

韓國日本語學會 第45回 國際學術大會

< 主 題 >

データ基盤の日本語学研究
-Culture Mining, AI, 日本語教育-

- ・ 日時：2022年 3月 19日(土)
- ・ 場所：徳成女子大學校
(ZOOMを利用したオンライン学術大会)
- ・ 主催：韓國日本語學會, 聖潔大學校文化フレームビッグデータ研究所
- ・ 主管：韓國日本語學會
- ・ 後援：徳成女子大學校, JAPAN FOUNDATION SEOUL ,
(株)時事Books, AIとクリエイティブ・ラーニング研究会



韓 國 日 本 語 學 會
The Japanese Language Association of Korea

韓國日本語學會

第45回 國際學術大會

■ 日時：2022年 3月 19日(土)

■ 場所：徳成女子大學校

(ZOOMを利用したオンライン学術大会)

<Program>

13:00-13:15	・開會式・總會 <司會：辛銀眞(仁川大)> 開會辭：盧姪鉉(徳成女子大 / 韓國日本語學會 會長) 祝 辭：加藤剛(日本国際交流基金ソウル文化センター 所長) 張根壽(祥明大 / 韓国日語教育学会 會長) 聖潔大學校文化フレームビッグデータ研究所MOU締結
13:15-14:00	・企画発表 1 李竣瑞(聖潔大學校文化フレームビッグデータ研究所 所長)
14:00-14:45	・企画発表 2 林尙諄(聖潔大)
14:45-14:55	・研究倫理講演 <司會：辛銀眞(仁川大)> 蔡盛植(高麗大)
14:55-15:15	休息
15:15-16:00	・企画発表 3 落合由治(淡江大)
16:00-16:45	・企画発表 4 檢校裕朗(極東大)
16:45-17:00	・閉會式 <司會：辛銀眞(仁川大)> 閉會辭：盧姪鉉(徳成女子大学)

企画発表

韓国日本語學會, 聖潔大學校文化フレームビッグデータ研究所,
AIとクリエイティブ・ラーニング研究會(AI-CL) 共同企画

時間	発表者	主題	発表言語	討論	司会
13:15-14:00	■企画発表 1 李俊瑞 (聖潔大學校文化フレームビッグデータ研究所所長)	컬처마이닝(Culture Mining)	韓	金嘯泳 (同徳女子大) 金恵娟 (聖潔大)	中村有里 (仁川大)
14:00-14:45	■企画発表 2 林尙諄 (聖潔大)	문화이미지프레임망 (Cultural Image Frame Network, CIFN) 소고(小考)	韓	尹榮珉 (延世大) 孫榮爽 (濟州大)	郭銀心 (京畿大)
14:45-14:55	■研究倫理講演		韓	蔡盛植 (高麗大)	辛銀眞 (仁川大)
14:55-15:15	休息				
15:15-16:00	■企画発表 3 落合由治 (淡江大 AI-CL顧問)	AIとテキストマイニング	日	曾秋桂 (淡江大) 尹楨勛 (釜山外大)	朴江訓 (全州大)
16:00-16:45	■企画発表 4 檢校裕朗 (極東大 AI-CL会長)	AIと外国語クリエイティブ・ラーニング	日	李偉煌 (靜宜大) 李敬淑 (亜洲大)	飯干和也 (祥明大)

[企画発表]

컬처마이닝(Culture Mining).....	1
李竣瑞 (聖潔大學校文化フレームビッグデータ研究所所長)	
문화이미지프레임망(Cultural Image FrameNetwork, CIFN) 소고(小考).....	5
林尙諄 (聖潔大)	
AIとテキストマイニング	9
落合由治 (淡江大, AI-CL顧問)	
AIと外国語クリエイティブ・ラーニング.....	17
検校裕朗 (極東大, AI-CL会長)	

컬처마이닝(Culture Mining)

이준서

(성결대 문화프레임빅데이터연구소 연구소장)

초연결, 초경계, 초융합된 다문화/이문화/다언어 사회에서 기하급수적으로 쌓여나가고 있는 빅데이터 속에서 보다 더 풍부한 부가가치(value)와 데이터 분석을 통한 지식정보(insight)를 찾아내는 방안으로 컬처마이닝(Culture Mining) 기법을 제안한다.

Culture Mining is the process of finding useful or interesting ‘cultural elements’ -life styles, patterns, fashions, trends, models, beliefs, rules, frames etc.- of a specific region or a generation from unstructured text, various image sets by comparing different languages and cultures.

컬처 마이닝이란 인간의 문화를 표현하고 있는 다양한 미디어(텍스트, 이미지, 동영상 등)를 분석하여 언어/지역/성별/세대별 문화를 구성하는 문화 요소와 그들간의 관계를 발견해가는 과정이다.

특허: 420200819840

컬처마이닝은 기능, 성능, 신뢰도 등 기본적으로 어떤 재화나 서비스의 상품 가격에 반영되는 경제적인 상품가치와는 별도로 소비자의 감성을 자극하여 감동과 공감을 얻어냄으로써 상품가치를 더욱 더 극대화시키는 ‘감성가치’라고 하는 부가가치를 창출할 수 있다.



그림 7 컬처마이닝이 창출하는 경제 가치

이준서(2021:71)_『R컬처마이닝』박문사

소비자의 니즈가 다원화되고 있고, 다문화, 글로벌화가 가속화되고 있는 작금의 시대에 컬처마이닝을 통해 특정되는 문화요소는 다양한 부가 가치를 만들어낼 수 있을 것으로 기대됨은 물론이며, 근본적으로는 다문화·다언어 사회의 이해의 폭을 증대시킬 수 있는 이문화(異文化) 커뮤니케이션 및 다양한 교육적·연구적·실용적 활용가치를 담지하고 있다.

カルチャーマイニング(Culture Mining)

李峻瑞

(聖潔大學校文化フレームビッグデータ研究所所長)

超連結、超境界、超融合された多文化/異文化/多言語社会において、幾何級数的に蓄積されたビッグデータの中から、より豊富な付加価値(value)とデータ分析による知識情報(insight)を見出す方策として、カルチャーマイニング(Culture Mining)手法を提案する。

Culture Mining is the process of finding useful or interesting ‘cultural elements’ -life styles, patterns, fashions, trends, models, beliefs, rules, frames etc.- of a specific region or a generation from unstructured text, various image sets by comparing different languages and cultures.

カルチャーマイニングとは人間文化を表現する多様な媒体(テキスト、イメージ、動画など)を分析し、言語/地域/性別ごとに文化を構成する文化要素と、それらの相互関係を見つけ出していくプロセスである。

特許: 420200819840

カルチャーマイニングは機能、性能、信頼度など、基本的に特定の財貨やサービスの商品価格に反映される経済的な商品価値とは別に、消費者の感性を刺激し感動と共感を得ることで、商品価値をさらに極大化させる「感性価値」という付加価値を創出することができる。

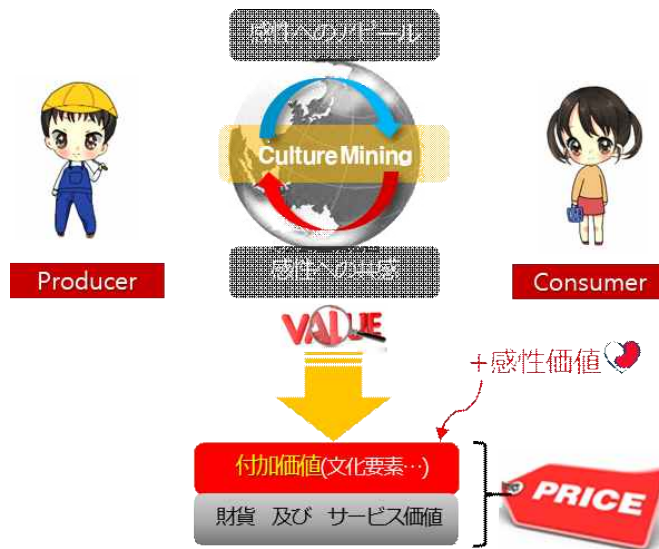


図1 カルチャーマイニングが創出する経済価値

李俊瑞(2021:71)_『Rカルチャーマイニング』博文社

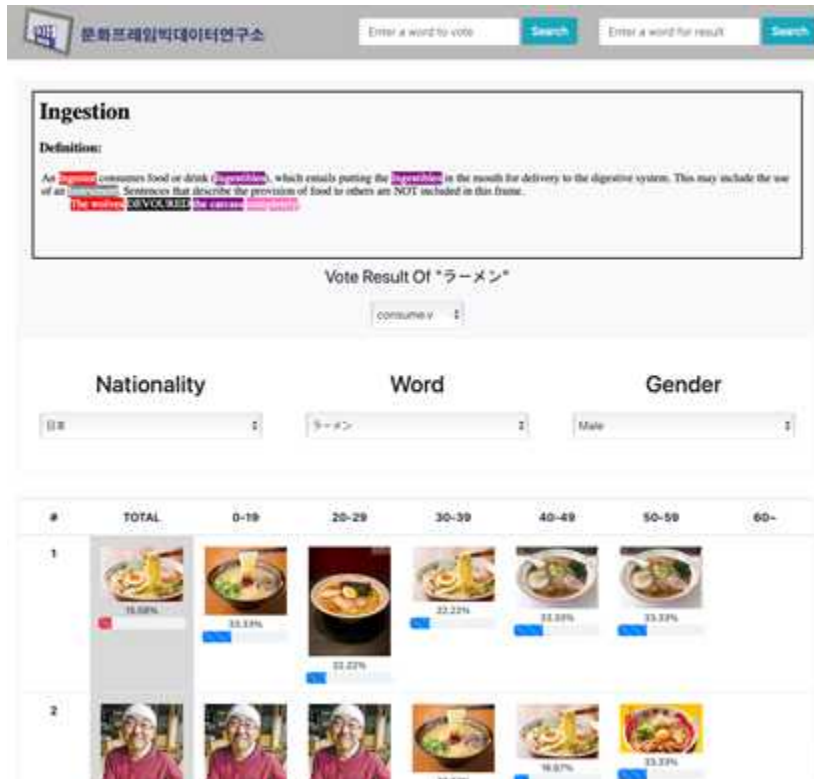
消費者のニーズが多元化し、多文化、グローバル化が加速している昨今の時代において、カルチャーマイニングによって特定される文化要素は多様な付加価値を生み出すことが期待されるのはもちろん、根本的には多文化・多言語社会の理解の幅を広げる異文化コミュニケーションや、多様な教育的・研究的・実用的活用価値を盛り込んでいる。

문화이미지프레임망(Cultural Image Frame Network, CIFN) 소고(小考)

임상순
(성결대)

최근 텍스트로 구성된 다양한 데이터를 수집하고 가공하여 이를 기반으로 요소 정보를 추출하는 연구가 활발히 진행 중이다. 예를 들어, 텍스트 기반 리뷰데이터를 수집하고 해당 리뷰데이터를 통해 사용자의 감성을 분석하여 이를 추천시스템과 같은 전문화된 시스템에 적용하는 등 지능화된 서비스 구현에 적용되고 있다. 하지만, 텍스트 기반의 정보에서 요소 정보를 추출하는 것과 함께 이미지 데이터에서 요소 정보를 추출하여 병합한다면 보다 다양한 정보 연관성을 추출하고 활용할 수 있다. 본 연구는 이미지 기반 문화 요소 추출시스템에 관한 연구로서 사용자들이 검색하는 단어에 대해 대표 이미지 정보를 나열하고 선호도를 조사할 수 있는 프레임워크를 제공한다. 한국, 중국, 일본의 특정 연령대, 성별에서 조사되는 대표 이미지에 대한 선호도를 손쉽게 수집함으로써 향후 관련 빅데이터를 구축하고 한국, 중국, 일본 세 나라의 문화 요소를 이미지 데이터 기반으로 추출해 낼 수 있는 장점을 가지고 있다. 따라서, 본 시스템을 통해 얻어지는 이미지 기반 문화 요소 정보를 향후 텍스트 기반 문화 요소 정보와 결합하여 문화 요소의 특징을 다양한 관점에서 도출할 수 있는 환경을 제공할 수 있다.

<대표 그림>

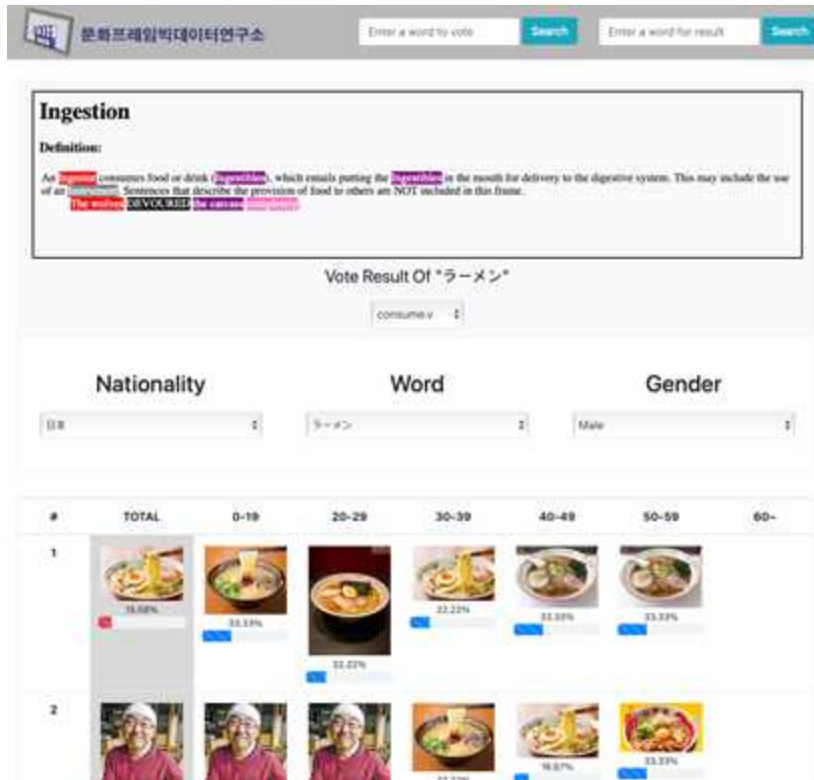


文化イメージフレーム網(Cultural Image Frame Network, CIFN) 小考

林尚諄
(聖潔大)

最近ではテキストで構成された様々なデータを収集、加工し、これを基に要素情報を抽出する研究が活発に行われている。例えば、テキストベースのレビューデータを収集することで当該レビューデータからユーザの感性を分析し、それを推奨システムなどの専門化されたシステムに適用するなど知能化されたサービス実装に適用されている。しかし、テキストベースの情報から要素情報を抽出するとともに、イメージデータから要素情報を抽出して併合すれば、より多様な情報関連性を抽出して活用することができる。本研究はイメージ基盤の文化要素抽出システムに関する研究であり、ユーザーが検索する単語について代表的なイメージ情報を羅列し、選好度を調査できるフレームワークを提供する。韓国、中国、日本の特定の年代、性別で調査される代表的なイメージに対する選好度を容易に収集することで、今後、関連ビッグデータを構築し、韓国、中国、日本の3国の文化要素をイメージデータ基盤に抽出できる長所を持っている。したがって、本システムによって得られるイメージ基盤の文化要素情報を今後テキスト基盤の文化要素情報と結合し、文化要素の特徴を多様な観点から導き出し得る環境を提供することができる。

<代表 図>



AIとテキストマイニング

落合由治

(台湾 淡江大学日本語文学科)

1. はじめに

論者は所属学科の同僚と共に、台湾の大学の日本語文学科で日本語教育と日本語による日本語学、文学、歴史、社会文化などの分野の人文系教育を今まで行っている。台湾でも世界各国の動きと連動し、急激に進む高度情報化社会に対応できる人材育成に方向転換し、現在、各レベルの教育機関で情報化教育を取り入れる試みが進んでいる。2030年の世界を予想して取り上げられた「五大新常态STEEP」(Society, Technology, Economy, Environment, Politics)を前提に、「高教学習創新計画」が募集されて、教育内容の革新を前提に教育方針が立てられている。¹⁾また、台湾で開催した2020年の「行政院第十一次全国科学技術會議」の報告書では、世界と台湾がそれぞれ直面すべき課題を検討し、「台湾2030—創新、包容、永續」のビジョンを掲げている。特に教育と関わりのある科学技術においては、AI(Artificial Intelligence、人工知能)とDX (Digital transformation、デジタルトランスフォーメーション)の導入を緊急の課題として、対応を呼び掛けている。²⁾論者の所属校でも、大学レベルで「AI創智學院」を各学科の区別を超えて創立し、現在の情報化技術を

-
- 1) 経済界では、今後の世界の変化を5つの常態として対応を呼びかけている。CIO Taiwan (2021) 「推動台灣數位經濟發展，2022年企業需關注的五大趨勢」
<https://www.cio.com.tw/top-five-trends-in-taiwans-economic-development-and-business-needs-in-2022/>
 (2022年1月10日加筆閲覧)。2021年の教育部の高等教育支援計画では、新時代の「五大新常态」を前提にした教育内容創造が求められている。「教育部辦理補助素養導向高教學習創新計畫徵件須知」
https://rdfiles.ym.edu.tw/public/0134428A00_ATTCH1.pdf (2021年10月1日閲覧)。
- 2) 台湾科技部「第十一次全國科學技術會議總結報告」
<https://web.most.gov.tw/tc/11th/upload/202012/160879188382087.pdf> (2021年10月15日閲覧)。

教育や産業に応用する教育と支援を行うことになり、2020年9月の新学年度から開放され、マイクロソフトと協働して今後毎年6000名のAI人材を育成する計画であるが、台湾各地の大学でも情報関係教育の導入が広がっている。3)

一方、こうした時代の変化に対応するため、人文系の日本語文学科でも応用できる情報処理技術および自然言語処理技術（以下、情報学とする）について学習を進め、現在まで継続して、導入と応用の試みを行っている。2018年から始め2021年までに、1) 自然言語処理の基礎について学ぶ勉強会やワークショップの開催、2) 日本語許育関係学会での情報処理関係講演者による講演と研究発表、3) 大学の授業として外国語学部生向けの入門授業の開設、4) 大学院での自然言語処理の基礎に関する授業の開設、5) 関係教員による個人の研究や学生指導への応用の試行という5つの方向で、今まで接触のなかった情報科学や自然言語処理という新しい分野について、関係する知識や情報との接点を開拓し、学科の教育内容や卒業生の進路開拓につなげようとさまざま試みをおこなっている。4)

本論文では、今まで実施してきた導入と応用の各試行の中から、特に第三世代のAI技術の発展で、自然言語処理でも従来のレベルから次第に新しい領域へのアプローチが行われてきている中で、実社会にも進出しているテキストマイニング技術について取り上げて、応用の可能性を考えていきたい。

2. 自然言語処理と言語データの関係

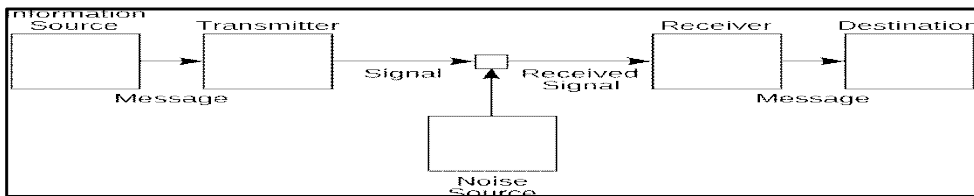
現在、第三世代AIの発達で現代社会と文字通り密接不可分な関係になっている情

3) 淡江大学(2020)「AI創智學院孕育智慧人材淡江大小事」<https://events.tku.edu.tw/FocusDetail.cshtml?sn=1700> (2021年10月16日閲覧)。以後、台湾各地の大学で情報関係教育の導入が模索されている。中時新聞網(2021)「跨領域人才卡位大企業！台灣「十間多元系所大學」培養職場力」<https://www.chinatimes.com/realtimenews/20220108001038-260405?chdtv> (2022年1月10日閲覧)。

4) 学科でのAI関係の内容の導入についての試行は、曾秋桂 (2021)「日本語教育のつながりとひろがり—AIとHIを兼ね備えた外国語(日本語)人材2.0の育成を目指して」日本語教育研究 54, pp23-36、落合由治、曾秋桂、王嘉臨、葉菱(2020)「人文系教育への情報処理・自然言語処理技術の導入と応用」『淡江日本論叢』42号pp.45-63を参照。韓国でも落合由治 (2021)「日本語関係学科での自然言語処理技術の導入と応用の可能性」『韓國日本語學會第43・44回 國際學術發表大會予稿集』pp.149-154等で紹介している。

報処理技術は1950年代に生まれた情報処理とコミュニケーション理論に基づいている。基本モデルとなっている情報処理論はシャノン&ウィーバーモデルで、1948年にクロード・シャノンが発表した「通信の数学的理論」⁵⁾によって情報理論分野が創設された。

図1 シャノン「通信の数学的理論」による情報通信モデル⁶⁾



シャノンは、電気的情報通信は以下の5要素で構成できると提起した。①メッセージを生成する情報源(information source)、②メッセージを操作して通信路を介して送信可能な信号を生成する送信機(transmitter)、③メッセージを構成する情報を含む信号が送信される媒体である通信路(channel)、④配信された信号をメッセージに変換し直す受信者(receiver)、⑤メッセージの送り先として意図されている人または機械である宛先(destination)である。同時に、情報量(エントロピー)と冗長性の概念を導入し、ビットという用語を情報単位として導入した。電話、電信からラジオ、無線通信、テレビ、そして電気信号を利用するあらゆる製品分野から、各種コンピューターさらに現在のインターネットとWi-Fi通信の発展まで、すべてこのモデルが元になっている。19世紀末に生まれた電気文明が現在のような高度情報通信社会に発展したのも、このモデルによる。現在、メタバースで人間の現実社会をインターネットのヴァーチャル空間に転移させる構想も同じモデルを前提にしている。

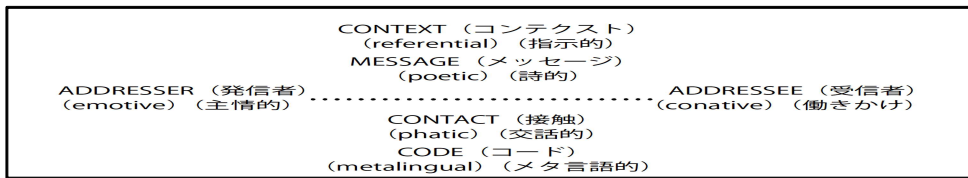
もう一つの重要なモデルは、ローマン・ヤーコブソンのコミュニケーション理論

5) Shannon, Claude E. (July 1948). "A Mathematical Theory of Communication". Bell System Technical Journal 27 (3): 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x. これは、1949年に The Mathematical Theory of Communication(0-252-72546-8)というタイトルで出版され、ウォーレン・ウィーバーによる解論文が含まれており、シャノン&ウィーバーモデルと呼ばれるようになった。

6) Wikipedia" A Mathematical Theory of Communication" https://en.wikipedia.org/wiki/A_Mathematical_Theory_of_Communication (2022年2月18日閲覧)

である。ヤーコブソンは、1958年の「“Metalanguage as a Linguistic Problem”」と1960年の「Closing statements: Linguistics and Poetics」で、人間のコミュニケーションは発信者と受信者の相互行為で、コミュニケーションに不可欠な六つの因子（発信者、受信者、メッセージ、コンテキスト、コード、コンタクト）を挙げ、その一つ一つに相異なる言語機能（主情的、働きかけ、詩的、指示的、メタ言語、交話的）が対応していることを指摘し、その後の対人コミュニケーション理論の先駆けになった。⁷⁾

図2 ヤーコブソンの6機能コミュニケーションモデル⁸⁾



大学での日本語関係人文社会系教育（以下、人文系）に情報処理関係教育（以下、情報学）を導入し応用する場合、二つの課題が考えられる。一つは人文系の教育と研究の特徴を情報学の導入と応用によって今以上に活用し発展するにはどうすればよいか、もう一つは具体的に情報学の各種内容について何をどのように人文系の教育と研究に応用するかというということである。この問題を考える時、コミュニケーションモデルの問題として自然言語処理と言語データの関係を考察しておく必要がある。人間の使用する言語を資料に情報処理をおこなう分野は自然言語処理であるが、完全に人間と同じ言語運用能力がAIに実現すれば、図1のようにAIが人間と同じように情報信号を処理して発信者と受信者の役割を果たすことが可能になると考えられている。図1の関係は量的問題であり、ビッグデータなどが大きな機能を果

7) 1958年の発表は研究大会のもので、現在は、Roman Jakobson, “Metalanguage as a Linguistic Problem,” in Roman Jakobson, *Selected Writings*, vol. 7 (Berlin: Mouton Publishers, 1985)、および Sebeok, Thomas Albert (1960). “Closing statements: Linguistics and Poetics” (PDF). In Jakobson, Roman (ed.). *Style in language*. New-York: M.I.T. p. 470。日本語では、朝妻恵里子 (2009) 「ロマン・ヤコブソンのコミュニケーション論—言語の「転位」」『スラヴ研究』No. 56 (2009) pp.197-213が詳細に論をまとめている。

8) 同注7朝妻恵里子 (2009) に拠る。

たすようになったのは、こうした処理ができる情報や信号（物理的現象の数量的処理）をAIが大量に扱えるようになったためである。また現在の自然言語処理の目覚ましい発展は、図1で理解できるような、言語の形式や信号で直接的明示的に表現されている部分を対象に特徴抽出技術が開発され、機械学習や深層学習でAI学ぶことができるようになったためと考えられる。

しかし、人間のコミュニケーションには元々、図2のように言語の形式や信号では直接的明示的に表現されていない（あるいは表現できない）、各種の条件と言える言語機能（ヤーコブソンでは主情的、働きかけ、詩的、指示的、メタ言語、交話的）があり、自然言語処理が言語データを処理する場合、それらを反映することがまだできていないため、期待する処理が実現できない問題が起こっていると言える。人文系の研究や教育は図1のように何らかの対象や資料を複製、再現することではなく、図2のように言語の形式や信号では直接表現されていない各種の言語機能を言語資料から読み取れることを研究対象にし、それを表現し、同時にそれが読み取れる母語や多言語の人材を育成しているのである。人文系での情報学の教育また自然言語処理スキルの導入では、図1の情報通信モデルと図2のコミュニケーションモデルの双方に目を向けることで、情報学で何ができ、何ができないかを定位していくことで、今後の新しい発展を目指すことができよう。以下では、テキストマイニングを例にして、図1の情報通信モデルと図2のコミュニケーションモデルの双方に目を向けることで何ができるか考えてみたい。

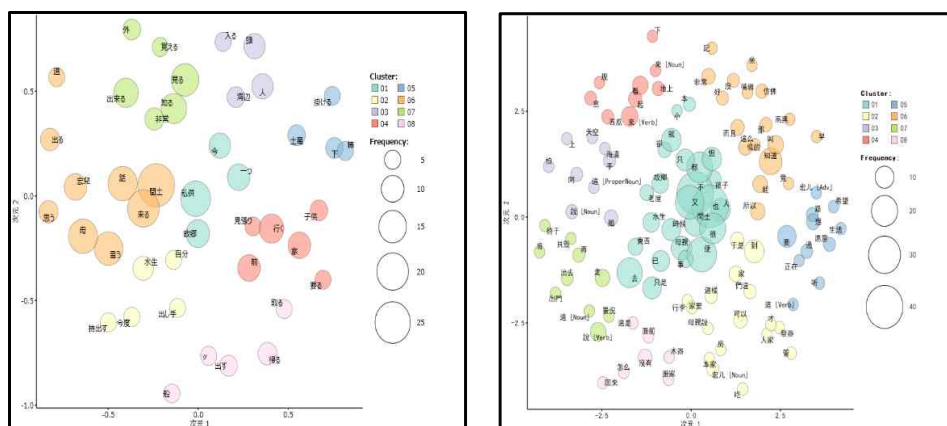
3. テキストマイニングの人文系での応用

元々、自然言語処理は図1のように量的処理に基づく言語処理が中心だったため、実際に社会的に運用されている言語表現を扱うようになった現在、使用される分野、場面等によって、言語表現には異なる形式や表現があり、それに応じた意味や用法の違いがあることが2000年代からドメイン問題として認識されるようになってきた。ドメイン問題は図2のモデルで生じるような言語の多様性であり、人文系では当然のことであるが、自然言語処理にとっては今後も対応してゆくべき課題に

なっている。9)こうした分野では、人文系の研究が直接、接続できる機会も生まれると言える。今回は、ドメインである自然言語と文章ジャンルの相違についてテキストマイニングの結果を検討していきたい。韓国語、日本語、中国語などの自然言語の相違は自明の事実で、また文学、評論等の文章ジャンルの差異も自明な社会的表現ジャンルの違いであるが、これらは自然言語処理ではやはり異なるドメインの問題になる。

まず、文学作品として、日本語と中国語が対照しやすい魯迅「故郷」の日本語訳と中国語版を比較してみた。以下の図3-1が結果である。いずれも結果は左が日本語、右が中国語である。多次元尺度構成法では双方とも同じように登場人物に関する語が中心に来て、動詞などが周辺に分布する。これらは、直接、内容理解の手掛かりにはならない語彙で、これだけでは有効なキーワードにならない。

図3-1日本語と中国語の小説の多次元尺度構成法の結果

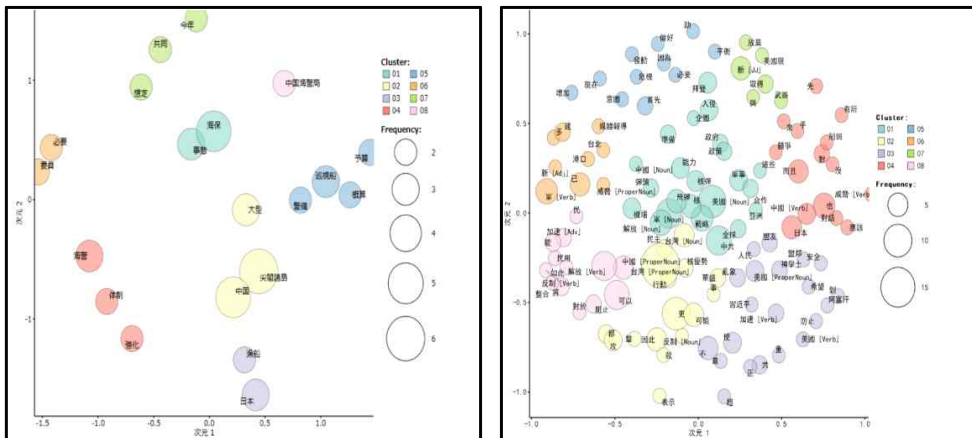


次に、日本語と中国語の評論作品を取り上げてみる。資料は、ともに時事問題を論じた内容で、日本語は読売新聞「海上保安庁 尖閣諸島を守る体制強化せよ

9) 自然言語処理では社会的言語使用分野別のドメインへの技術対応が継続して研究課題になっている。森信介 (2012) 「自然言語処理における分野適応(<特集>知識の転移)」『人工知能学会誌』27-4pp.365-372、井佐原均 (2021) 「自然言語処理と社会実装」『電子情報通信学会基礎・境界ソサイエティ Fundamentals Review』14-4pp.297-307等参照。テキストマイニングには、樋口耕一 (2020) 『社会調査のための計量テキスト分析—内容分析の継承と発展を目指して』ナカニシヤ出版のKH Coderを使用した。

2021/09/15」¹⁰⁾、中国語は自由時報の論壇「台灣可以戰勝近期的最大危機」をとりあげた。10)以下の図3-2が結果である。多次元尺度構成法では日中両言語とも同様に、事件に関する場所、主体などの語が中心に来て、事件の個々の局面での登場物や問題等に関する名詞が周辺に分布する。社会問題を論じた評論作品では、文学作品よりも、事件の主体、場所、問題がクラスターのまとまりで各個に表示されていると考えられ、問題点や話題がかなり明確に理解できる。主な名詞のクラスターのまとまりで、話題が要点として表示されることで、各クラスターのまとまりで、主要な話題を理解できる。

図3-2日本語と中国語の評論の多次元尺度構成法の結果



さらに、文学作品よりもキーワードを的確に抽出できるので社会的文化的に注目する点を抽出することが容易であり、評論全体を読まないでも、主要な話題だけを読むことが可能で、テキストマイニングを元に質的理解に拓ける手掛かりが最も明確に明示されていると言える。さらに言えば、違う論者は別の意見なので、一定の論者の傾向で一文書を精読するよりも、話題を中心に多面的な理解をする可能性に繋がると言える。今回見た二ジャンルの中では、テキストマイニングの内容理解への応用が、最も容易にでき、また創造性読解に発展できる分野と考えられる。

10) 読売新聞(2021)「海上保安庁 尖閣諸島を守る体制強化せよ 2021/09/15」<https://www.yomiuri.co.jp/editorial/20210914-OYT1T50313/>、自由時報(2021)「星期專論」台灣可以戰勝近期的最大危機」<https://news.ltn.com.tw/news/politics/paper/1472334> (2021年10月12日閲覧)。

4.おわりに

言語類型論では文法的差異が強調されるが、テキストマイニングを使うと言語間の差異はなく、文章ジャンルの違いのほうが明確に表示された。今回比較した文学と評論では、一回限りの出来事を描く文学と、概念の構築で説明や意見表明をするジャンルである論説との間で、大きな差異があり、これは言語の相違によらず、見られる傾向であることが明らかになった。

文章ジャンルや会話ジャンルの違いによるテキストマイニングの実施を考える時、量的指標と質的研究の相補性を確保することで、新しい人文社会系の研究の可能性が見え、同時に異なる言語間でも比較文化研究、異文化理解研究の展開が期待できる。人文社会系でのデータマイニング技法の利用は、その結果の利用方法自体が大きな課題、同時に新しい挑戦であり、特に言語間の相違を超えて比較文化的考察にテキストマイニングを応用する方法を開発することは、今後の大きな課題と言えよう。

註記：参考文献は注をご参照。本論文は科技部研究案MOST 109-2410-H-032 - 061 -MY3の研究成果である。

AIと外国語クリエイティブ・ラーニング

検校裕朗
(極東大学校)

1. はじめに - Society5.0 (創造社会)

「Society5.0」とは、日本が発案している未来社会のコンセプトを表した言葉である。科学技術基本法に基づき、日本の内閣府が5年ごとに改定、公開している科学技術基本計画の第5期（2016年度から2020年度の範囲）でキャッチフレーズとして登場した。その「第5期科学技術基本計画」の中で「Society5.0とは、サイバー空間（仮想空間）とフィジカル空間（現実空間）を高度に融合させたシステムにより、経済発展と社会的課題の解決を両立する人間中心の社会（Society）」と定義されている。狩猟社会（Society1.0）、農耕社会（Society2.0）、工業社会（Society3.0）、情報社会（Society4.0）に続く新たな社会として、Society5.0（創造社会）は定義される。現在、Society4.0には様々な課題が存在するのであるが、それらの課題を先端技術の応用によって超克していくことを目指した概念といえる。Society4.0が課題として抱えているのは、知識・情報などの共有が不十分であり、他分野とうまく連携させることができず、繋がりが不足しているという点である。AIが活用され始める前までは、基本的に人間だけの能力に頼らなければならなかったが、その為、膨大な情報の中から必要な情報を抽出する作業や、抽出した情報を分析、処理していく作業などにおいて、大きな負担がかかってしまっていた。近年、特に少子高齢化に伴う人材不足が深刻化していることもあり、労働力の確保に困難をきたしている企業も少なくない。こういった課題を克服するために提唱されたのが、Society5.0である。政府はSociety5.0の実現に向けて、〈1.デジタル市場のルール整備 2.フィンテック・金融分野 3.モビリティ 4.コーポレート・ガバナンス〉の4つの方針を打ち出して国内

成長力の強化を図っており、2021年には「デジタル庁」が設置され、今後さらにDX（デジタルトランスフォーメーション）に向けた動きを加速させようとしている。

Society 5.0を実現するDXとして、5G（第5世代移動通信システム）、IoT（Internet of Things）、AI（Artificial Intelligence）、ビッグデータ、AR（Augmented Reality：各超現実）、VR（Virtual Reality：仮想空間）、各種ロボット、ブロックチェーン（分散台帳技術）、自動運転、ドローン、リアルタイム通翻訳などの先端技術の活用が必要不可欠となる。高精度の分析や予測を実現する技術によって、これまで人の能力に頼らざるを得なかった作業が補完され、さらなる効率化や生産性向上に繋がられる。それらの技術を応用し、例えば、エネルギー、医療、ヘルスケア、モビリティ（IT技術を駆使し、人々の移動を効率化し、快適なものにする）、観光、町づくり、ものづくり、スマート農業、食品、物流、消費、セキュリティ、監視、警備、防災等、各分野での取り組みが期待されている。

Society5.0は、日本が最初に提唱したものであるが、筆者は世界に通じるコンセプトでもあると考える。国連が2030年までに達成することを目標として提唱した「SDGs：Sustainable Development Goals（持続可能な開発目標）」も、Society5.0と深く関係している。SDGsは、貧困問題・エネルギー・教育格差などの「世界が到達すべき17の目標」と、それらの目標を達成するための169のターゲットから成り立っている。経団連では、革新技術の活用によって「経済発展」と「社会的課題の解決」を両立させ、SDGsを達成するためのコンセプトとしてSociety5.0（創造社会）を提案した。SDGsとSociety5.0は、現在の日本を含む世界が抱える課題を克服し、より暮らしやすい環境を実現するという目標があるため、非常に深い関係性があるといえる。〈検校(2022)〉

経団連が発表した、『Society5.0 - ともに創造する未来 - 』¹⁾では、Society5.0時代を生きていく上で「世の中を変える『想像力』とそれを実現する『創造力』」が必要に

1) 経団連はこれまで「超スマート社会」とされてきたSociety5.0を「創造社会」と呼称することを提唱した。Society5.0で目指すべき人間中心の社会では、利便性や効率性の実現を主目的とするのではなく、デジタル技術・データを使いながら、人間が人ならではの多様な想像力や創造力を発揮して、社会を共に創造していくことが重要であると考え、Society5.0とは、創造社会であり、「デジタル革新と多様な人々の想像・創造力の融合によって、社会の課題を解決し、価値を創造する社会」であるとした。

なる。」と述べられている。そして、Society5.0を達成する上で願われる人材像として、「自ら主体的に課題を見つけ、AIを活用して課題を解決できる人材」が求められている。Society5.0を実現するためには、AIを使いこなすことができるリーダーが欠かせない。様々なプロジェクトを遂行するチームの中でリーダーシップを発揮する為には、責任感と協調性と洞察力を持って課題と向き合っていく力が必要であり、それに加え、新しい価値を生み出す能力を備えていれば、より良い「もの」やサービスを作り出すことが可能になる。

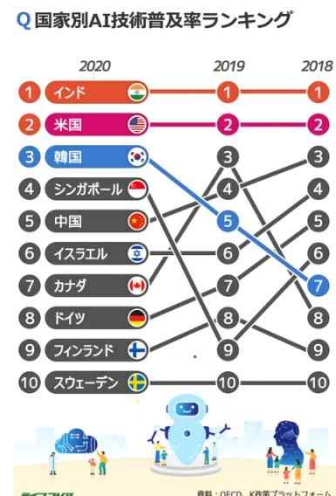
それらの観点からも、未来社会に向けて教育内容も発展する必要がある、教育学の1つである外国語教育学も、日本語教育学も発展していく必要があると考える。

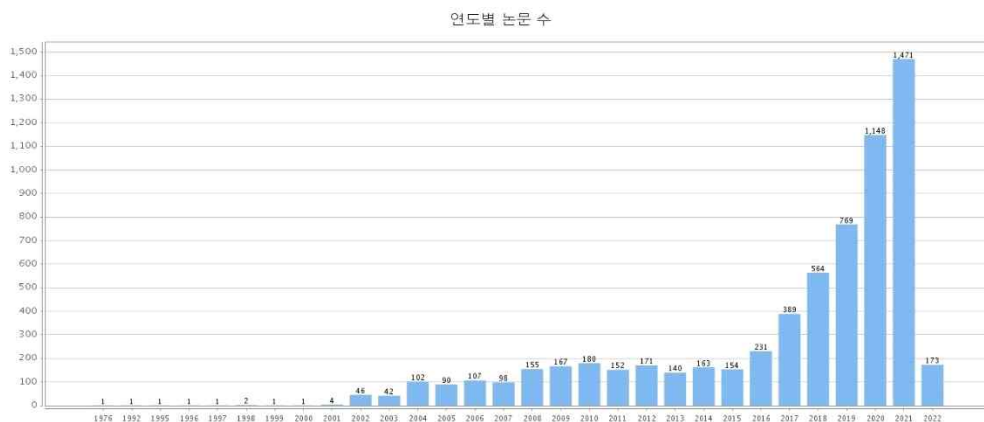
2. 韓国の日本語学・日本語教育学におけるAI

民間シンクタンクK政策プラットフォームが経済協力開発機構（OECD）の資料を分析した結果によると、韓国は2020年国家別AI技術普及率ランキングで3位になった。1位はインド、2位は米国、そのほかシンガポール、中国、イスラエル、カナダ、ドイツ、フィンランド、スウェーデンがトップ10入りを果たした。

該当のランキングはOECDが2016年から2020年のグローバル人的資源（HR）サービスであるLinkedInのデータベースを活用し、各国のAI技術普及率と標本内のすべての国家・地域の平均AI技術普及率の比率を推定し、算出したものである。2016年に4位だった韓国は、2017年には6位、2018年には7位まで下降したが、2019年に5位に上昇し、昨年には3位となり、歴代最高順位を記録した。・2年間でランキングが7位から3位へ、4ランクアップした。一方、デジタル先導企業と中小企業間の生産性の格差はよりいっそう拡大したという懸念も提起されている。（朝鮮日報日本語版 2021.4.2）

このように、IT分野やAI技術に関しては普及してい

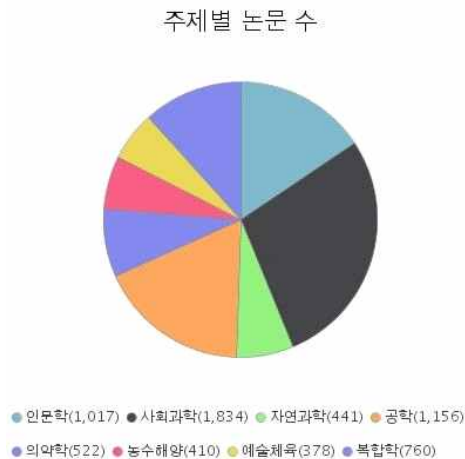




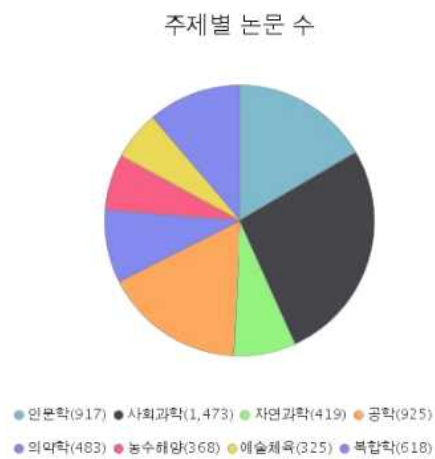
る韓国であるが、日本語学・日本語教育学における活用はいかなる状況であろうか。韓国研究財団の学術誌引用索引(KCI: Korea Citation Index)に登載されているAI全体の論文数は、現在のところ6,525本(2022.2基準)であり、グラフを見てもわかるように、年々増加の傾向にある。(2021年度、1471本。)また、2016年以降急激に増加している。韓国でAI(人工知能)が社会一般の注目を浴びるようになったのは、2016年のアルファGoと李世乭九段との大局の影響が大きいと考えられるが、それ以外にもグーグル翻訳がビックデータを使ってマシンラーニングの結果を取り入れるようになり、ネイバーの翻訳サービスPAPAGOも同様の翻訳技術を導入した。新聞のデータベースの検索結果では2016年以降AI関連記事が増加しているが、学問研究でも同様に、2016年を前後にして大きな変化が見受けられ、量的増加の影響が表れている。

分野別には1.社会科学(1,834)、2.工学(1,156)、3.人文学(1,017)、4.複合学(760)、5.医薬学(522)、6.自然科学(441)、7.農水海洋(410)、8.芸術体育(378)のように多岐に渡る。下の6か月前主題分野別論文数と比べると増加数順に、社会科学<1,473→1,834(+361)24.51%up>、工学<925→1,156(+231)24.97%up>、複合学<618→760(+142)22.98%up>、人文学<917→1,017(+100)10.91%up>、芸術体育<325→378(+53)16.30%up>、農水海洋<368→410(+42)11.41%up>、医薬学<483→522(+39)8.07%up>、自然科学<419→441(+22)5.25%up>のように、社会科学、工学分野の増加が多い。人文学は更なる増加が望まれる。「AI、日本語」に該当する論文は

16本であるが、内容を見ると日本語学・日本語教育分野に関するものは5本である(2022.2.基準)。〈落合由治(2020)、申忠均(2020)、曾秋桂(2021)、신민철(2021)、金ソヒ(2021)〉



<Figure 1 KCI主題分野別論文数 2021.8.>



<Figure 2 KCI主題分野別論文数 2022.2.>

2016年から、人工知能の開発などに直接関わる理工系の学問分野ではない、人文分野においても人工知能をめぐっての様々な議論や研究報告が行われるようになり、ついに、2018年には『人工知能人文学』という人文系の学術誌が創刊されるに至った。人文学においては、哲学分野での議論が活発であるが、言語研究、外国語教育に関しては、まだ人工知能を活用した外国語教育の胎動期と言え、外国語教育の分野では、特に英語教育分野の研究成果が量的に断然多く、その中でも特に、チャットボット、AIスピーカー、機械翻訳に関する論議が活発である。英語教育以外の分野では、まだそれほど注目を集めてはいないのが現状であり、それぞれの言語教育の分野で数編の研究結果発表を見るにとどめ、また、分野的にも偏りがあり、その多くが機械翻訳・通訳の分野のものである、機械翻訳には古典漢文、とりわけ韓国内の歴史資料の機械翻訳への関心が高く、韓中/中韓翻訳に関する考察が目立つ。また、日本語教育関連では、人工知能関連の研究は未開拓の分野であり、機械翻訳関連の分野は、比較的早い時期から関心が寄せられており、日本語教育への応

用なども提案されているが、今後、AI音声アシスタントの活用などは積極的に取り入れる必要がある。〈申忠均(2020)〉

落合(2020)は、テキスト・マイニングプログラムを活用して、人文系研究と教育が日本語教育と結びついて、日本語教育現場やカリキュラムでのAI技術応用の端緒になっていく可能性をまとめた。曾(2021)は、AIとHIを兼ね備えた外国語人材2.0の育成といった目標を達成させるために実行した成果を述べ、신(2021)は、人工知能(AI)翻訳機を活用して漢語の特徴分析を試みており、金(2021)はテキストマイニングを活用した「X{まで}」構文の語彙分析を行っている。

翻訳分野の中で日本語教育に通じる研究には、李朱利愛・朴惠璟・朴星姝(2017)「逐次通訳授業のスマートラーニング活用案構想-日本語逐次通訳授業の例を中心に-」、李朱利愛・朴惠璟・尙禹延(2018)「同時通訳授業の補助道具としてのアプリケーション設計-課題提出及びフィードバック機能を中心に-」等がある。

更に、論文化されていないが、プログラムの中でAIが動いていると思われるアプリケーションの活用方法と実践を報告として行われているものは、金侑嬪(2021)「学習ツールを活用したオン・オフラインの新世界活動授業」、趙湫羅(2020)「コロナ時代 中学校実際授業事例-クラスカードを中心に-」、趙湫羅(2021)「中学校でのエデュテックの活用事例-クラスカードとパドレット及び様々なクイズ形式の学習ツールについて-」等、大学教育よりも中学・高校の中等教育分野において先行している状況である。

今後も、AIを導入した日本語教育の更なる発展が見込まれ、それを研究面からも深めていく必要性から、2021年9月の韓国日本語学会国際学術大会の時に「AIとクリエイティブ・ラーニング研究会(AI-CL)」を創立し、同時に韓国日本語学会ともMOUを締結した。

3. クリエイティブラーニング

筆者があげる『クリエイティブ・ラーニング』は、2019年に日本で出版された『クリエイティブ・ラーニング：創造社会の学びと教育』(慶應義塾大学出版会)の編

著者である慶應義塾大学総合政策学部教授・井庭崇氏による新しい教育概念である。井庭氏はクリエイティブ・ラーニング全体を提唱しており、外国語教育に応用する試みについて具体的な言及はないが、筆者は2020年の「日本語教育研究国際シンポジウム」（主題：クリエイティブ・ラーニングを目指す日本語教育）でのパネリストとしての参加を契機に、井庭氏とやり取りし「外国語（日本語）クリエイティブ・ラーニング」を目指すことにした。2021年「AIと日本語教育国際シンポジウム」（主題：クリエイティブラーニングを目指すAIと日本語教育）で、筆者が行った講演とパネルディスカッションを通して、この分野の研究の必要性和重要性を更に痛感した。

井庭氏は、クリエイティブ・ラーニング（創造的な学び）とは、「つくることで学ぶ」という新しい学びのスタイルであるとする。「自ら知識を構成する」学習観は、「アクティブ・ラーニング」や「プロジェクト型学習」「探究学習」のように、学び手自身による活動がベースとなるが、クリエイティブ・ラーニングは、何かを「つくる」ことをより一層重視する。そして、これからの学校は、創造的に学ぶための「つくる」経験を積む場となり、教師は、生徒が「つくる」ことを支援するだけでなく、一緒に問題に挑戦し、一緒につくることに取り組む仲間である「ジェネレーター」となる必要があることを述べている。

「序章 構成主義の学びと創造—クリエイティブ・ラーニング入門」では、「デューイの「経験の連続性」と「経験の再構成」等について述べられている。プラグマティズムの教育哲学者であるジョン・デューイは「教育者は他のどのような職業人よりも、遠い将来を見定めることにかかわっている」と述べた。

筆者の学部時代の恩師は、日本イギリス理想主義学会会長、日本道德教育学会会長代行、日本デューイ学会常任理事等を歴任した行安茂氏である。行安茂(2018)『アクティブ・ラーニングの理論と実践』では、デューイが「なすことによって学ぶこと」(Learning by Doing)の主張者であったことを想起しながら、アクティブ・ラーニングの先駆者であったと見ることができるとし、デューイの哲学から現代のアクティブ・ラーニングに通じる理論を読み解いている。

井庭氏は、教わる時代から、話す時代、そして、創る時代になり、今後さらに創る

ことを通して学ぶ時代になると主張する。それはまさに「Learning by Creating」であり、「Learning through Creating」であり、つくことは変えること、学ぶことである。そこでは、学生が「無我の創造」の経験をするのが大切であり、パターンランゲージの活用が重要であり、創造的読書（クリエイティブ・リーディング）の推進も求められ、創造社会における人々の創造実践の支援が肝要になってくるとしている。

4. 創造の7つの原理

創造の第一の原理は、つくり手の意図・作為によるコントロールを手放すということである。

この感覚について、実際に、様々な分野の作家や芸術家が語っている。例えば、村上春樹は、自らの創作について、次のように語っている。

「本を書き始めるとき、僕の中には何のプランもありません。ただ物語がやってくるのをじっと待ち受けているだけです。それがどのような物語であるのか、そこで何が起ころうとしているのか、僕が意図して選択するようなことはありません。物語が何を求めているのかを聴き取るのが僕の仕事です。」

日常的な感覚では、「つくる」ということは、つくり手が「こうしたい」という思いや、「こうしよう」という頭のなかの青写真や設計図にもとづいて、手を動かし、それを外部に表現することだと捉えられることが多い。しかし、ここで語られている創造は、つくり手の作為によるのではなく、作品に引っ張られて進んでいくようなつくり手像である。

このように、意識的な意図や作為によるコントロールを手放し、無我（egoless）な状態で「つくる」ということを、井庭氏は「無我の創造」（egoless creation）と呼んでいる。以下、7つの原理は、2. 生成と分化（植物が絶えず一つの全体として育つように、つくっているものが育つことを支援する。）3. 内側から感じる（つくっている対象になりきって感じる。）4. 冒険的創造（先が見えないなかで探索的に進んでいく。）5. あるべきかたち（創造の制約・ルールを手掛かりに、パズルのピー

スをはめてつめていく。) 6. 発掘したものを磨き上げる (深く埋もれていた状態から掘り起こしたものの良さが出るように、徹底的に手直しをして美しい状態に整える。) 7. つくることで生きる (つくることが生きることそのものに重なり、生きることの本質となるように、自分の人生を生きる。) となっており、この7つの原理が重要である。〈井庭崇(2019)〉

5. 外国語(日本語)クリエイティブラーニング実践

- 5-1. クリエイティブCM制作プロジェクト
- 5-2. クリエイティブ・ロールプレイ
- 5-3. クリエイティブ・ウェブ・ミニドラマ
- 5-4. クリエイティブ・カナコン
- 5-5. アバター・バーチャル・ホームステイ
〈Flat Stanley・バーチャル・ホームステイ〉
- 5-6. メタバース上での可能性
Gather town, Spotの事例

6. ビッグデータと日本語教育

各種コーパス
 峯松信明: OJAD / 韻律読み上げチュータスズキクン
 李在鎬: jReadability, jwriter, JEV 他
 機械翻訳 (翻訳サイト構築プロジェクト参与経験から)
 シンギュラリティ(「技術的特異点」)の問題

7. おわりに AIとクリエイティブ・ラーニング研究の意義と重要性

AIを活用した日本語教育も、日本語教育におけるクリエイティブラーニングもま

だ、草創期の段階にあり、更なる研究が望まれるところである。今後、日本語教育への応用・発展につながることを切に願い、実践と研究を推進していきたい。

参考文献

- ・ 신민철(2021) 「인공지능(AI) 번역기를 활용한 한일 한자어 비교의 가능성 모색」 『일본근대학연구』 (71), 한국일본근대학회, pp.97-112
- ・ 井庭 崇(2019) 『クリエイティブ・ラーニング:創造社会の学びと教育』 慶應義塾大学出版会
- ・ 井庭 崇(2020) 「クリエイティブ・ラーニング:創造社会の学びと教育」 『淡江大学設立70周年記念行事シリーズ活動 2020年度台湾日本語教育研究国際シンポジウム—クリエイティブ・ラーニングを目指す日本語教育—国際会議予稿集』 台湾日本語教育学会, pp.17-52
- ・ 井庭 崇(2021) 「クリエイティブ・ラーニング:創造社会の学びと教育」 『韓国日本語学会 第43・44回国際学術発表大会論文集』 韓国日本語学会, pp.166-171
- ・ 落合由治(2020) 「AI技術からみた日本語学、日本語教育研究の展望と課題- 日本語教育の繋がりと協働の新領域をめざして-」 『日本語教育研究』 第50輯, 韓国日語教育学会, pp.23-34
- ・ 落合由治(2021) 「AIと日本語教育について」 『韓国日本語学会 第43・44回国際学術発表大会論文集』 韓国日本語学会, pp.159-165
- ・ 金ソヒ(召소희)(2021) 「텍스트마이닝을 활용한 「X{まで}」 構文의 語彙分析—「X{さえ}」「X{も}」との比較・対照を中心に—」 『日本学報』 (126), 韓国日本学会, pp.121-143
- ・ 檢校裕朗(2020) 「クリエイティブ・ラーニングを目指す日本語教育について」 『2020年度台湾日本語教育研究国際シンポジウム—クリエイティブ・ラーニングを目指す日本語教育』 台湾日本語教育学会
- ・ _____ (2021a) 「つながりの日本語教育と日本語クリエイティブラーニング」 『韓国日本語教育研究会 全国日本語教師授業研究発表大会』 基調講演、韓国日本語教育研究会
- ・ _____ (2021b) 「韓国における日本語教育の成果と今後の展望—韓国日語教育学会(KAJE)(20周年期:2019~20年)における会長経験者のナラティブ分析をもとに—」 『日本語教育研究』 第54輯、韓国日語教育学会, pp.39-56
- ・ _____ (2021c) 「韓国の日本語教育におけるAIとクリエイティブラーニング」 『2021年 AIと日本語教育 国際シンポジウム—クリエイティブラーニングを目指す AIと日本語教育—』 招待講演、台湾日本語教育学会, pp.35-44
- ・ _____ (2021d) 「韓国における「AIとクリエイティブ・ラーニング研究会」の創立とその意義」 『韓国日本語学会 第43・44回国際学術発表大会論文集』 韓国日本語学会, pp.172-176
- ・ _____ (2022) 「地球市民教育(GCED)を目指した日本語教育の実践」 『韓国日本学会 第103回国際学術発表大会Proceedings』 韓国日本学会, pp.97-100
- ・ 曾秋桂(2021) 「日本語教育のつながりとひろがり—AIとHIを兼ね備えた外国語(日本語)人材2.0の育成を目指して—」 『日本語教育研究』 (54), 韓国日語教育学会, pp.23-36

- ・辛銀眞(2020)「クリエイティブ・ラーニングを引き出す授業設計」『2020年度台湾日本語教育研究国際シンポジウム―クリエイティブ・ラーニングを目指す日本語教育』台湾日本語教育学会
- ・申忠均(2020)「人工知能時代の日本語教育と研究-諸言語教育・研究分野の成果から考える-」『日本語文学』1(86)、韓国日本語文学会, pp.117-136
- ・日本経済団体連合会(2018)『Society5.0ーともに創造する未来ー』日本経済団体連合会, pp.1-36
- ・行安茂(2018)『アクティブ・ラーニングの理論と実践』北樹出版
- ・李朱利愛・朴惠璟・朴星姝(2017)「逐次通訳授業のスマートラーニング活用案構想 -日本語逐次通訳授業の例を中心にー」『日本言語文化』第41輯、韓国日本言語文化学会, pp.227-251・李朱利愛・朴惠璟・尙禹延(2018)「同時通訳授業の補助道具としてのアプリケーション設計ー課題提出及びフィードバック機能を中心にー」(동시통역 수업의 보조도구로서의 애플리케이션 설계-과제 제출 및 피드백 기능을 중심으로-)『日本語教育研究』第42輯、韓国日語教育学会, pp.99-116
- ・金旻嬪(2021)「学習ツールを活用したオン・オフラインの新世界活動授業」(학습툴을 활용한 온·오프라인의 신세계 활동 수업)『韓国日本語教育研究会 全国日本語教師授業研究発表大会』韓国日本語教育研究会, pp.71-88
- ・趙恩羅(2020)「コロナ時代 中学校実際授業事例ークラスカードを中心にー」(코로나 시대 중학교 실제 수업 사례ー클래스카드를 중심으로-)『韓国日語日文学会2020年秋季国際学術大会 発表論文集』韓国日語日文学会, pp.108-119
- ・趙恩羅(2021)「中学校でのエデュテックの活用事例-クラスカードとパドレット及び様々なクイズ形式の学習ツールについて-」『韓国日本語学会 第43・44回国際学術発表大会論文集』韓国日本語学会, pp.177-182