

## 研究動向

# 外国語教育に関する研究動向

## —— 自然言語処理研究との接点に着目して ——

名畑目 真 吾\*

Shingo NAHATAME

### 1. はじめに

本稿では、近年の外国語教育の研究動向について、特に自然言語処理研究との接点に着目して記述することを目的とする。具体的な研究動向の詳細を述べる前に、まずは以下で外国語教育研究の学際性と自然言語処理研究との関わりについて簡単に説明する。

#### 1.1 外国語教育研究の学際性

「外国語教育研究」とは文字通り外国語教育に関わる研究のことであるが、従来の科研費の審査区分（平成30年度以前）で外国語教育は「人文学」の「言語学」分野の1つとして含まれていた。また、平成30年度から新たに適用された審査区分においても、外国語教育は「文学、言語学、及びその関連分野」の1つとして記載されている。このように、外国語教育研究は一般的には人文学、及び言語学に属する研究分野の1つとして捉えることができる。しかしながら、外国語教育研究として実際に行われる内容は多岐にわたり、外国語教育は人文学や言語学の枠に収まらない幅広さを持っている。外国語教育研究の中で最も頻繁に扱われる言語に英語があるが、望月（2018）は英語教育研究について「言語学の一分野ではなく、独自の学問であり科学である。言語学、心理学、教育学などの知見を利用するが、いかにして英語を教えるかの立場に立ってそれらの知見を再構築する科学である」（p.12）と述べ、外国語（英語）教育研究の学際性と独自性を強調している。

また、外国語教育研究と最も関連の深い学問として応用言語学（Applied Linguistics）があり、特に日本国内では応用言語学と外国語教育研究が同義で用

---

\*筑波大学

いられることも少なくない。応用言語学における主要な研究分野の1つとして第二言語習得 (Second Language Acquisition; SLA) が含まれるのは事実であるが、応用言語学の世界最大の学会であるアメリカ応用言語学会 (American Association for Applied Linguistics) の発表分野を見ると、2018年12月現在で表1に示す21の小区分を設けている。

表1に示される通り、応用言語学の中には言語習得に関するトピックだけでなく、認知科学や脳科学に関連したもの、言語政策やイデオロギーなどの社会学と関連したもの、コーパスやテキスト分析といった情報科学に関連したものなど、多様な学問分野と関連した内容が含まれていることが分かる。これらの研究は母語を扱ったものもあるが、その多くは外国語を扱ったものである。よって、広義に外国語教育研究を捉えた場合、その包括する研究内容は言語学の一分野には収まらず、また、純粋な第二言語習得だけでも限定されず、非常に多岐にわたるといえる。

表1. アメリカ応用言語学会における研究発表の区分

---

Analysis of Discourse and Interaction
Assessment and Evaluation
Bilingual, Immersion, Heritage, and Minority Education
Corpus Linguistics
Educational Linguistics
Language Cognition and Brain Research
Language and Ideology
Language and Technology
Language Maintenance and Revitalization
Language Planning and Policy
Language, Culture, Socialization and Pragmatics
Phonology/Phonetics and Oral Communication
Reading, Writing, and Literacy
Research Methodology
Second and Foreign Language Pedagogy
Second Language Acquisition, Language Acquisition, and Attrition
Sociolinguistics
Teacher Education, Beliefs, and Identities
Text Analysis (Written Discourse)
Translation and Interpretation
Vocabulary and Lexical Studies

---

## 1.2 外国語教育研究と自然言語処理研究との関わり

前節で述べたように、外国語教育研究は学際的な性格を持ち、多様な研究内容が行われる分野である。そのような外国語教育研究で近年特に注目を浴びているのが、自然言語処理（Natural Language Processing）の知見を活かした研究である。自然言語処理とは人間が扱う言語（自然言語）の特徴をコンピューターによって自動的に抽出する手法であり、この手法を以て様々な課題の達成を目指す。その主たる課題の1つに、人間がどのように言語を理解し、処理するのかに関する情報を、コンピュータープログラムの開発を通して収集することがある（Crossley, Varner, Kyle, & McNamara, 2014）。

外国語教育分野の主要な国際学術誌の1つである Wiley 社発行の Language Learning においても、2017年に“Experimental, Computational, and Corpus-Based Approaches to Language Learning: Evidence and Interpretation”という特集が組まれるなど、外国語教育研究において自然言語処理のようなコンピューターを用いたアプローチが注目されている。この特集に含まれている Rebuschat, Meurers, and McEnergy (2017) では、言語習得を研究するアプローチとしてこれまでは実験的な手法が主であったが、それにコンピューターやコーパスを活用したアプローチを統合することで、実験的なアプローチでは明らかになり得ないことまで解明することができると主張されている。コンピューターやコーパスを活用した研究には、大規模な学習者コーパスを自然言語処理の手法を用いて分析することで学習者の言語習得パターンを見出すものなどが多いが、その他にも様々な形で自然言語処理の知見を活かした外国語教育研究は行われている。次節では、具体的な自然言語処理の手法に触れながら、これらの研究事例をレビューしていく。

## 2. 自然言語処理の手法を活かした外国語教育研究の事例

### 2.1 潜在意味解析

潜在意味解析（latent semantic analysis; LSA）は心理学の分野で広く用いられてきた自然言語処理の手法であり、近年になって外国語教育研究の分野でも用いられるようになってきた。この手法は、大規模な言語コーパスに対して複雑な統計的処理を行うことで単語や文の意味を数学的に示すものである（Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998）。また、そのような意味に基づ

いて、単語や文の概念的な類似度を数値として算出することもできる。

LSAの最も重要な側面は、その表象する意味が複数の文脈における単語の用法 (usage) と他の単語との関係 (reference) に基づいているということである (Landauer & Dumais, 1997)。この理論的基盤となっているのが、Harris (1970) の分布仮説 (distributional hypothesis) である。この仮説では、意味的に関連している単語は同じ文脈で直接的に共起するだけでなく、同じようなトピックの別々の文脈において共通の他の単語と共起するとされる。例えば、「授業の準備には教材研究が必須である」と「生徒が主体的に思考する授業が求められている」という2つの文を個別の文脈として考えた場合、「授業」と「教材」という単語は同じ文脈で直接的に共起しているため、意味的に関連していると判断するのは容易い。一方、「教材」と「生徒」は同じ文脈には表れていないものの、我々の直感に基づけば関連した意味を持つ。これらの単語は「授業」という単語を介して2つの文脈で間接的に共起しており、LSA ではこのような文脈における単語の直接的・間接的な共起情報に基づいて単語の意味を表象する。このような意味に基づくことで、LSA では表面的な意味の関連だけでなく、より深いレベルの間接的な意味の関連まで捉えることができ、結果的に人間の直感により近い意味の関連を示すことができる。このように LSA は文脈中の単語の用法と他の単語との関係に基づいて意味を表象するため、「単語の用法に基づく学習」、つまり、文脈における使われ方に基づいて単語の意味を学ぶプロセスをシミュレーションする知識モデルとしても活用される (猪原, 2016)。

数学的な観点から述べれば、LSA はコーパスによって形成される高次元の意味空間内において単語や文などの分析対象をベクトルとして表象するものである。大まかな手順としては、元となるコーパスをもとに (1) 単語を行、文書を列とし、各セルが各文書における単語の出現回数を示す単語-文書行列を作成し、(2) 多数の文書に出現する高頻度語の影響を弱め、少数の文書にのみ出現する低頻度語の影響を強めるような重みづけを行い、(3) 特異値分解 (因子分析に似た手法)、(4) 次元の縮約、という4つのステップを経て意味の表象が行われる (Günther, Dudschig, & Kaup, 2015; Günther, & Kaup, 2016)。実際の計算式など、LSA のより詳細な手続きの説明については、猪原 (2016) や Landauer and Dumais (1997) を参照されたい。

LSA では得られた単語の意味ベクトルをもとに、それらの間のコサインが算出

され、その値が単語間の意味的な関連度とされる。関連度の値は理論的には-1.0から+1.0を取るが、実質は0に近いほど無関連であり、1.0に近づくほど概念間の意味的な関連度が高いと解釈される。たとえば、LSA で最もよく用いられるTASA (Touchstone Applied Science Associates, Inc.) コーパスをベースに、意味的に関連していると思われる“education”と“schools”の2つ単語のコサインを算出した場合、その値は0.80であるのに対し、意味的にはあまり関連していないと思われる“education”と“lions”については0.03と算出される。また、表2は“education”をターゲットとして、それと意味的関連度が高い上位10語をTASA コーパスから抽出したものである。この例では、ターゲットが現れる文脈と同じ、もしくは類似した内容の文脈で共起し易いと考えられる語が高い関連度を持つと評価されていることが分かるだろう。

また、LSA では単語だけでなく、文や段落など様々な単位を分析対象にできる。単語どうしの意味的関連性はその類義関係などから判断しやすいが、文以上の単位で意味的な関連性を客観的に判断するのは難しい。しかし、LSA では単語だけでなく文や段落、文章のような相対的に大きな単位での意味的関連度の算出が可能であり、この場合、文や段落を構成する個々の単語のベクトルの合計や平均をとることが典型的に行われる (Foltz, Kintsch, & Landauer, 1998; Wolfe, Magliano, & Larsen 2005)。表3は、Leo Lionni が1963年に発表した絵本 *Swimmy* の冒頭数文について、隣接する文間の意味的関連度を算出したものである。この例では、文2と文3の意味的な関連度が高く評価されていることが分かる。

表2. LSA によって“education”と意味的関連度が高いと判断された単語

意味的関連度の値	語
0.84	educational
0.80	schools
0.78	educators
0.78	schooling
0.77	vocational
0.67	colleges
0.67	educating
0.67	esea
0.66	curricular

表 3. LSA による文間の意味的関連度

意味的関連度	文
.13	1: A happy school of little fish lived in a corner of the sea somewhere.
.46	2: They were all red.
.15	3: Only one of them was as black as a mussel shell.
.10	4: He swam faster than his brothers and sisters.
	5: His name was Swimmy.

注. テキストは *Swimmy* (Leo Lionni, 1963) の冒頭部分から抜粋した。

ここまで、LSA による単語や文間の意味的な関連度の算出や、意味的に関連した単語のコーパスからの抽出を実例とともに説明してきた。これらの LSA の機能は、ウェブベースのアプリケーション (Dennis, 2007) や R などの統計ソフトのパッケージを用いて実行することができる (Günther et al., 2015)。

外国語教育研究における LSA の活用事例の 1 つとして、学習者の発話データを LSA の観点から分析したものがある。Crossley, Salsbury, and McNamara (2010) は、米国に滞在している英語学習者の発話データを長期的に収集し、学習者が産出した発話と発話間の意味的な関連度を LSA によって分析した。その結果、学習者が時間経過とともに英語に習熟するにつれて、産出される発話間の意味的関連度が高まることが示された。これはつまり、英語の習熟度が高まるにつれて、学習者は前後で意味的に関連した内容を発話するようになった、言い換えれば、意味的に一貫した発話をするようになったということである。このことは、目標言語への習熟の過程で、学習者の語彙知識における単語の結びつきが LSA で表象されるような深いレベルのものに変わっていたことを示唆する (cf. Crossley, Subtirelu, & Salsbury, 2013)。

また、LSA を活用して外国語教材におけるテキスト分析を行った研究も多く存在する。Crossley, Louwerse, McCarthy, and McNamara (2007) は初級英語学習者向けの教材に含まれるテキストのうち、学習者向けに言語が簡素化されたテキストと、学習者向けの言語の調整がされていない真正性の高い (オーセンティックな) テキストを対象として、LSA による分析を行った。これらのテキストにお

ける隣接する文の意味的関連度を算出したところ、簡素化されたテキストよりもオーセンティックなテキストにおいてその関連度が低かったことを報告している。また、Crossley and McNamara (2008) は中級学習者向けのテキストを分析し、同様の結果を得ている。さらに、Crossley, David, and McNamara (2012) では初級、中級、上級学習者向けの簡素化されたテキストについて、レベルが上がるごとに隣接する文間の意味的関連度が低くなったことを報告している。これらの研究では一貫して、難易度が高いと想定されるテキストほど隣接する文間の意味的な関連度が低くなることが示されており、英文の読みやすさには単語の用法に基づく文間の意味的な関連度が影響していることが示唆される。

上記の事例のように、LSA を外国語教育研究に活用することで「文脈における単語の用法に基づく意味」という新たな観点から、学習者の言語習得や語彙知識の発達、教材となるテキストの特徴などを明らかにすることが可能になる。

## 2.2 多様な言語的特徴の評価

上記で述べた LSA は、いわば単語や文の関連度というテキストや発話の言語的特徴の一側面のみを捉えるものである。自然言語処理では、その他にも様々な手法を駆使して多様な言語特徴量を抽出することが試みられてきた。複数の言語特徴量を自動的に抽出するコンピューターツールの代表的なものに、ウェブベースのソフトウェアである Coh-Metrix がある。このツールでは、様々なデータベースや自然言語処理の手法を活用して、文章に含まれる言語の難しさや理解し易さに関する多数の指標を算出する (Graesser, McNamara, Louwerse, & Cai, 2004; McNamara, Graesser, McCarthy, & Cai, 2014)。これらの指標にはテキストの基本的な情報である記述統計 (e.g., 総語数, 単語や文の長さに基づくリーダビリティ) だけでなく、語彙情報 (e.g., 頻度, 親密度, 心像性), 語彙の多様性 (e.g., タイプトークン比), 文の統語特徴 (e.g., 名詞句における修飾語数の平均), 文と文のつながり (結束性; e.g., 文間の名詞の重複) などに関連した指標も含まれる。従来は総語数や文の長さなどテキストの表面的な特徴のみに基づいてテキストの難易度が評価されていたところを, Coh-Metrix では語彙や文構造, 結束性などの観点から多面的にテキスト難易度を評価することを目的としている。図 1 は実際に Coh-Metrix によって英文を解析した際の画面である。画面右側に様々な指標が算出されていることが分かる。なお, 2018年12月現在, Coh-Metrix では英

Enter your input

What is Coh-Metrix?  
Coh-Metrix is a system for computing computational cohesion and coherence metrics for written and spoken texts. Coh-Metrix allows readers, writers, educators, and researchers to instantly gauge the difficulty of written text for the target audience.

What is cohesion?  
"Our definition of cohesion consists of characteristics of the explicit text that play some role in helping the reader mentally connect ideas in the text" (Graesser, McNamara, & Louwerse, 2003).



Type text in the image

Save Data

Number	Label	Label V2.x	Text	Full description
<b>Descriptive</b>				
1	DESPC	READNP	6	Paragraph count, number of paragraphs
2	DESSC	READNS	9	Sentence count, number of sentences
3	DESWC	READNW	117	Word count, number of words
4	DESP	READAPL	1.5	Paragraph length, number of sentences in a paragraph, mean
5	DESPld	n/a	0.837	Paragraph length, number of sentences in a paragraph, standard deviation
6	DESSL	READASL	13	Sentence length, number of words, mean
7	DESSLd	n/a	8.676	Sentence length, number of words, standard deviation
8	DESWLsy	READASW	1.761	Word length, number of syllables, mean
9	DESWLsyd	n/a	1.064	Word length, number of syllables, standard deviation
10	DESWLl	n/a	5.410	Word length, number of letters, mean
11	DESWLld	n/a	3.091	Word length, number of letters, standard deviation
<b>Text Easeability Principle Component Scores</b>				
12	PCNArz	n/a	-0.447	Text Easeability PC Narrativity, z score
13	PCNArp	n/a	33	Text Easeability PC Narrativity, percentile
14	PCSNz	n/a	0.871	Text Easeability PC Syntactic simplicity, z score
15	PCSNp	n/a	80.780	Text Easeability PC Syntactic simplicity, percentile
16	PCCNz	n/a	-0.191	Text Easeability PC Word concreteness, z score
17	PCCNp	n/a	42.470	Text Easeability PC Word concreteness, percentile
18	PCRFz	n/a	-0.846	Text Easeability PC Referential cohesion, z score
19	PCRFp	n/a	20.050	Text Easeability PC Referential cohesion, percentile
20	PCDCz	n/a	1.214	Text Easeability PC Deep cohesion, z score
21	PCDCp	n/a	88.690	Text Easeability PC Deep cohesion, percentile
22	PCVERBz	n/a	-1.792	Text Easeability PC Verb cohesion, z score
23	PCVERBp	n/a	3.670	Text Easeability PC Verb cohesion, percentile
24	PCCONnz	n/a	-1.571	Text Easeability PC Connectivity, z score
25	PCCONnp	n/a	5.820	Text Easeability PC Connectivity, percentile
26	PCTEMPz	n/a	0.045	Text Easeability PC Temporality, z score
27	PCTEMPp	n/a	51.600	Text Easeability PC Temporality, percentile
<b>Referential Cohesion</b>				
28	CRFNO1	CRFBN1um	0.5	Noun overlap, adjacent sentences, binary, mean
29	CRFAO1	CRFBA1um	0.5	Argument overlap, adjacent sentences, binary, mean
30	CRFSO1	CRFBS1um	0.5	Stem overlap, adjacent sentences, binary, mean
31	CRFNOa	CRFBNaum	0.333	Noun overlap, all sentences, binary, mean
32	CRFAOa	CRFBaum	0.361	Argument overlap, all sentences, binary, mean
33	CRFSOa	CRFBsaum	0.333	Stem overlap, all sentences, binary, mean
34	CRFCWO1	CRFPC1um	0.071	Content word overlap, adjacent sentences, proportional, mean
35	CRFCWO1d	n/a	0.079	Content word overlap, adjacent sentences, proportional, standard deviation
36	CRFCWOa	CRFPCAum	0.061	Content word overlap, all sentences, proportional, mean
37	CRFCWOad	n/a	0.094	Content word overlap, all sentences, proportional, standard deviation
<b>LSA</b>				
38	LSASS1	LSAaas1	0.073	LSA overlap, adjacent sentences, mean
39	LSASS1d	LSAassd1	0.153	LSA overlap, adjacent sentences, standard deviation
40	LSASSp	LSAasp	0.187	LSA overlap, all sentences in paragraph, mean
41	LSASSpd	LSAaspd	0.131	LSA overlap, all sentences in paragraph, standard deviation
42	LSAPP1	LSAppa	0.083	LSA overlap, adjacent paragraphs, mean
43	LSAPP1d	LSAppd	0.122	LSA overlap, adjacent paragraphs, standard deviation
44	LSAGN	LSAGN	0.179	LSA given/new, sentences, mean

図1. Coh-Metrix による英文解析結果

文に加えて中国語の文章を解析するツールも提供されている。

以下では、Coh-Metrix を活用した外国語教育研究である Crossley et al. (2012) で用いられていた指標を例として、Coh-Metrix で算出される多様な言語的特徴の指標を説明する。

## 1 Number of words

テキストに含まれる語数の合計である。Coh-Metrix が算出する総語数は、



構文解析ツールによって特定された構文木（8 統語的類似性を参照）と呼ばれる各文の構造に基づいている。

## 2 Lexical diversity (vocd)

テキストにおいてどの程度多様な語彙が用いられているかを表す指標である。語彙的な多様性が高い場合、内容を理解しながら新たな単語を文脈に統合する機会が増えるため、テキストの理解が難しくなり得る。*vocd* は語彙的多様性の指標として一般的に用いられるタイプトークン比（延べ語数に対する異なり語数の比率）を統計処理によって修正し、テキストの長さ（総語数）の影響を考慮できるようにした指標である。この値が高い場合は、テキストで異なり語が多く用いられている（語彙的多様性が高い）ことを示す。

## 3 CELEX Word frequency (content words)

分析対象のテキストに含まれる語彙（内容語）の頻度レベルを示す指標である。頻度レベルが高い語（高頻度語）は親密度が高く、処理も素早く行われるため、テキストの理解しやすさに貢献する。本指標は CELEX 語彙データベースに基づいており、これは COBUILD コーパスに基づくデータベースで約1790万語を含むものである。

## 4 Word familiarity (content words)

分析対象のテキストに含まれる語彙（内容語）の親密度の指標である。この指標は MRC 心理言語学データベースに基づいており、このデータベースでは親密度や具象性など英語語彙の様々な認知的側面について、大人の協力者による7段階の評定値を提供している（値は実際の評価値を100倍した100から700の間をとる）。親密度については、協力者が7段階で評価した9,392語に対する親密度が報告されている。たとえば、“water”（641）は“otter”（391）よりも親密度が高いとされる。

## 5 Word meaningfulness (all words)

分析対象のテキストに含まれる語彙（内容語）の意味性の指標であり、これも MRC 心理言語学データベースが提供する値に基づいている。意味性は

2,627語について報告されており、人間の評価に基づいて算出された、ある単語が持つ連想語の数である。つまり、その値が高い語は他の語と強く結びついており、低い語は他の語との結びつきが弱いと考えられる。たとえば、“people”の意味性は612であるのに対し、“abbess”は218である。

## 6 Polysemy

分析対象のテキストに含まれる語彙の多義性を表す指標で、WordNet と呼ばれるデータベースに基づいている。Coh-Metrix では分析されるテキストの内容語のうち、多義語がどのくらい含まれているかを算出する。この値が高ければ、テキストの意味が曖昧であることを示す。しかしながら、多くの高頻度語は多義語であるため、この値は高頻度語の割合を示すことにもなり得る。

## 7 Verb hypernymy

分析対象のテキストに含まれる動詞の特定性に関する指標で、WordNet に基づいている。WordNet では単語が上位・下位の階層構造にまとめられており、たとえば名詞の例であるが、“chair”は“camp chair”→“chair”→“seat”→“furniture”→“artifact”→“object”→“entity”という階層構造を持つ。単語が持つ特定性の値は階層構造の中でその語よりもいくつの上位語があるかを表し、その値が高いほど単語の特定性が高く、値が低いほど特定性が低いということになる。

## 8 Syntactic complexity

文の統語構造の複雑さの指標であり、Crossley et al. (2012) では主動詞（文の主節の動詞）に先行する単語数の平均が扱われている。たとえば、*They all enjoyed the dinner.* であれば主動詞前の語数は2語、*When he was a child, he had a dog* であれば6語となる。主動詞前の語数が少ないほど統語構造が単純であり、文が処理し易いと想定される。

## 9 Sentence syntax similarity

文間の統語的な類似性を測定する指標で、テキスト内での統語構造の単一性

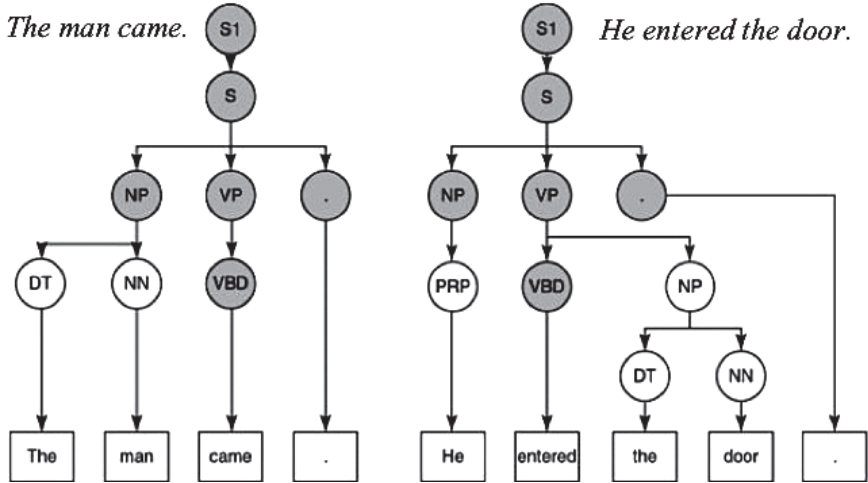


図 2. 文間の統語的類似性の算出 (McNamara et al., 2014, p.71 を一部改変)

や一貫性を示す。この指標の算出にあたっては、まず2つの文を構文木で示し、文間で共通しているノードの数をカウントする。そして、2つの構文木が持つ全ノード数から共通しているノード数を引いた値で共通しているノード数を割る。図2を例にすると、2つの文を合計した全ノードは18であり、灰色で示された2つの文で共通のノードが6つあるため、統語的類似性は  $6 / (18 - 6) = 0.5$  と算出される。テキストの統語的類似性が高ければ、単純な統語構造を持つ文が繰り返し用いられていることが示唆され、統語的な複雑さが低いと判断される。

## 10 LSA adjacent sentences

テキストの意味的な結束性を示す指標であり、先述の潜在意味解析 (LSA) に基づいて2つの文がどの程度意味的に類似しているのかを量的に評価する。たとえば、“house”を含む文と“table”を含む文は互いに意味的に関連していると判断され、このような文間の意味的な関連度が高いほど、テキストは一貫した内容やトピックを記述しており、その理解が容易となり得る。

## 11 Motional prepositions

“through” や “into” など移動を表す前置詞の数を示す指標である。この指標はテキストの空間性 (spatiality) を表す指標であり、空間性はテキストの豊かな理解の形とされる状況モデルを構成する複数の次元の1つである (Zwaan, Langston, & Graesser, 1995)。よって、空間性の高いテキストは状況モデルが構築しやすいテキストとなる。

## 12 Noun overlap

テキストの照応的な結束性を示す指標であり、テキスト内の文のうち同じ名詞を含む文の割合を算出する。文間で同じ名詞を共有している割合が高いほどテキストの照応的な結束性が高いと判断され、照応的な結束性は読み手・聞き手が文と文のつながりを構築することに貢献し、その結束性が低い場合は理解を阻害することもある。

## 13 Logical operators (negative connectives)

テキストの論理的な結束性に関わる指標であり、1000語あたりにおける接続詞の出現頻度を表す。接続詞のタイプごとに算出が可能であり、Crossley et al. (2012) では否定 (逆接) を表す接続詞 (“however” “but”) に関するものが算出されている。

## 14 Causal cohesion

テキストの因果的な結束性の指標であり、因果的な接続を表す語句 (e.g., “because”, “to”, “by”, “as a result”, “consequence of”) と因果的な動詞 (状況の変化を示す動詞; e.g., “break”, “freeze”, “impact”, “hit”, “move”, “kill”) の比率を表す。因果的な接続を表す語が少なく因果的な動詞が多い場合は、テキストに記載された出来事の因果関係が明示的に記述されていないことが多いということであり、読み手は因果関係を自身で推論する必要性が高くなる。このような場合、文章の因果的な結束性は低いということになる。

上記の1はテキストの記述統計 (表層的な特徴) の指標, 2~7が語彙の特徴, 8~9が文の統合構造, 10~14がテキストの結束性に関わる指標というように分

類される。

さらに、Graesser, McNamara, and Kuilkowich (2011) は大規模なコーパスデータをもとに、Coh-Metrix で算出される54の指標を以下の8つの成分へと縮約している：Narrativity, Referential cohesion, Syntactic simplicity, Word concreteness, Causal cohesion, Verb cohesion, Logical cohesion, Temporal cohesion. これら8つは、上記の語彙や文、結束性に関する個別の指標とそれらに関連した別の複数の指標を主成分分析によって統合したものである。これら8つの主成分は文章理解の5つのレベル（語彙、統語、テキストベース、状況モデル、ジャンル）を想定する理論モデルとも整合性が高い（Graesser & McNamara, 2011）。

外国語教育研究の分野でも、この Coh-Metrix を活用した研究が近年多く報告されている。様々な研究課題について Coh-Metrix は活用されてきたが、最も多く活用されてきたのが外国語教材に含まれるテキストの分析である。Crossley et al. (2007) では英語学習者用の初級レベルのテキストについて、簡素化されたテキストはオーセンティックなテキストと比較して含まれている語彙の頻度レベルが高いことや、文間の名詞の重複が多いことなどが明らかにされている。中級レベルのテキストを分析対象として同様の検証を行った Crossley and McNamara (2008) でも、概ね同じ結果が得られている。さらに Crossley et al. (2012) では、初級・中級・上級の簡素化されたテキストについて、レベルの低いテキストほど語彙の頻度レベルや文間の意味的な関連度が高く、語彙的多様性や統語的な複雑さが低くなることなどが明らかにされている。また、Plakans and Bilki (2016) では特にテキストの結束性に焦点を当てた分析が行われており、中上級レベルよりも初級レベルのテキストで結束性が高くなる傾向が示されている。しかしながら、この研究ではリーダビリティや統語的な複雑さなどの指標と比べると、結束性のレベル間の差は小さいことが指摘されている。

### 3. 国内における研究事例

前節では具体的な自然言語処理の手法やツールの説明とともに、それらが外国語教育研究でどのように活用されてきたかを述べた。本節では、日本人の英語学習者や日本人向けの英語教材を対象とした国内の研究事例を、筆者の過去の研究を中心に述べていく。

### 3.1 潜在意味解析を用いた研究事例

国内における外国語教育研究では、名畑目（2012）が先駆的に潜在意味解析（LSA）を用いている。英検の読解問題のうち空所補充型の問題に使われるテキストを対象として、空所を含む文とその前後の文の意味的な関連度を潜在意味解析によって算出した。その結果、上位の級になるほど空所を含む文とその前後の文との意味的関連度が低くなることが明らかにされた。難易度が高いと想定される文章ほど隣接する文間の意味的関連度が低くなるという傾向は、Crossley らの研究結果と一貫している。

また、名畑目（2018）では日本人の小学生向けに作成された英語ストーリー教材を対象として、各教材における隣接する文間の意味的関連度を算出した。その結果、大人の学習者向けの教材を主な分析対象とした Crossley らの研究で報告されている値よりも、全体として高い値が得られた。よって、この研究の結果も、難易度が低いと想定される文章ほど文間の意味的な関連度が高いという先行研究の知見を支持するものである。さらにこの研究では、算出された文間の意味的な関連度に基づいて個別の教材の文脈の記述内容を考察し、文間の意味的な関連度に基づけば文脈の理解し易さの観点からストーリー教材を評価できることや、より意味的に関連した文脈を持つように教材を作成・改善できる可能性があることを示唆した。

上記の研究は LSA によって英語学習者用のテキストを分析したものであるが、Nahatame（2017, 2018）では LSA によって得られる文間の意味的な関連度が実際の英語学習者の文章理解に与える影響を検討している。これらの研究では、日本人大学生に表 4 のような英文を提示して読解させた。この英文は 2 文 1 組であり、1 文目が 2 文目との因果的・意味的な関連の程度によって a～d の 4 つのパターンがある。因果的な関連度は 1 文目に述べられた出来事と 2 文目に述べられた出来事の因果関係の強さを表すものであり、英語教育を専攻する大学院生による評定値に基づいている。意味的な関連度は、LSA による 2 つの文のコサインの値に基づくものである。

Nahatame（2017）では、これら 2 文 1 組の英文について書かれた内容がどの程度一貫しているかを協力者に 5 段階で評定させた。その結果、LSA による意味的な関連度が高い英文ほど内容が一貫していると評定されやすく、それは特に文の因果的な関連度が低い場合や協力者の英文読解力が低い場合で顕著だった。この

表 4. Nahatame (2017, 2018) で用いられた英文

1a. Mary could not find anything to read in the library.	(CR-High / SR-High)
1b. Mary wanted to look for recipes for her dinner party.	(CR-High / SR-Low)
1c. Mary went to the library to look for something to read.	(CR-Low / SR-High)
1d. Mary was having a dinner party for her office.	(CR-Low / SR-Low)
2. She went to the bookstore to get new books.	

*Note.* CR = Causal Relatedness; SR = Semantic Relatedness.

結果は、LSA のように統計的に算出される英文の意味的な関連度が、英語学習者による文章の一貫性の知覚と関連していることを示した点で重要である。

Nahatame (2018) では、同様の文章を用いて協力者の読解時間と読解後の記憶テスト（筆記再生課題）の分析を行った。線型混合モデル（Linear Mixed Model）による分析を行った結果、文間の意味的な関連度が高い文章ほど学習者の記憶に残り易いことが明らかにされた。また、読解時間に対しても意味的な関連度の促進効果は見られるものの、因果的な関連度のほうが影響が大きく、因果的関連度が高い場合にのみ意味的な関連度の効果が顕著であった。これらの結果は、LSA によって算出される英文の意味的な関連度が、英語学習者の一貫性の知覚だけでなく、テキスト内容の記憶、及び限定的ではあるがテキストの処理時間とも関連していることを示している。

Nahatame (2017, 2018) は、LSA によって算出される文の意味的な関連度と英語学習者の行動データとの関わりを実証した点で重要である。言い換えれば、これらの研究結果は、LSA による文の意味的な関連度の心理学的妥当性を外国語読解の文脈で示したものである。

### 3.2 Coh-Metrix を用いた事例

Coh-Metrix を用いた研究は国内でもいくらか行われてきたが、それらの多くは国外の研究と同様に英語学習教材を分析したものである。たとえば Coh-Metrix を用いた先駆的な研究である水本 (2013) は、高等学校の英語教科書 1 冊を分析対象とし、Graesser et al. (2011) が特定した 8 つの主成分に関する得点を全 15 レッスンに含まれるテキストに対して算出した。さらにその他の分析ツールから指標を追加して、これらの指標の持つ情報量を主成分分析により圧縮して 2

つの変数で示し、その変数に基づいてクラスター分析を行うことで、各レッスンのテキストを難易度（読みやすさ、具体性、文体）の観点から分類している。

名畑目・木村（印刷中）は、水本（2013）と同様のアプローチで日本人の小学生向けに作成された英語ストーリー教材を分析した。文部科学省が作成した英語ストーリー教材を含む12の教材を対象として、Graesser et al.（2011）が特定した8つの主成分に関する得点、及び総語数などその他の指標をCoh-Metrixによって算出した。そしてそれらを主成分分析によって縮約し、その主成分に基づくクラスター分析を行って教材を複数のクラスターに分類した。このような教材の言語的特徴の分析とその特徴に基づく分類を行うことで、同じ小学生向けの英語教材であっても言語的特徴の観点からその相違点を知ることができる。たとえば名畑目・木村（印刷中）の主成分分析の結果では、第一主成分として総語数や平均的な文の長さといった表面的な特徴に基づくテキストの難易度、因果・論理関係の記述量、語彙の具体性を統合した指標が得られた。この指標に基づけば、これらの言語的特徴に基づいて教材の難易度を把握したり、指導の工夫を考えることが可能となる。たとえば、表面的な難易度が低く、因果・論理関係の記述量が少ない教材は、英語に慣れていない児童に対する入門的なストーリー教材として適切である。一方で、因果・論理関係の記述が少ないため、出来事の因果関係の理解のためにイラストやジェスチャー、発問を活用するといった工夫が必要になるかもしれない。反対に、表面的な難易度が高く、因果・論理関係の記述が多い教材は、その活用の際に児童の外国語に対する習熟度を把握・考慮する必要がある。内容理解が難しいような場面は単語や文を繰り返す、短い文で言い換える、前の内容を振り返らせるなどの理解への足場かけが必要となるだろう。一方、このような教材は、因果・論理関係を表す外国語の語句や表現に児童が文脈の中で繰り返し触れることで、それらの語句や表現を習得することを可能にする教材でもある。これらのような教材の言語的特徴を把握することで、指導目的や学習者の習熟度に応じた教材の選定や活用が可能になる。さらに、このような知見は、指導目的や学習者の習熟度に応じた教材開発にも示唆を与えるものである。



#### 4. 今後の展望

自然言語処理の知見を活かした外国語教育研究は、近年国内外で多く発表されている。本稿で扱った研究はあくまでそれらのほんの一部であり、本稿で扱った以外にも自然言語処理の知見を活かした外国語教育研究は多く存在する。そして、外国語教育研究においてそのような研究はこれからますます増加すると予想される。自然言語処理の技術は日進月歩であるため、様々なツールや手法が日々開発されており、今後はこれまでよりも多様な形で外国語教育研究における自然言語処理の手法の活用が見られるだろう。たとえば、本稿では代表的なテキスト分析のツールとして Coh-Metrix を取り上げたが、近年その限界点を改善したツールが発表されている。Coh-Metrix では1つずつのテキストに対して分析を行う必要があったが、数百のテキストを分析するような際には膨大な時間が掛かる。また、Coh-Metrix には含まれていない、より新たな言語的特徴の指標も近年開発・提案されている。これらの限界点を踏まえて開発されたのが、TAALES (The Tool for the Automatic Analysis of Lexical Sophistication) や TAACO (The Tool for the Automatic Analysis of Cohesion) である (Crossley, Kyle, & Dascalu, in press; Kyle, Crossley, & Berger, 2018)。これらのツールではコンピューターにソフトウェアをインストールすることで、複数のテキストを同時に分析することが可能である。また、TAALES は多数の語彙指標を算出するツールであり、Coh-Metrix に含まれていない n-gram (n 語の組み合わせ) の頻度情報や Academic word list に基づく指標なども算出することができる。TAACO は多数の結束性に関する指標を算出するツールであり、LSAよりもさらに洗練した意味的関連度の指標である Latent Dirichlet Allocation (LDA) や Word2vec を使って文間の意味的な関連度を算出することができる。本稿執筆時点においてはまだ Coh-Metrix ほどには外国語教育研究には浸透していないが、今後確実にこれらのツールを使用した研究は増えていくだろう。

また、上記のようなツールは、英語学習者のスピーキングやライティングにおける習熟度の予測を通して、算出される指標の妥当性が検証されている。たとえば、ある課題に対する英語学習者の作文データを収集し、専門的な評価者がそれに対して包括的な評価、あるいは語彙や結束性といった個別の言語特徴に関する評価を与える。同時に、その作文データに対して上記のようなツールを用いて様々な言語的特徴の指標を算出する。そして、その指標が人間による評定値をど

の程度予測するかを検証するという流れである。このようなアプローチや研究は、最終的にはスピーキングやライティングの自動採点に向けた取り組みへとつながっていく。実際に国内においても、言語特徴量に基づいて英語学習者の発話データの自動採点を行う試みはすでになされている（近藤・石井, 2017）。

さらに、近年では Coh-Metrix 及びその関連ツールを統合・発展させた *Reader-bench* が開発され、このツールではテキストの言語的な複雑さの分析や学習者が作成した要約文の自動採点などを行うことができる（Dascalu, Dessus, Trausan-Matu, Bianco, & Nardy, 2013）。また、このツールの機能の1つとして Automated model of comprehension (AMoC) があり、この機能では文の統語的な情報などに基づくテキストの字義的な理解と、読み手の推論に基づくテキストのより深い理解がどのように構築されていくかをシミュレーションできる（Dascalu, Paraschiu, McNamara, & Trausan-Matu, 2018）。今後改善の余地はあるが、このような機能を使えば人間の文章理解の認知メカニズムをより明確に示すことができる可能性を秘めている。

これらのような自然言語処理の手法に基づく言語解析ツールがさらに開発・普及していくことで、より多様な自然言語処理の手法や知見を取り入れた外国語教育研究が行われていくだろう。そして、そのような研究が増加・発展していけば、外国語教育研究、そして外国語教育自体の在り方も大きく変わっていくだろう。

## 引用文献

- Crossley, S. A., Allen, D., & McNamara, D. S. (2012). Text simplification and comprehensible input: A case for an intuitive approach. *Language Teaching Research, 16*, 89–108.
- Crossley, S. A., Kyle, K., & Dascalu, M. (in press). The tool for the automatic analysis of cohesion 2.0: Integrating semantic similarity and text overlap. *Behavioral Research Methods*.
- Crossley, S. A., Louwse, M., McCarthy, P. M., & McNamara, D. S. (2007). A linguistic analysis of simplified and authentic texts. *Modern Language Journal, 91*, 15–30.
- Crossley, S. A., & McNamara, D. S. (2008). Assessing second language reading texts at the intermediate level: An approximate replication of Crossley, Louwse, McCarthy, and McNamara (2007). *Language Teaching, 41*, 229–409.
- Crossley, S. A., Salsbury, T. & McNamara, D. S. (2010). The development of semantic relations in second language speakers: A case for latent semantic analysis. *Vigo International Journal of Applied Linguistics, 7*, 55–74.

- Crossley, S., Subtirelu, N., & Salsbury, T. (2013). Frequency effects or context effects in second language word learning: What predicts early lexical production? *Studies in Second Language Acquisition*, 35, 727–755.
- Crossley, S. A., Varner, L. K., Kyle, K., & McNamara, D.S. (2014). Analyzing discourse processing using a simple natural language processing tool (SiNLP). *Discourse Processes*, 51, 511–534.
- Dascalu, M., Dessus, P., Trausan-Matu, Ș., Bianco, M., & Nardy, A. (2013, July). ReaderBench, an environment for analyzing text complexity and reading strategies. In *International Conference on Artificial Intelligence in Education* (pp. 379–388). Springer, Berlin, Heidelberg.
- Dascalu, M., Paraschiv, I. C., McNamara, D. S., & Trausan-Matu, S. (2018, September). Towards an Automated Model of Comprehension (AMoC). In *European Conference on Technology Enhanced Learning* (pp. 427–436). Springer, Cham.
- Dennis, S. (2007). How to use the LSA web site. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis* (pp. 57–70). Mahwah, NJ: Lawrence Erlbaum Associates.
- Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis *Discourse Processes*, 25, 285–307.
- Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, 3, 371–398.
- Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, 40, 223–234.
- Graesser, A., McNamara, D., Louwerse, M., & Cai, Z. (2004). *Coh-Metrix: Analysis of text on cohesion and language. Behavioral Research Methods, Instruments, and Computers*, 36, 193–202.
- Günther, F., Dudschig, C. & Kaup, B. (2016). Latent semantic analysis cosines as a cognitive similarity measure: Evidence from priming studies. *Quarterly Journal of Experimental Psychology*, 69, 626–653.
- Günther, F., & Kaup, B. (2015). LSAfun - An R package for computations based on latent semantic analysis. *Behavior Research Method*, 47, 930–944
- Harris, Z. (1970). Distributional structure. In *Papers in Structural and Transformational Linguistics* (pp. 775–794). Dordrecht, Holland: D. Reidel.
- 猪原敬介 (2016). 『読書と言語能力：言葉の「用法」がもたらす学習効果』 京都大学学術出版会.
- 近藤悠介・石井雄隆 (2017). 「英語学習者の発話自動採点システムの開発と英語教育プログラムへの導入可能性の検討」『Language Education & Technology』第54号, 23–40.
- Kyle, K., Crossley, S. A., & Berger, C. (2018). The tool for the analysis of lexical

- sophistication (TAALES): Version 2.0. *Behavior Research Methods*, 50, 1030–1046
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104, 211–420.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25, 259–284.
- McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. New York, NY: Cambridge University Press.
- 水本篤 (2013). 「英文解析プログラムから得られる各種指標を使ったテキスト難易度の推定—教材作成への適用可能性—」『外国語教育メディア学会 (LET) 関西支部 メソドロジー研究部会 報告論集』第3号, 141–150.
- 望月昭彦 (2018). 「第1章 英語教育と英語教育学」望月昭彦 (編著) 『新学習指導要領にもとづく英語科教育法 第3版』(pp. 3–13) 大修館書店.
- 名畑目真吾 (2012). 「Latent Semantic Analysis (LSA) による空所補充型読解テストの解明：文レベルの意味的関連度を観点として」. *STEP (the Society for Testing English Proficiency) Bulletin*, 24, 42–58.
- Nahatame, S. (2017). Standards of coherence in second language reading: Sentence connectivity and reading proficiency. *Reading in a Foreign Language*, 29, 86–112.
- Nahatame, S. (2018). Comprehension and processing of paired sentences in second language reading: A comparison of causal and semantic relatedness. *Modern Language Journal*, 102, 392–415.
- 名畑目真吾 (2018). 「小学生向けストーリー教材の文脈の分析：文間の意味的な関連度に基づいて」『小学校英語教育学会誌 (JES Journal)』第18号, 84–99.
- 名畑目真吾・木村雪乃 (印刷中). 「言語的特徴に基づく小学生向けストーリー教材の分類：英文解析プログラムを用いた多面的な分析」『小学校英語教育学会誌 (JES Journal)』第19号.
- Plakans, L. & Bilki, Z. (2015). Cohesion features in ESL reading: Comparing beginning, intermediate and advanced textbooks. *Reading in a Foreign Language*, 28, 79–100.
- Rebuschat, P., Meurers, D., & McEnery, T. (2017). Language learning research at the intersection of experimental, computational, and corpus-based approaches. *Language Learning*, 67, 6–13.
- Wolfe, M. B. W., Magliano, J. P., & Larsen, B. (2005). Causal and semantic relatedness in discourse understanding and representation. *Discourse Processes*, 39, 165–187.
- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological science*, 6, 292–297.