慶應義塾大学学術情報リポジトリ
Keio Associated Repository of Academic resouces

Title	自然言語処理を用いた学びの振り返りについて
Sub Title	Self-assessment of foreign language learning with the help of natural language processing
Author	矢田部, 清美(Yatabe, Kiyomi)
Publisher	慶應義塾大学外国語教育研究センター
Publication year	2022
Jtitle	慶應義塾外国語教育研究 (Journal of foreign language education). Vol.18, (2021.) ,p.103- 112
JaLC DOI	
Abstract	
Notes	研究ノート
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AA12043414-20210000-0103

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

自然言語処理を用いた学びの振り返りについて

矢田部　清　美

1　はじめに

本稿では、外国語学習法研究における自然言語処理の活用について論ずる。まず、自然言語処理の概略と外国語学習への応用を述べる。さらに、重要性を増す外国語学習における自己評価において、自由記述文による自己評価を含んだ総合的評価の必要性を述べ、自由記述文の自然言語処理例を検討する。最後に事例を踏まえ、外国語学習における自然言語処理応用の長短と今後の展望を述べる。

2　外国語学習と自然言語処理技術

自然言語処理とは、人間が日常的に使用している自然な言語（例えば日本語や英語といった言語）を（計算機などの利用のもとで）自動的に処理したり処理した結果を応用することを目標とした技術分野である。自然言語処理の基本的な処理技術には、形態素解析や構文解析、それらを応用したテキスト処理技術が存在する。形態素とは自然言語において意味を持つ最小の言語単位を指す。形態素解析器は文を形態素に分割し、機械可読辞書を用いて意味のある情報を取得する。分割結果には各単語の品詞や語形変化が付与されている場合もある。さらに構文解析器が形態素解析で得られた分割された単語間の関係性について解析を行う。これにより各単語の自然言語文内の機能、例えば文の主語などが特定され、構文が理解される。これらの結果は自然言語文章の意味解析や文を超えた文脈解析に用いられることもある。

外国語学習においては、大きく三つの対象に対して自然言語処理の応用、もしくは今後の応用が期待されると考えられる。

まず一つ目は、学習者の学習目標言語、つまり母語話者の使用する言語を対象としたものである。母語話者によって実際に使用されている（authentic な）言語を集め、その用例を用いて言語活動に利用することが増えている。そこで用例の収集、表示、学習用の問いの生成に自然言語処理技術が使用されることがある。例えば、用例中の自然言語を解析し、単語や単語列の例を表示するデータ集合に、コンコーダンスというものがある。コンコーダンスは、自然言語文章中のある特定の文字列（単語）について、その単語が含まれる文全体や前後の単語の位

置情報を備えたものである。コンコーダンスを備えたコンコーダンサーの検索欄に、学習者が用例を知りたい単語を入力すると、入力単語を中心としてその単語が含まれる文全体や前後の単語が表示される。また、自然言語処理を用いて半自動的に収集され注釈を付与され可視化された外国語のデータも目標言語の学習用教材として利用できる可能性がある。例えば、単語の関係性には継承関係があることを利用した電子辞書 WordNet[1]がある。こうした自然言語処理を用いた目標言語の学習に役立つ教材用のツールの開発は古くからおこなわれ最適化されており、使いやすいユーザーインターフェースを備えたものが多い。教育者が自然言語処理ツールを用いて収集した用例を手動でクイズ作成に利用することもある。例えばクイズ作成にはラーニングマネージメントシステムとしての Moodle や Google Classroom などでのクイズ作成機能の他、クイズ作成可能なフリーウェアの Hot Potatoes[2]や、一部商用であるが多種のクイズが作成できる Quiz Generator[3]などのアプリケーションも提供されている。

二つ目は、学習者の使用する言語を対象としたものである。学習者がの語彙や文法を習う際に生成した誤り含みの外国語に対して自然言語処理が利用されることが考えられる。学習者の母語の特性や習熟段階に応じて学習者が発する単語や文は異なるという仮定の下、学習者モデルが解析結果を用いて更新されながら処理される。さらにモデルに従った適切なフィードバックも試みられている。これらの解析処理には学習者の使用する言語文を集めた学習者コーパスの研究も役立っている。例えば International Corpus of Learner English（ICLE）は異なる母語を持つ中上級学習者クラスの大学生などによる英作文550万単語を納めた学習者コーパスである（Granger, 1998）。前述したコンコーダンス機能を用い、学習者の母語や年齢、性別などを選択し、学習者の使用する言語を分析することもできる。図 1 は日本語を母語とする学習者の英作文の中から very という単語を用いた文を very の単語位置で揃えて ICLE で表示した結果である。用例を俯瞰することで very の誤用にどのような特徴があるかを簡便に把握する点が有用であると思われる。

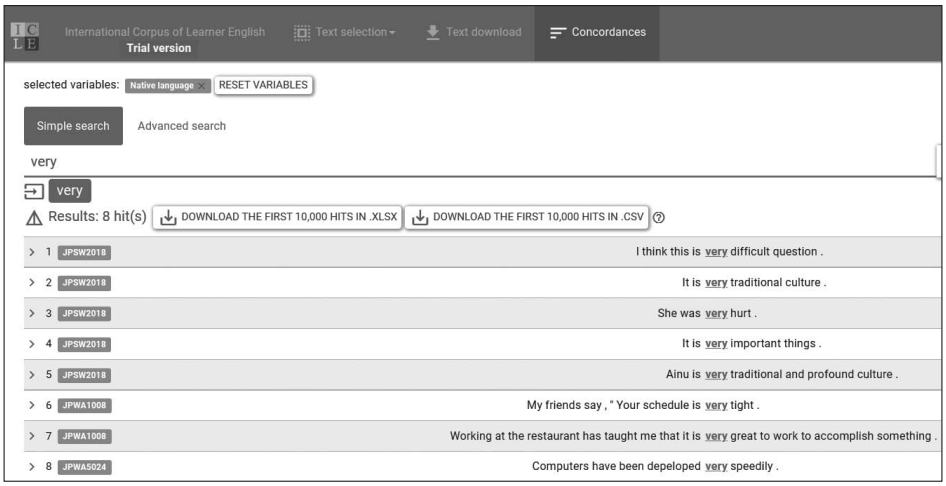

図１　コンコーダンスを用いた学習者コーパスの例

自然言語処理の外国語学習応用として、三つ目の対象は、学習者が学習を振り返り自己の言葉で述べた文章になるのではないかと考えられる。学習者や教育者は各学習段階で、学びや教えを振り返り評価を行う。多くの学習者の振り返りを自由に文章で記述したデータを自然言語処理を用いて分析することは、それ以降の言語学習活動の設計と学習者の発達に関しての洞察を与えてくれると考える。三つ目の対象に対する自然言語処理については次章で詳述することとする。

３　自然言語処理を用いた学びの振り返り

３.１　外国語学習における総合的自己評価における自由記述の重要性

筆者らはこれまでの研究実践事例で外国語学習における自己評価は自由記述の方式を含む複数の評価からなる総合的評価を行うことが望ましいことを報告している（吉村・矢田部・境, 2019, 2020）。

この報告の中で、履修二年目にある第二外国語学習者は約三カ月間のパフォーマンス課題の学習中、自身の能力を振り返る自己評価を行った。ここで、パフォーマンス課題とは、実際に存在するもしくはそれを模した文脈において、知識やスキルを総合して使いこなすような課題を指す（西岡・田中，2009）。実際のパフォーマンス課題では学習者らは母語話者に外国語を用いて学校内の施設を案内した。その自己評価においては、ルーブリック、CEFR の共通参照レベル（Council of Europe, 2001）、自由記述といった三種類の異なる評価をあわせて総合的評価をした。ルーブリックとは学習者の達成度を判断する評価要素について数段階のレベルの目安を記述した評価基準表を指す（當作・中野，2013）。実際のルーブリック評価では教育

者がパフォーマンス課題における複数の評価要素について「目標以上達成」「目標達成」「もう少し頑張ろう」「努力が必要」の 4 つのレベルを設け、評価要素の各レベルに該当する目安を記述文として作成した。CEFR の共通参照レベルの自己評価では、読む、話すなどの外国語能力について数段階のレベルの目安として「自分は〜できる」といった能力記述文があり、学習者は自身の能力のレベルを選択した。自由記述における自己評価では、学習者が自分ができる（できた）あるいはできない（できなかった）と思うことやそれに関する感想などを自由に記述した。CEFR の共通参照レベルと自由記述の両方で、外国語学習開始時と課題実施後に自己評価を行った。その結果、学習者は CEFR の共通参照レベルにおける自身の能力評価について開始時より実施後により低い評価を行う傾向があった。一方、自由記述における自身の能力評価では学習開始時は漠然とした記述であった回答が、課題の実施後はより明確な分析や学習方略を意識した記述を伴った回答となり、記述量も増したことがわかった。

この研究実践事例を通じ、学習者の自己評価は、予め定義されたレベルを選択するという方式だけでなく、自由に記述するという方式も加える、つまり総合的評価を行うことで、学習者が学びの振り返りを深めることが可能になるということが示唆された。選択式の学習者の CEFR の共通参照レベルの自己評価が外国語学習開始時と課題実施後に低くなったのは、現実に起こりうる文脈で知識やスキルを総合して使いこなすようなパフォーマンス課題に取り組んだことで学習者がより客観的に自己の能力をみつめるようになったのではないかと推察された。それは自由記述文による自己評価の変化をみることでより確信性が高まった。

このように自己評価ではレベルを選択する評価形式だけでなく、自由記述も含めることで、方略などの学習者の言語発達に洞察を与える機会がもたらされると考えられる。学習能力は、どのように学ぶかを学ぶという方略、方法知を豊かにする上で重要である（Halász & Michel, 2011）。例えば、CEFR でも学習者個人にかかわる一般的な能力の一種として学習能力の名で学習能力は言語学習の動機を高め、学習方略を豊かにするものとして紹介されている。外国語学習では、このような学習能力が学習者の学習の基礎に育つことで、言語間、つまり新しい言語のより容易な学習、使用が可能になると考えられている（de Sousa, Greenop & Fry, 2010; Kuo & Anderson, 2012）。そのため学習能力の発達を促すような貴重な自由記述文の内容は学習者の声として分析・活用することが必要である。

自由記述文を自己評価に加えるということは学びの振り返りを深めるという利点がある一方で、自由記述による評価文章を読み分析することには時間や労力、読み取るための分析力が必要であると考えられる。先の研究実践事例では研究者が各学習者の自己評価文を読み、学習者にどのような傾向があるのかといった考察を行ったが、自由記述文の回数や人数が多くなると負担が多くなることが予想された。そこでこうした学習者や教育者が学習の振り返り自分の言葉で述べた自由記述文の分析支援に自然言語処理技術が可能なのではないだろうかと考えた。

3.2　自由記述による評価文章の自然言語処理による分析例

本節では、自己や複数の人の振り返りである自由記述の分析を簡便な自然言語処理で支援することが可能かどうか四つの例を用いて検討する。

まず記述量についてである。これは目視でもある程度可能であり、ごく簡単に自由記述文の文字数をカウントするツールを用いることが可能である。Microsoft Office Word などの文書作成ソフトウェアにおいてもその機能は実装されている。例えば「文字数をカウントする」といった文は10文字の文字数である。もし文字数ではなく単語のような意味や形式の単位で記述量をカウントする場合には、第一章で述べたような機械可読辞書と形態素解析技術を用いる必要がある。形態素解析器や機械可読辞書は自由記述文中の言語用のものを選択する。例えば日本語の自由記述文においては日本語用の形態素解析器や辞書が必要である。日本語用の形態素解析器には、例えば Juman、Chasen、Janome、MeCab といったものがある。インターネット上で文章を入力するとこれらの形態素解析器を用いた分析結果を表示するサイトもある。例えば日本語教育の研究推進をおこなう「やさしい日本語」科研グループのやさしい日本語のサイト[4]がある（庵 他，2014）。このサイトでは文章を入力すると解析器 MeCab と辞書 UniDic による形態素解析結果（ここでは形態素ではなく辞書に登録されている単位の語彙素を使用）を表示する。図２にこのサイトを利用した「単語数をカウントする」といった文の解析結果を例として示す。この解析結果では文が「単語 / 数 / を / カウント / する」のように「/」の位置で５つの単位に分割され、単位ごとに発音や読み、品詞や活用形が表示されている。このような自然言語処理の基礎となる解析により、それをもとにした文章の解析が可能になる。そのため、文章の解析をする際にはどの形態素解析を用いてどのような単位の分割になっているのかに注意する必要がある。

書字形	発音形	語彙素読み	語彙素	品詞	活用型	活用形	語形	書字形基本形	語種
単語	タンゴ	タンゴ	単語	名詞-普通名詞-一般			タンゴ	単語	漢
数	スー	スウ	数	名詞-普通名詞-一般			スー	数	漢
を	オ	ヲ	を	助詞-格助詞			オ	を	和
カウント	カウント	カウント	カウント-count	名詞-普通名詞-サ変可能			カウント	カウント	外
する	スル	スル	為る	動詞-非自立可能	サ行変格	終止形-一般	スル	する	和

図２　日本語形態素解析結果の例

記述量だけではなく、出現する単語の頻度によって、自由記述文を分析することも可能である。例えばワードクラウドがその一例である。ワードクラウドとは文章中に多く出現する単語などを、出現頻度の高さに応じて文字の大きさを変更して表示する手法をいう。形態素解析結果を用いて自分でワードクラウドを作成することも可能であるが、インターネット上で文章を入力するとこれらの形態素解析器を用いたワードクラウドの分析結果を表示するサイトもある。図３はワードクラウド[5]の結果である。これは筆者が所属する機関のあるシンポジウムに参加した複数の教育関係者がシンポジウムの内容を評価し自由に記述した文章を入力して得られた。今回はあらかじめ別の形態素解析結果を用い前出の図２にみられるような形態素解析上の格助詞「を」やサ変動詞の語尾「する」などを除くように事前に編集した。これらは機能語と呼ばれるもので、名詞・形容詞・動詞・副詞のように実質的な内容を表す内容語に比較し、出現頻度が高いが意味的な情報が少ない。そこで、そのような意味的な情報が少ない語がワードクラウドで大きな文字として表示されないように予め除去した。ワードクラウドは結果の図を一枚みるだけで自由記述文で主にどのような単語が多くあらわれたのかということを簡単に理解することを支援する点で有用であるといえる。一方で、欠点としては、文書中に多くの単語がある場合、頻度が少ない単語は文字サイズが極端に小さく読みづらくなる、あるいは表示されないということも起こる可能性がある。また、今回おこなったように情報価値が低いと思われる機能語をあらかじめ除くなどの編集が必要な場合が考えられる。

図３　出現頻度に応じた文字の大きさで表示された自由記述文中の単語例

これまでは自由記述文中の文字列の数や出現頻度を計算したのみの結果であった。こうした出現する度数のみの結果でなく、その文字列が文書中でどのような特徴を持つのかについてさらに解析を行う必要がある。

文書内のある特定の単語の重要性を測る指標に Term Frequency − Inverse Document Frequency（TF−IDF）がある。TF−IDF はある文書 d 内のある特定の単語 i の重要性を、複数の別の文書内のその単語と他の単語の出現頻度に基づいて相対的に評価する指標である。TF−IDF で重要度が高いとされる単語は文書内でその単語が多く出現し、かつその単語は他の文書ではほとんど出現しないように計算される。TF−IDF は Term Frequency（TF）と Inverse Document Frequency（IDF）の積により求められる。TF は語の出現頻度である。単語 i の文書 d における出現回数を文書 d における全単語の出現回数の和で分割して求められる。IDF は逆文書頻度である。単語 i が複数の別の文書で共通して出現するかもしくはある文書 d でのみ出現するか、つまりを単語 i の希少性（Specificity）を示す指標である。IDF は単語 i が出現する総文書数の逆数によって計算される。その対数をとることで TF の重み付けに利用される。この重み付けを使用することで、ほとんどの文書で出現する機能語や一人称の私などの単語の重要度を抑制されるようになっている。最終的には TF と IDF の積で計算されるため IDF が 0 になると TF によらず値が 0 になってしまうので IDF には 1 が加算されることが多い。TF−IDF は一般的な文書要約（テキストマイニング）にも使用される指標であり、他の人の自由記述文と比較し、ある自由記述文内で特徴のある単語の選択に有用である。例えば、複数の人が書いた自由記述文「多読をした」「リスニングをした」「コメントをした」があった場合、「多読をした」では「多読」「を」「した」の出現頻度が各一回で同じであるが、「を」「した」は他の自由記述文「リスニングをした」「コメントをした」にも現れている一方で「多読」は他の自由記述文に現れていないため、特徴のある単語と計算された。TF−IDF を簡便に使用できるアプリケーションは見つからなかったが、python といったプログラム言語で利用できる文書集合を TF−IDF 特徴量の行列に変換するインターフェース[6]が提供されている。

TF−IDF は文書内の特徴的な単語をあらわす指標であったが、別の方法で意味的に似通った単語を塊（クラスター）として表示することもできればよいかもしれない。例えば意味的に類似する単語を表示することができる技術に word2vec がある（Mikolov, Sutskever, Chen, Corrado, Dean, 2013）。word2vec は「ある単語の意味はその単語の周囲の単語によって規定される」という分布仮説に基づいた自然言語処理手法の一つである。この手法では、文書内の各単語はある一定の長さのベクトルに配置され、近くに同じ単語が現れる単語同士は類似したベクトルに対応するよう、コーパス文章中の近接単語の出現確率を用いて一定の次元数のニューラルネットワークによりベクトルの値が学習される。word2vec も簡便に使用できるアプリケーションは見つからなかったが、python プログラム言語で利用できる学習済みデータ

や類似単語ベクトルを検索するインターフェース[7]は提供されている。今回使用した学習済みデータは「日本語 Wikipedia エンティティベクトル」（鈴木・松田・関根・岡崎・乾，2016）で、日本語版 Wikipedia の本文全文を MeCab を用いて形態素解析した結果から word2vec を用いて学習した200次元の分散表現ベクトルのデータである。図４はワードクラウド用に利用した各単語のベクトルの値を学習済みデータから検索し、２次元にマッピングした結果である。すべての単語が表示されているために図の中央部分の単語群が重なり読みにくくなっている。そこで図中の（4）の拡大図を図５とした。これらをみると、自由記述文中で意味的に似通った単語は近距離に配置されている場合が多いことがわかる。例えば図中の（1）は研究や調査に関わる意味的単語のクラスター、（2）は思考や感覚といった認知活動にかかわる意味的単語のクラスター、（3）は評価軸にかかわる意味的単語のクラスター、（4）は言語活動にかかわる意味的単語のクラスターといった具合である。複数の人の自由記述にどのような傾向があるのかを理解し易くなると考えられる。ただし、このままではそうしたクラスターのつながりやどのクラスターの単語が文書内で出現したのか等の情報が得られず理解しづらいため、さらにワードクラウドで用いた頻度情報等を組みあわせて再表示する必要があると思われる。また機械が学習した結果であるので、人間が意味的に集合とみとめる内容と異なる結果も含まれている。これについては今後検討が必要と思われる。

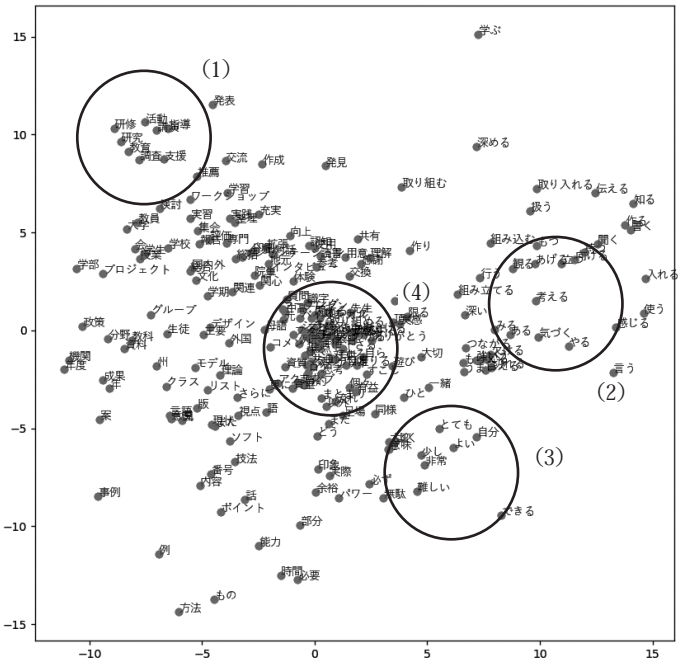

図４　意味的な類似度により配置された自由記述文の単語群の例

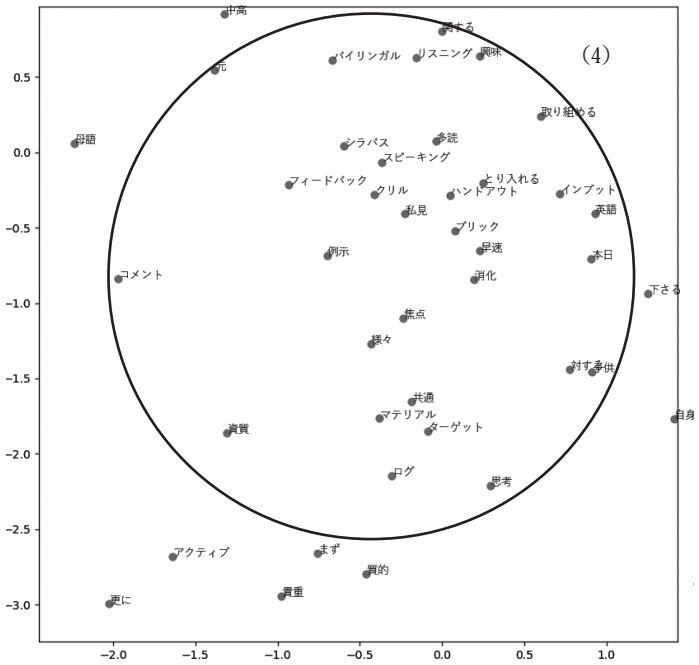

図５　図４の中央部分を一部拡大した図

４　まとめと展望

本稿では、外国語学習法研究における自然言語処理の活用について論じた。自然言語処理の応用分野の中で、外国語学習における自己評価の大切さについて着目し、自然言語処理でそれらの文章を対象とする提案をした。さらに自由記述文による自己評価を含んだ総合的評価の必要性を述べ、自由記述文の自然言語処理例を検討した。

その結果、自己や複数の人の自由記述をそれぞれ、もしくはあわせて自然言語処理技術によって処理することで、個人の自由記述の特徴や集合としての自由記述の特徴を分析できる可能性が示唆された。今後は学習者の時系列的発達や集合としての言語活動の傾向をさらに分析できるよう、処理結果を統合する試みが必要と考える。具体的には、外国語学習が進むにつれて自由記述の記述量やワードクラウドの内容が時系列でどのように変化しているのかを記録表示できるようにすること、さらにそれらの結果に TF−IDF や word2vec の技術と組み合わせることで個人と集合の関係を検討すること等である。

自然言語処理はこうした処理を人手による処理よりも大量かつ短時間のデータ処理で可能とするという利点がある一方で、今後の利用時に注意する点、検討が必要な点も明らかになった。それらは、自然言語処理で初期の処理となる形態素解析の結果を十分確認する必要があること、直感的にわかりやすく情報量の高い分析結果の表示をするために工夫が必要なこと、機

械的な学習結果をそのまま用いると人間の意味的解釈と異なるので検討が必要なこと、学習者や教育者といったユーザが簡便に使える自然言語処理ソフトウェアの開発が今後も必要なことである。今後はこうした点に留意して自然言語処理の利点を生かしつつ、外国語教育研究を発展させる必要がある。

引用文献

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge, U.K.: Cambridge University Press.

de Sousa, D. S., Greenop, K., & Fry, J. (2010). The effects of phonological awareness of Zulu-speaking children learning to spell in English: A study of cross-language transfer. *British Journal of Educational Psychology, 80*(4), 517−533.

Granger, S. (ed.) (1998). *Learner English on Computer*. Harlow: Longman.

Halász, G., & Michel, A. (2011). Key Competences in Europe: interpretation, policy formulation and implementation. *European Journal of Education, 46*(3), 289−306.

Kuo, L.-J., & Anderson, R. C. (2012). Effects of early bilingualism on learning phonological regularities in a new language. *Journal of Experimental Child Psychology, 111*(3), 455−467.

西岡加名恵・田中耕治（2009）『「活用する力」を育てる授業と評価　中学校』，学事出版.

Mikolov, T, Sutskever, I., Chen, K., Corrado, G.S., Dean, J.（2013）. Distributed Representations of Words and Phrases and their Compositionality. *NIPS*, 26:3111−3119.

鈴木正敏・松田耕史・関根聡・岡崎直観・乾健太郎（2016）Wikipedia 記事に対する拡張固有表現ラベルの多重付与．言語処理学会第22回年次大会（NLP2016），March 2016.

當作靖彦・中野佳代子（2013）『外国語学習のめやす』国際文化フォーラム.

吉村創・矢田部清美・境一三（2019）『パフォーマンス課題における自己評価の効果について―CEFR の共通参照レベル，ルーブリック，自由記述による総合的評価の提案』，言語教育エキスポ2019，p46−47.

吉村創・矢田部清美・境一三（2020）『高等学校ドイツ語授業におけるパフォーマンス課題の実践報告』，外国語教育研究，16，72−93.

注

1 https://wordnet.princeton.edu/

2 http://hotpot.uvic.ca/

3 https://quizgenerator.net/

4 http://www4414uj.sakura.ne.jp/Yasanichi1/unicheck/

5 http://wordcloudjp.com/

6 https://scikit learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html

7 https://radimrehurek.com/gensim/models/keyedvectors.html