) Frames and the semantics of understanding. *Quaderni di Semantica*, 6 (2). pp.222-254.

Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson and J. Scheffczyk (2016) *FrameNet II: Extended Theory and Practice*. <https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf>

投野由紀夫(編)(2013)『CAN DO リスト作成・活用：英語到達度指標 CEFR J ガイドブック』大修館書店。

内田 諭(2017)「英単語の意味内容と CEFR レベル—フレーム意味論に基づいた学習用語彙リストの分析」『言語科学』52, 75-85.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 2 件)

1. 内田 諭(2018)「word2vec による類義語抽出と FrameNet の比較: 言語研究のための質的検証」『言語統計を用いた認知言語学研究へのアプローチ (統計数理研究所レポート)』41-51.

2. 内田 諭(2017)「英単語の意味内容と CEFR レベル—フレーム意味論に基づいた学習用語彙リストの分析」『言語科学』52, 75-85.

[学会発表] (計 5 件)

1. 畔元里沙子・内田 諭(2018)「高校英語教科書の CEFR レベル: CEFR-J Wordlist に基づいた語彙の数量的分析」言語処理学会第 24 回年次大会 (岡山コンベンションセンター)

2. S. Uchida (2017) “Semantic Frames in the CEFR-J Wordlist: A Frame Semantic Approach to the Lexical Level”, JACET (青山学院大学)

3. 内田 諭(2017)「コロケーションスコアに基づいた英単語の変換手法」, メソドロロジー研究部会 (西南学院大学)

4. 内田 諭・高田祥平・水嶋海都・荒瀬由紀 (2016)「CEFR レベルに基づいた英単語の変換: 英文難易度の最適化を目指して」, 英語コーパス学会 (成城大学)

5. 内田 諭(2015)「連続するイベントに関する辞書記述について: フレーム意味論の観点から」, ワークショップ『語用論と言語教育』(九州大学伊都キャンパス)

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称 :
発明者 :
権利者 :
種類 :
番号 :
出願年月日 :
国内外の別 :

○取得状況 (計 0 件)

名称 :
発明者 :
権利者 :
種類 :
番号 :
取得年月日 :
国内外の別 :

[その他]

ホームページ等

<http://fllc.kyushu-u.ac.jp/~uchida/research/index.php/study/>

6. 研究組織

(1) 研究代表者

内田 諭 (UCHIDA, Satoru)

九州大学・大学院言語文化研究院・准教授
研究者番号 : 20589254

(2) 研究分担者 なし

()

研究者番号 :

(3) 連携研究者なし

()

研究者番号 :

(4) 研究協力者 なし

()