

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 第 16 号 報告論集  
(2024 年 3 月)

---

01. コーパス研究の思考法—Sketch Engine を用いたデータの抽出・可視化—  
神原 一帆 (立命館大学) pp. 1-33
02. 我が国の motivational strategies の研究動向の概観：CiNii データベースを用いた  
システムティックレビュー  
田中 博晃 (近畿大学) pp. 34-50
03. Methods of Research on Depth of Processing for Written Corrective  
Feedback: A Scoping Review Purposes  
Yoko Fujisawa, Atsushi Doi, & Natsuko Shintani (Kansai University)  
pp. 51-68
- 

発行	2024 年 3 月 31 日
編集・発行	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
代表	浦野 研 〒062-8605 北海道札幌市豊平区旭町 4-1-40 北海学園大学 経営学部 Tel: 011-841-1161 E-mail: urano@hgu.jp

# コーパス研究の思考法

## —Sketch Engine を用いたデータの抽出・可視化—

神原 一帆

立命館大学（言語教育センター, R-GIRO）

---

### 概要

本稿はコーパスを利用した研究におけるデータの取得法, 処理法, そして, その簡易な可視化の手順について論じるものである。本来これらの処理には高度なプログラミング技術が必須であるが, 高性能なコーパスインターフェイスである Sketch Engine を利用することである程度のタスクを遂行することができる。本稿ではコーパス研究における基礎知識を概観した上で, 複雑なデータ抽出を可能にする正規表現の利用法, そして Sketch Engine の活用法のそれぞれを, 具体例と共に提示する。

**Keywords:** コーパス研究, 正規表現, Sketch Engine, メソドロジー

---

### 1. はじめに

理論言語学において, コーパスのデータを利用した実証研究の重要性は繰り返し論じられてきましたが (Fillmore 1990, Glynn & Fischer 2010, Glynn & Robinson 2014, Gries 2010), これは研究者の内省判断にもとづく研究が主流であったことが背景にあります (Stefanowitsch 2020)<sup>\*1</sup>。理論的な動機をもつものだけに限らず, 応用志向の研究であっても, それを経験的に進めるためには適切な観察手法が必須となります。

本稿の目的は, (i) コーパスを言語研究などにもちいるための基礎知識と, (ii) 強力なコーパスインターフェイスである Sketch Engine の利用法を概観することです。本稿は次のように構成されます。2. 節ではコーパスを研究にもちいる際の基本事項を確認し, その使用の際の注意点を確認します。3. 節では Sketch Engine で実行可能な様々な機能の特徴について概観します。4. 節では本稿のまとめと, 更に先に進むための文献紹介をおこ

ないます。

## 2. コーパス言語学の基礎

本節ではコーパス言語学の基礎知識として、複雑なデータ抽出を可能にするアノテーションの有用性と、データ抽出をおこなう際のいくつかの留意点を確認していきます。コーパス (corpus) とは (i) 書き言葉や話し言葉などの現実の言語を、(ii) 大規模に、(iii) 基準に沿って網羅的・代表的に収集し、(iv) コンピューター上で処理できるデータとして保存し、(v) 言語研究に使用するものと特徴付けられます (石川 2021, 13)。近年ではコーパスを利用するための様々なサービスが整備されており、簡単にデータを取得することができるようになってきました。しかし、このようなツールを適切に利用するためには、そのようなサービスが提供する機能の理解だけでなく、サービスの基本的な構造の理解も必要となります。本節では、2.1 節にてコーパスの構造について確認し、2.2 節にてコーパスを利用する際の留意点について導入します。

### 2.1 コーパスをつくる「技術」

本節では「コンピューター上で処理できるデータ」というものの形がどのようなものなのかを確認していきます。コンピューター上で利用可能な形式には様々なものがありますが、任意のデータを抽出するためには品詞などのメタ言語情報を付与したものが最も頻繁に利用されます。このようなメタ言語情報は一般的に注釈 (annotation) (または原語をカタカナにしたアノテーション) と呼ばれますが、本節ではこの情報の有用性とその利用法の二点について確認していきます。2.1.1 節ではデータ抽出の基礎として正規表現の基本を確認します。そして、2.1.2 節にてコーパスとコーパスインターフェイスの区別を説明します。

#### 2.1.1 データ抽出の基礎としての正規表現

まずはコーパスの利用にメタ言語情報が有用になることを確認するための例として、ある作家による英語で書かれた小説のテキストを一つの text ファイルに集約した原始的なコーパスを考えてみましょう。このファイルは.txt という拡張子を持ち、手持ちのテキス

トエディターによって簡単に検索をすることができるものと仮定します。ここで現代英語の定 (definiteness) に関心をもつ K さん (仮名) は不定冠詞 *a* または定冠詞 *the* の後ろにどのような名詞がくるのかを観察することを決めたとしましょう \*2。

このタスクにおいて必須となる技術が正規表現 (regular expression) です。例えば、あなたのコーパスで *woman* という名詞がどのように使われているのかを調べたいとしましょう。その時、“woman” という文字列をそのまま検索すると、(当たり前ですが) 複数形の “women” を検索結果として得ることはできません。ここで求められるのは “woman” と “women” に含まれる a と e のどちらでもマッチするような検索をすることですが、これは正規表現を使って `wom[ae]n` のように実現することができます。正規表現は難しいものの、マスターすれば非常に便利な道具として活用することができます。正規表現の詳しい使い方を説明することは本稿の目的から逸れますが、代表的な表現を浅尾・李 (2013, 259) を参考にまとめたものを表 1 に掲載します。これらの詳細な使い方については適宜 Web 検索等で補ってください \*3。

ここで本来のタスクの遂行に戻りましょう。今回のタスクの目的は不定冠詞 *a* または定冠詞 *the* の後ろにあらわれる名詞を当該のコーパスから抽出することです。まず、定冠詞の表現は簡単に `(a)|(the)` で検索することができます \*4。英語であれば語と語の間に一文字分の半角スペースが入りますが、何かのミスや元ファイルのエラーなどで二文字以上の半角スペースが入ったら欲しいデータがとれなくなってしまうので、+ を用いて `(a)|(the) +` と表現するようにしましょう。

これだけで問題は解決しません。K さんは自分のコーパスに含まれる全ての冠詞と名詞の組み合わせを観察することを目的としています。ここで K さんは元のコーパスにどんな名詞が含まれるのかを知らないものとする、(1) のように名詞の単複を選言の | によってつなげていく必要がでてきます。ここで必要になるのは現代英語における名詞の包括的なリストですが、それを列挙することは到底不可能だけでなく、テキストファイルに含まれる名詞を目視で書き留めていくのは現実的な作業とはいえません \*5。

- (1) a. `((a)|(the)) +(dogs?)`
- b. `((a)|(the)) +((dogs?)|(wom[ae]n))`
- c. `((a)|(the)) +((dogs?)|(wom[ae]n)|(m[ae]n))`
- .....



表1 主な正規表現とその機能

正規表現	機能	例
<code>?</code>	直前の文字があってもなくてもよい	<code>dogs?</code> によって“dog”や“dogs”にマッチする
<code>.</code>	任意の一文字にマッチする	<code>d.g</code> によって“dog”や“dug”, “dig”などにマッチする
<code>+</code>	直前の文字の1回以上の繰り返しにマッチする	<code>no+</code> によって“no”や“noooo”などにマッチする
<code>*</code>	直前の文字の0回以上の繰り返しにマッチする	<code>dog*</code> によって“do”や“dog”, “dogg”などにマッチする
<code>[...]</code>	[ ]の中に列挙したどれか一文字にマッチする	<code>wom[ae]n</code> によって“woman”や“women”にマッチする
<code>[^...]</code>	否定	<code>[^s]</code> によって“s”以外の一文字に, <code>[^A-Za-z]</code> によってアルファベット以外の一文字にマッチする。
<code>... ...</code>	の前後に生起する文字列の選言	<code>dog cat</code> によって“dog”や“cat”などにマッチする。
<code>(...)</code>	文字のグループ化	<code>(dogs?) (cats?)</code> によって“dog”, “dogs”, “cat”, “cats”などにマッチする。
<code>\w</code>	アルファベット, アンダーバー, 半角数字 ( <code>[a-zA-Z_0-9]</code> という記述と同じ)	<code>\w+</code> によって“dog”, “kazy3024”, “whisky”などにマッチする。

ここで便利になるのがアノテーション付きのコーパスです。アノテーションとは何らかの基準によって分割された事例に対して何らかのメタ表現を付与したものを指します (cf. Pustejovsky & Stubbs 2013, Ch.1)。このようなアノテーションの代表例としては品詞情報が挙げられます。例えば, 「名詞」のような品詞は{*dogs, hypothesis, noodle, ...*}のような表現を包摂するような一種のクラス (class) に対応します。言語学において, 品詞とい

うものは統辞的な関係 (syntagmatic relation) と範列的な関係 (paradigmatic relation) にもとづいて同定されます (黒田 2003, 守田 2013)。近年では自然言語処理 (Natural Language Processing; NLP) の発展によって品詞情報を自動で付与するプログラムが利用可能となっていて、様々なものを利用することができます。ここで一つ簡単な具体例を (2) に挙げます。(2a) は元の文, (2b) は品詞解析機<sup>\*6</sup> を適用した結果です。(2b) では (2a) に含まれるそれぞれの語に対して `<...>...</...>` という形で `nnp`, `vbd` などの品詞タグが付与されています。ここで重要なことは `nnp`, `vbd` などの品詞タグは元々のデータに含まれているものではなく, 元のデータの特徴を記述するための表現として機能しているという点です。仮に K さんが自分のデータに対して同様の解析機を適用し, その結果に対して `<det>[a-zA-Z]+</det> +<nn[a-zA-Z]*>[a-zA-Z]+</nn[a-zA-Z]*>` のような正規表現を用いた検索をすれば, 簡単に冠詞の直後に名詞が生起する事例を検索することができます。

- (2) a. Alice thought that Bill said that Charlotte believed that David was a liar  
 b.

```
<nnp>Alice</nnp> <vbd>thought</vbd> <in>that</in> <nnp>Bill</nnp> <vbd>said</vbd> <in>that</in> <nnp>Charlotte</nnp> <vbd>believed</vbd> <in>that</in> <nnp>David</nnp> <vbz>was</vbz> <det>a</det> <nn>liar</nn>
```

ここでのアノテーションとはあくまで当該の表現に対して付与されるメタ言語情報であり, その情報は品詞のような文法的な情報だけに限らず, 音韻的, 統語的, 意味的, 語用論的なものを付与することも (理論的には) 可能です。しかし, そのようなアノテーションをおこなうためには常に一貫した基準が必要で, 特に意味に関わるようなアノテーションについては様々な議論があることには注意しましょう (cf. Fellbaum & Baker 2013)<sup>\*7</sup>。

### 2.1.2 コーパスとコーパスインターフェイス

アノテーションを付与されたデータを利用すれば全ての問題が解決するわけではありません。(2b) では簡単な事例を挙げましたが, 一般的にコーパスとして公開されているデータには非常に多くのメタ言語情報が含まれます。例えば, 品詞 (e.g., *VERB*), 語彙素 (e.g., *break*), 屈折形 (e.g., “*broke*”), そのデータの出典 (使用域 (*resister*)) といった膨大なアノテーションが付与されたデータを人力で解釈することは到底不可能です。この理

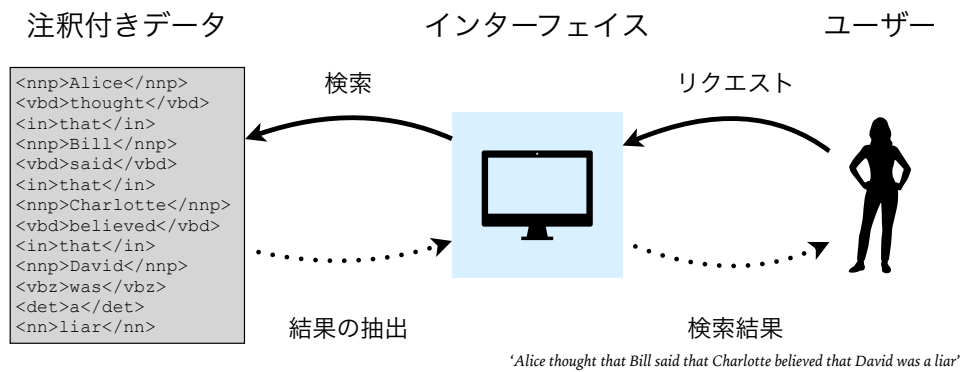


図1 コーパス、コーパスインターフェイス、ユーザーの関係

由から、コーパスのデータを扱うためには必要に応じてコーパスのデータを処理するための何らかの機構が必要になります。このようなコンピュータによる処理を前提としたデータを扱うための機構のことは一般的に、インターフェイス (interface) と呼ばれます<sup>\*8</sup>。この機構を簡単に図示すると図1のようになります。

コーパスインターフェイスとは機械処理が前提となっているコーパスのデータを人が扱いやすいように処理する機構を指します。図1では、ユーザーが(2a)の文をコーパスのデータから抽出するためのリクエストを送信し、インターフェイスが当該の条件に合致する事例である(2b)を検索・抽出し、それをユーザーが閲覧可能な(2a)のような形に成形する過程を示しています。

本来であれば、コーパスインターフェイスの構築にはプログラミング言語の習得が必要不可欠です。なぜならば生のアノテーション付きのコーパスデータを目視で確認することは現実的ではないからです。しかし、幸いなことにコーパスのデータにアクセスするためのサービスが数多く存在します。本稿で扱う Sketch Engine もその一つで、3. 節にてその使い方を簡単にみていきます。他にも言語学において頻繁に用いられるサービスとしては English-Corpora.org が挙げられますが、その詳細は長谷部 (2020) を参照してください。

ただし、Gries (2017, 4) が主張するように、有料・無料の是非を問わず、Web 上のコーパス検索アプリケーションは急な仕様変更や機能停止に直面する可能性が否めません。一定の検索結果が常に得られないという事態は、最悪の場合再現性が担保できなくなるという研究倫理上の大きな問題に直面する可能性が常に残ります。この理由から Gries はプログラミング言語の習得が全ての言語学者にとって必須の技能であるという立場をとりま<sup>\*9</sup>す。

現実的には、(言語教育や理論研究を含む広義の)言語研究に携わる全ての研究者がある程度のプログラミングの技能を習得するまでにはまだ時間がかかるでしょうし、中にはその過程を苦痛に感じる研究者も多いでしょう<sup>\*10</sup>。しかし、自分が普段利用しているアプリケーションに利用したいデータが含まれない場合、データの取得は自力でおこなう必要がでてくることは念頭に置くべきでしょう。また、使用するデータについてはどのような事例をどのような基準で分析し、どのような結果が得られたのかについて論文等で詳細に論じるだけでなく<sup>\*11</sup>、実際に利用したデータをすぐに提示できるようにする工夫などが必要となります。

仮に、データの取得、分析、考察といった各サイクルの透明化の試みを極限まで発展させるのであれば、言語学におけるオープンサイエンス (open science) 化を促進することができるでしょう。特に既存のコーパスを用いた言語研究であれば(構築に関わらない限り)倫理的な問題に抵触することはほとんどありません<sup>\*12</sup>。また、収集したデータに対して意味的なアノテーションを与えるような研究では再現性が特に問題になるため、データの公開は非常に重要な試みとして評価することができるでしょう<sup>\*13</sup>

## 2.2 コーパスをつかう「技術」

前節にて詳しく述べたように、コーパスの元データには品詞などのメタ情報が付与されている場合が多く、Sketch Engineに限らず、多くのコーパス検索アプリケーションはこのメタ情報を用いた検索が可能になっています。本節ではメタ情報の利用は複雑な条件に基づく検索を可能にする反面で、検索条件の厳しさと検索結果の量は一般的にトレードオフの関係にあることを確認します。

コーパス検索アプリケーションは様々な条件にもとづく検索を可能にしますが、複雑な検索にはある程度の慣れと技術が必要になります。どのようなデータを抽出するのかは各々のタスクによって異なりますが、一般的に検索条件の厳しさと検索結果の量はトレードオフの関係にあります。つまり、厳しすぎる検索条件では十分なデータが取れませんし、簡単すぎる検索条件では必要以上のデータが取れてしまいます。この関係は集合論における外延の多さと内包の豊富さの関係と並行的なものとして理解することができます (i.e., 被覆率と精度の関係)<sup>\*14</sup>。

この関係がどのようなものであるのかを実際の事例を用いて経験的に確認してみましょ

表2 各語の頻度と条件の厳しさに応じた頻度

	<i>kill</i>	<i>read</i>	<i>seek</i>	<i>test</i>	<i>watch</i>
1	15,148	27,841	16,606	22,685	21,912
2	14,973	27,710	16,605	6,896	18,934
3	1,423	4,354	2,006	1,198	4,430
4	1,036	3,315	1,357	934	3,362

う。ここで恣意的ではあるものの、*kill*, *read*, *seek*, *test*, *watch* という五つの語を挙げ、それぞれの事例に対して ‘British National Corpus (BNC), tagged by CLAWS’ を用いて (3) に挙げた四つの条件で検索した際に頻度がどのように変動するのかを調査してみました。その結果、表2にあるような頻度が確認されました。この検索条件の厳しさと頻度を R Core Team (2022) をもちいて可視化したものが図2です。ここでは条件が厳しくなるほどその頻度が少なくなるという大まかな傾向が確認できます。

- (3)
- i.  $\{test, tests, tested\}$ のように、品詞を問わない様々な屈折形を含んだ語 *xxx* のあらゆる事例 (i.e., `[lemma="xxx"]`)
  - ii.  $\{test, tests, tested\}$ のような様々な屈折形の語のなかでも、動詞としてタグづけされた語 *xxx* のあらゆる事例 (i.e., `[lemma="xxx" & tag="V.*"]`)
  - iii.  $\{test, tests, tested\}$ のような様々な屈折形の語のなかでも、動詞としてタグづけされた *xxx* の直後に冠詞{a, the}が後続するあらゆる事例 (i.e., `[lemma="xxx" & tag="V.*"] [tag="AT0"]`)
  - iv.  $\{test, tests, tested\}$ のような様々な屈折形の語のなかでも、動詞としてタグづけされた *xxx* の直後に冠詞{a, the}と名詞が後続するあらゆる事例 (i.e., `[lemma="xxx" & tag="V.*"] [tag="AT0"] [tag="N.*"]`)

研究者は仮説を検証、または形成するためにデータを取得、ないしは分析します。どのような目的の下で調査をおこなうにせよ、研究者は得られた結果から確認できる何らかの一般化を提示する必要があります。この一般化という過程を簡単に特徴付けると、当該のデータに含まれる共通性を見出すことに対応します<sup>\*15</sup>。あるデータからどれだけ「興味深い」傾向を見出せるかは分析者の理論的立場（や経験年数、センス）などによって大きく異なります。もしある傾向がその分析者の理論から予測できないようなものであれば、そ

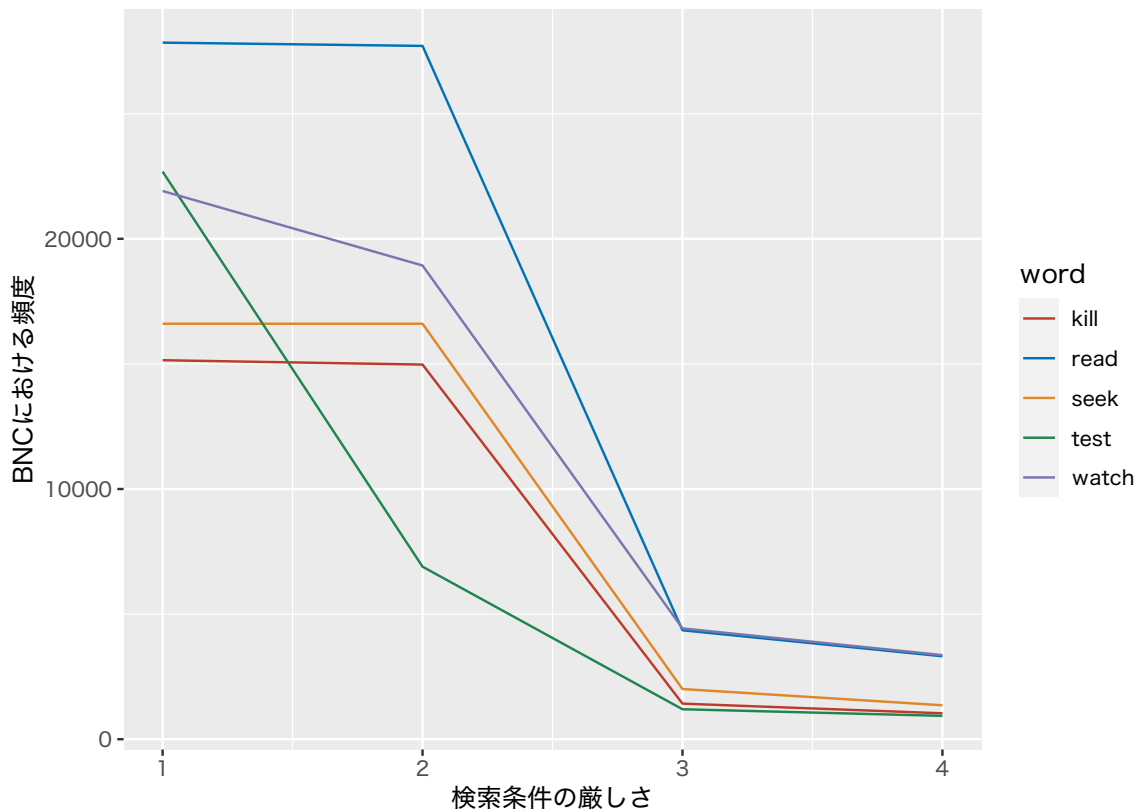


図2 各語の頻度と条件の厳しさの関係

これは（典型的に）理論的な議論の余地がある「興味深い」ものとなりますが、そうでない場合は「興味深い」ものとしてみなされない可能性があります<sup>\*16</sup>。しかし、一般的にそのデータの規模と一般化のしやすさには上でみたようなトレードオフの関係が成立することが容易に想像できます。つまり、データの規模が大きくなればなるほど、そこから共通性を見出すことが難しくなり、データの規模が小さくなればなるほど、そこから共通性を見出すことが簡単になりやすいのです。

ここで分析対象の量とその分析結果の一般化の難易度の関係をまとめると表3のようになる。一般化が困難であるというのは、その一般化がより複雑なものになりうるということに対応する。つまり、適切な規模のデータからは「 $P$ ならば $Q$ である」のように前件と後件が一つの命題となっている単純な一般化を導くことができますが、必要以上に大規模なデータからは「 $P_1$ かつ $P_2$ , または $P_3$ であれば $Q$ である」<sup>\*17</sup>のように前件と後件が複合的な命題からなるような複雑な一般化を導かざるを得ません<sup>\*18</sup>。

少ない事例から一般化を蓄積することの方が分析者の負担量は減る反面で、少ない事例

表 3 検索条件の厳しさと一般化の難易度

検索条件	事例の数	一般化
ゆるい	多い	困難
厳しい	少ない	容易

からの一般化は言明の一般性 (generality) の欠如という問題に直面することになります。言語学は言語の科学である以上、特定の個人の言語知識がどうなっているのか、ということよりもある言語  $L_i$  を使用する話し手全体がもつ傾向を明らかにしようとするものです。つまり、あまりに特殊な事例ばかりを扱っていても、 $L_i$  の全体の傾向を明らかにすることはできないということです。統計的には、調査を実施するまえに (i) 有意水準 ( $\alpha$ )、(ii) 効果量、(iii) 検定量 ( $1 - \beta$ )、(iv) サンプル・サイズのうちの、(i-iii) を決めてしまえば (iv) を決定することができる (水本・竹内 2010, 56-57)。一つの解決法としては、自分の決めた調査水準によって得られるデータの総数が統計的に得られたこのサンプル数と概ね一致するかどうかを検討することが挙げられるだろう。ただし、修辭的な表現などは出現頻度が低い場合も多々あるため、「統計的にも妥当な手法を用いて比喩表現を探しましたが、見つかりませんでした」のような事態に陥るのは馬鹿らしいので、分析対象の抽出には慎重になるべきでしょう \*19。

このような検索条件の厳しさと検索結果の量の関係を踏まえると、コーパス基盤の研究には (4) に示す難しさが付きまといます。この問題をどのように克服するのかは各々が解決すべき問題で、非常に泥臭い作業になっていきます。これに関しては、それぞれが 3. 節で概観する Sketch Engine のようなツールを使って試行錯誤していければと思います。

- (4) 適切な厳しさの検索条件によって、言語学的に有意義な一般化が可能になるような量のデータを抽出する必要性

### 3. Sketch Engine

本節ではコーパス検索アプリケーションの一つである Sketch Engine の機能を概観していきます (Kilgarriff et al. 2004, 2014)。このアプリケーションは高度な検索が可能で、様々な目的をもった研究者にとって非常に有用なツールとなることは間違いないでしょう

(cf. 黒田 2017)。

2022年9月現在の Sketch Engine では 94 カ国語の 742 個のコーパスが利用可能となっています。以降では ‘British National Corpus (BNC), tagged by CLAWS’ をコーパスとして利用しながら、以下に挙げる機能についてその概要と利用可能性とその注意点について簡単に論じていきます。このリストは完全なものではありませんが、コーパスをもちいた研究をおこなうためには事足りるでしょう。なお、利用可能性は神原個人の関心に多大な影響を受けているため、どのようなタスクに各機能が有用かどうかは各自で考えてもらえると嬉しいです。

3..1. Word Sketch

3..2. Word Sketch Difference

3..3. Thesaurus

3..4. Concordance

3..5. Wordlist

3..6. N-grams

3..7. Keywords

### 3.1 Word Sketch

■概要 Word Sketch とは、検索対象が当該のコーパスにおいてどのような振る舞いをみせるのかを要約したものを指します。図 3 に名詞 *dog* の Word Sketch の検索結果を記載します。この Word Sketch は検索対象とする語の頻度によってもその結果の豊富さが大きく変わることにご注意してください。

■利用可能性とその注意点 ある表現の大まかな傾向を把握したいときに便利です。例えば、*replace* の主語や目的語には何が生じやすいのかということ調べるのには有効な手段となる。しかし、共起語の傾向の確認は大雑把な理解しかもたらさないことには注意が必要です。



The screenshot shows the 'WORD SKETCH' interface for the word 'dog' in the British National Corpus (BNC). The word is listed as 'dog as noun 12,098x'. Below the word, there are six columns of related terms, each with a header and a list of items:

usage patterns	modifier	object_of	subject_of	and/or	modifies
poss ...	stray ... stray dogs	leette ... leette dog	bark ... dog barked	cat ...	handler ... police dog handler
Sfin ...	pet ... pet dog	train ...	foul ... dog fouling	bitch ...	warden ... dog warden
VPing ...	guide ... guide dogs for the blind	walk ...	whine ...	wolf ...	chaser ... official dog chaser
VPto ...	prairie ... prairie dogs	wag ... wagging the dog	sniff ... dog sniffing	sledge ...	chaser ...
SwH ...	mad ... a mad dog	bark ... barking dogs	bite ... dog bit	pet ...	dog dirt
Sing ...	sleeping ... to let sleeping dogs lie	breed ...	howl ...	dog ...	breeder ... dog breeder
It+ ...	sniffer ... sniffer dogs	feed ...	chase ...	duck ... the dog and duck	turd ...
	hunter ... a hunting dog	pat ...	savage ...	monkey ...	owner ... dog owners
	top ...	guide ... guide dogs	yap ...	jackal ...	shit ... dog shit
		muzzle ...	leap ...	Englishman ...	col... dog
		exercise ...	roam ...	sheep ...	
				horse ...	

図3 BNCにおける名詞 *dog* の振る舞い

### 3.2 Word Sketch Difference

■概要 Word Sketch Difference とは、二つの語彙同士のコーパス上での振る舞いを要約したものに該当します。図4にBNC上での名詞 *dog* と *cat* の振る舞いを要約したものを記載します。緑色の着色は *dog* に特有なものを、赤色の着色は *cat* に特有なものを示しています。

■利用可能性とその注意点 Word Sketch Difference が最も有用になるのは語どうしに成立する意味関係 (semantic relation) の分析でしょう (cf. Cruse 1986, 2011, Murphy 2003, 2010)。一般的に意味の類似性は共起語の類似性と相関があることが知られていますが、この詳細についてはまだ不明瞭な点が多いです。例えば、Murphy (2010, 109) は類義関係にある語のペアとして *sofa* と *couch* を挙げています。これを Word Sketch Difference で調べてみると、形容詞 *soft* は *sofa* としか、形容詞 *brown* は *couch* としか生起しないことが簡単にわかります。

ですが、意味の類似性と共起語の類似性の相関関係は完璧なものではないことには注意が必要です。例えば、先の例を用いて「*soft* となりうる椅子は *sofa* だけである」や「*brown* となりうる椅子は *couch* だけである」という結論を導くのは明らかに間違っています

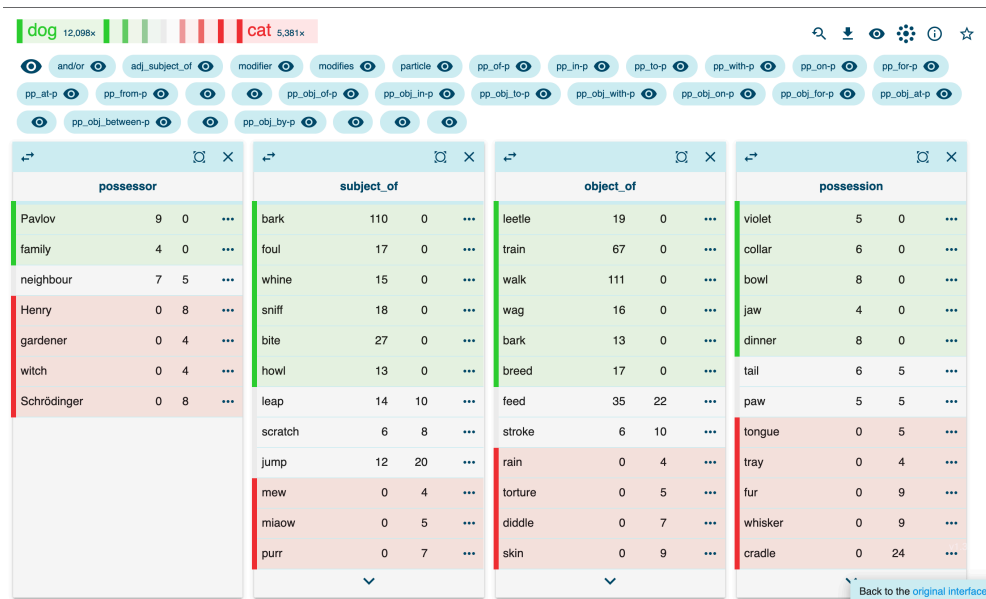


図4 BNCにおける名詞 *dog* と *cat* の振る舞い

(Kambara & Yamanaka 2023, 53–55)。コーパスをもちいた分析は研究者の直観が及ばないような事実を明らかにするものの、安直な結論に飛びつきやすくなることにも注意がいらいます。どのような頻度情報であればその表現の意味的な相違点を明らかにできるのかということはコーパスだけで明らかにすることはできません (cf. 神原 2021, 129–130)。

### 3.3 Thesaurus

■概要 Thesaurus は Sketch Engine によって自動的に生成された類義語のリストを指します。通常の単語リストとして出力することも可能ですが、図5のような可視化もできます。ここでは機械的に似た統語環境に発生しうる語を類義語としてまとめているだけなので、[*dog-doggy*] のような対が必ずしも得られるわけではないことに注意が必要でしょう。

■利用可能性とその注意点 Thesaurus の利用可能性としては、作例の支援などが挙げられます。黒田 (2011b) は、任意の数の変項を用意し、それぞれの変数に様々な表現をいれるという作例の作成方法を提案しています。(5) に動詞 *give* の例を挙げます。ここでは動詞 *give* に一つから三つの変項を用意し、それぞれの変項に様々な表現をいれています。

- (5) a. Alice gave \_\_\_\_ a car. ( $n = 1$ )

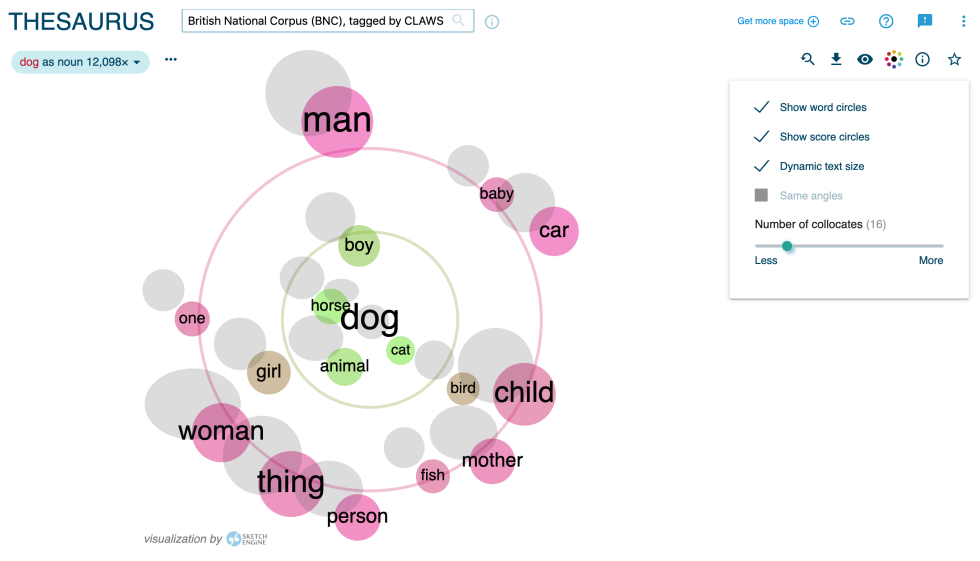


図5 名詞 *dog* の類義語

- i. Alice gave Bill a car.
- ii. Alice gave Charlotte a car.
- .....
- b. Alice gave \_\_\_\_ \_\_\_\_\_. ( $n = 2$ )
  - i. Alice gave Bill a bicycle.
  - ii. Alice gave Charlotte a book.
  - .....
- c. \_\_\_\_ gave \_\_\_\_ \_\_\_\_\_. ( $n = 3$ )
  - i. David gave Bill a bicycle.
  - ii. Elizabeth gave Charlotte a book.
  - .....

ここでどのような変項と表現を用いるのかは例文の作成者に委ねられることになりま  
す。ここで Thesaurus の機能をもちいて作成した単語リストから様々な表現をいれるこ  
とで、より客観的な作成をおこなうことができるでしょう、なお、当該の表現の形式的な  
類似性と容認度の変動にどのような対応関係がみられるのかを観察することは選択制限  
(selective restriction) の分析に有用なものとなります。

### 3.4 Concordance

■概要 Concordance は検索対象が含まれる事例を KWIC (*keyword in context*) の形、ないしは文単位で抽出する機能を指します。この機能は通常の検索機能とほぼ同じですが、より高度な検索も可能になっています。主な検索オプションは (6) の通りです。

- (6) a. **simple**: 語単位で自動的に事例を検索することができる。 *dog* であれば、名詞判定をし、複数形も含めた事例も抽出される。
- b. **lemma**: 語彙素単位で自動的に事例を検索することができる。名詞形の *test.n* ではなく、動詞形の *test* のみを検索する場合のみなどに使用することができる。
- c. **phrase**: フレーズ単位で自動的に事例を検索することができる。句動詞の *take out* を含む事例を検索することができるが、*took out* や *taking out* といった事例は含まれないことに注意。
- d. **CQL**: Corpus Query Language と呼ばれるデータベース検索用の言語を用いることで、複雑な条件の事例を検索することができる。

(6d) の CQL は正規表現を含む検索式によって複雑な検索を可能にする人工言語です。この言語では大括弧 [ ] が一つの単位をなし、その中に様々なタグを入力することができる<sup>\*20</sup>。Sketch Engine では通常の品詞情報に基づく検索式だけでなく、3.1 で導入した Word Sketch の情報も利用することができます。(7) にその使用例を挙げます。

(7) 便利な CQL:

- a. ‘*bring NP for NP*’ という形をとる事例:

```
[lemma="bring"] [tag!="IN"]{0,2} [tag="N.*"] [lemma="for"]  
[tag!="IN"]{0,2} [tag="N.*"]
```

- b. ‘*NP is like a shark*’ という形をとる事例:

```
[tag="N.*"] [lemma="be"] [lemma="like"] [tag="DT"] [lemma="shark"]
```

- c. 目的語に *dog, elephant, animal* のどれかを取る動詞 *kill* の事例<sup>\*21</sup>:

```
[ws("dog-n|elephant-n|animal-n", ".*object.*","kill-v")]
```

この機能によって得られた結果には (8) に示すような様々な操作を適用することができます。

- (8) a. DOWNLOAD: 得られた事例のダウンロードができる。形式としては.txt, .csv, .xlsx, .xml を選択することができる。現時点のブラウザの情報を pdf 形式で保存することも可能。
- b. CONCORDANCE ANNOTATION MODE: 各事例に対して任意のアノテーションをおこなうことができる。各事例の分類などに用いることができるが、このアノテーションの結果をダウンロードできないことには注意すること \*22。
- c. VIEW OPTIONS: 各事例に含まれる様々な属性や構造を表示することができる。これは元のコーパスファイルに含まれるタグによって生成される。
- d. GET A RANDOM SAMPLE: 最大で 10,000 件の任意の数のサンプルを抽出する。200 件のダウンロード結果と 201 件のダウンロード結果は異なるので注意すること。
- e. SHUFFLE LINES: 条件に合致する事例をランダムにシャッフルする。このランダム化はシャッフルする回数に応じて同じ結果が得られるようになっている \*23。
- f. SORT: 検索対象を含む前語三語の中から、特定の条件に従った並び替えを実施する。
- g. FILTER: 当該の結果から特定の条件に合致するもののみを抽出する。
- h. GOOD DICTIONARY EXAMPLES: Kilgarriff et al. (2008) によって定義された GDEX (*good dictionary example*) 順にもとづく並び替えをする。
- i. FREQUENCY: 当該の表現の前文脈, 当該の表現, 当該の表現の後文脈に生起する表現の頻度リストを (i) 語形, (ii) 品詞, (iii) タグ, (iv) レマの四つの基準によって生成する。
- j. COLLOCATIONS: 当該の表現の前後五語までに生起する表現の頻度リストを取得する。
- k. DISTRIBUTION OF HITS IN THE CORPUS: 当該の表現が使用しているコーパスに含まれるファイルのどの位置に生起するのかを可視化する。

Word	↓ Frequency ?	Word	↓ Frequency ?	Word	↓ Frequency ?
1 dog	7,846 ...	18 doggerel	24 ...	35 doges	9 ...
2 dogs	4,347 ...	19 dogmatically	22 ...	36 dog-like	9 ...
3 dogged	282 ...	20 dog-wheelk	21 ...	37 dogon	8 ...
4 dogma	254 ...	21 dog-leg	19 ...	38 dog-fighting	7 ...
5 dogmatic	214 ...	22 dog-tired	18 ...	39 doggo	7 ...
6 doggedly	111 ...	23 dogmatics	17 ...	40 dog-headed	7 ...
7 doggy	109 ...	24 doghouse	17 ...	41 dogfood	7 ...
8 dog-wheelks	66 ...	25 dogan	17 ...	42 dog-handlers	6 ...
9 dogmas	65 ...	26 dogwood	15 ...	43 dog-training	6 ...

図 6 名詞 *dog* の wordlist [words, starting with]

■利用可能性とその注意点 個人的に一番よく使う機能ですが、(i) 言語研究のためのデータ抽出と、(ii) 教材の開発で用いることができます。(i) に関しては、当該の条件に合致する事例を抽出し、それに対して意味的な特徴を中心とするアノテーションを加えるための一次データとして用います。(ii) については (8h) の GDEX によって並び替えをした上で単語テストを作成することが多いです。いずれの場合も検索結果をダウンロードする際にはブラウザ上の表示が KWIC か Sentence なのかを確認してからでないと、予想通りのデータが得られない場合があることに注意が必要です。

### 3.5 Wordlist

■概要 Wordlist は任意の条件にもとづく頻度表を作成する機能を指します。頻度表の作成には {lemmas, nouns, verbs, adjectives, adverbs, pronouns, conjunctions, prepositions} という 8 つの条件 (A) と {all, starting with, ending with, containing} という 4 つの条件 (B) の全 32 通り組み合わせが利用できます。図 6 は (A) words, (B) starting with の条件で *dog* を検索した結果のサンプルです。

■利用可能性とその注意点 ある表現の使用傾向などを確認するためには便利なものになります。例えば、英語の形態素 *re-* から始まる表現の頻度と、それが付与される語幹の関

Word	↓ Count ?	Word	↓ Count ?	Word	↓ Count ?
1 dogs and	340 ...	18 dog for	64 ...	35 dog would	38 ...
2 dog and	241 ...	19 dog food	64 ...	36 dogs with	37 ...
3 dog is	199 ...	20 dog in the	62 ...	37 dogs which	37 ...
4 dog to	180 ...	21 dog owners	61 ...	38 dogs of	37 ...
5 dog in	172 ...	22 dog had	58 ...	39 dog barked	36 ...
6 dog was	154 ...	23 dogs on	50 ...	40 dogs that	35 ...
7 dogs are	142 ...	24 dog out	50 ...	41 dogs or	35 ...
8 dogs were	137 ...	25 dog who	47 ...	42 dog from	35 ...
9 dogs in	114 ...	26 dogs for	45 ...	43 dog I	35 ...
10 dog with	90 ...	27 dogs have	43 ...	44 dogs in the	34 ...
11 dogs to	88 ...	28 dogs and cats	43 ...	45 dog which	32 ...

図7 名詞 *dog* から始まる  $n$  グラム ( $2 \leq n \leq 4$ )

係を調査するには有効な手段となると考えられるでしょう。しかし、このような形で得られた事例はあくまで形式的な共通性によって抽出されたものであるため、形式が類似しているからといって、意味的にも類似しているという保証がないことには注意すること。

### 3.6 N-grams

■概要  $n$  グラムとは、機械的に抽出される  $n$  組の形式のことを指します。この1組として計算される形式を仮に「語」とすれば、‘*Alice likes a dog*’という文の  $n$  グラムは (9) のようなものになります。

- (9) a. <sup>ユニ</sup> 1 グラム: {Alice}, {likes}, {a}, {dog}
- b. <sup>バイ</sup> 2 グラム: {Alice, likes}, {likes, a}, {a, dog}
- c. <sup>トライ</sup> 3 グラム: {Alice, likes, a}, {likes, a, dog}
- ...

Sketch Engine を用いることで、任意の  $n$  グラム ( $2 \leq n \leq 5$ ) を頻度順に取得することができる。図7は名詞 *dog* から始まる  $n$  グラム ( $2 \leq n \leq 4$ ) の頻度表です。これらの頻度表の取得には最低頻度の設定などの細かな設定が可能です。

British National Corpus (BNC), tagged by CLAWS

Spoken

**SINGLE-WORDS**

	Word	Focus corpus ?	Reference corpus ?	
1	er	88,425	961	...
2	erm	62,413	685	...
3	mhm	7,467	7	...
4	yeah	81,611	1,393	...
5	cos	15,883	354	...
6	gon	11,940	522	...
7	alright	7,994	335	...
8	ooh	4,390	176	...
9	mm	32,704	1,961	...

**MULTI-WORDS**  
Terms not available for this corpus

Back to the original interface

図8 BNC コーパスでの「書き言葉」と比較した時の「話し言葉」のキーワード

■利用可能性とその注意点  $n$  グラムは  $e_i$  という表現に後続する  $e_{i+n}$  の組を指すため、「 $e_i$  といえば、次は……」という情報に対応すると考えることもできるでしょう。頻度の高い  $n$  グラムはそれだけ経験する可能性が高い事例として扱うことができるため、反応速度が速くなる等の頻度効果が観察できるのかもしれませんが (cf. Baayen 2010)。このような観点から実験をおこなう際には刺激の作成にこの情報を利用することができるでしょう。

### 3.7 Keywords

■概要 Keywords では二つのコーパスにおいて、他方よりも一方のコーパスで頻出する語のリストを得ることができます。図8では、BNC のサブコーパスである「話し言葉」と「書き言葉」の間で、「話し言葉」に特有な語を挙げたものです。この比較するコーパスは様々なものを利用することができます。

■利用可能性とその注意点 コーパスを用いた研究に対するよくある批判として、「コーパスやレジスターが変われば分析結果が変わるだろう」というものがあります。Gries & Divjak (2010, 347–348) はこの批判への応答として、高頻度の表現の分布などはコーパスによって大きく変動するものの、態の交替 (i.e., 能動態 vs. 受動態) のような文法的な構文の分布などはコーパス間に統計的な差が観察できない研究が複数あることを指摘してい



ます。このようにコーパスごとの共通性と相違点を模索する研究などではこの機能は有用になるでしょう。しかし、この keywords だけでは詳細な分析を展開することができないことには注意する必要があります。

#### 4. おわりに

本 WS ではコーパスを利用する際に便利になるコーパス言語学の基礎として正規表現の基礎、アノテーションの必要性、コーパスとコーパスインターフェイスの区別について説明しました。その上で、Sketch Engine の代表的な七つの機能を概観しました。いずれの場合も内容としては必要最低限のことしか扱っていないため、研究や教育の場で利用するためには以下に挙げるような文献を参考にしてください。

■プログラミング関係 神原は Progate をオンライン教材としてプログラミングの勉強に使いました。どんな言語を用いるにせよ、基本的な利用法を学んでおくことは役立つでしょう。個人的には Ruby に一番慣れているのですが、Ruby を用いたテキスト分析に関しては田野村 (2012) が便利になるでしょう。言語学者向けの R へのイントロとしては Gries (2017) の二章が便利ですが、さらに先に進むためには松村ほか (2021) などが役立つでしょう。浅尾・李 (2013) は Python をもちいたテキスト処理に関して非常にわかりやすい解説になっています<sup>\*24</sup>。

■コーパス言語学・統計関係: McEnery & Hardie (2012), 石川 (2021) はコーパス言語学の教科書として代表的なもので、より近年のものであれば Stefanowitsch (2020) などが挙げられます。特に McEnery & Hardie (2012) は認知言語学をはじめとする機能主義的な言語学との関連を詳細に論じています<sup>\*25</sup>。また、Stefanowitsch は Gries と共にコロストラクション分析 (collostructional analysis) を提唱した研究者としてよく知られています (Stefanowitsch & Gries 2003)。

■統計・数学関係: 統計の教科書については (本当に) たくさんのものであるので、ここではコーパス分析に関わるようなものを中心に挙げます<sup>\*26</sup>。日本語であれば、石川ほか (2010) がよいですが、竹内・水本 (2023) も多くの分析法を具体例とともに解説しているため非常に分かりやすいです。英語であれば、Gries (2021) が非常に多くの事柄を扱っていて非常に勉強になりますが、R の操作を同時並行で学ぶという意味で難しい感じる方も

いらっしょると思います。Levshina (2015) や Glynn & Fischer (2010) などが参考になるでしょう。また、近年では統計モデルを用いたコーパス分析が主流となっていますが、その感覚を掴むためには久保 (2012) や Winter (2019) などが大変参考になります。

## 謝辞

本稿は 2022 年 9 月 10 日に電子語学教材開発研究部会 第 38 回研究会「これから始めるコーパス分析: Sketch Engine 活用術」で配布した資料を加筆・修正したものです。企画をしていただいた木村修平氏 (立命館大学), 世話人の近藤雪絵氏 (立命館大学), ならびにワークショップに参加して頂いた方々に感謝いたします。また、本稿の前身となった資料は京都大学の谷口研究室 (京都大学 認知言語学系研究室) で開催された自主ゼミでも「コーパス基盤アプローチへの招待: Sketch Engine の活用をとおして」というタイトルで利用しました。そこでの議論に参加された方々にも感謝します。最後に本稿の投稿を勧めて下さった水本篤氏 (関西大学) に感謝します。本稿に残る誤りは全て筆者によるものです。

## 注

\*<sup>1</sup> 本稿は 2022 年 9 月 10 日に電子語学教材開発研究部会第 38 回研究会「これから始めるコーパス分析: Sketch Engine 活用術」での配布資料を加筆・修正したものである。Workshop での配布資料という特性上、「ですます」調による記述になっている。

\*<sup>2</sup> これは仮定の研究方法であって、2.2 節でみる理由から、実際のコーパスで似たことをするととんでもない量のデータを扱うことになるのでお勧めしません。

\*<sup>3</sup> Web 検索だけでなく、近年の Chat GPT などの生成 AI サービスを利用することで簡単に正規表現の式を取得することができます。単純なものであればすぐに正解を取得することができますが、複雑な検索式を生成するのは難しくなります。他のタスク同様に、AI を利用したコードは鵜呑みにせず、表 1 を参照しながら適宜修正をおこなってください。

\*<sup>4</sup> 厳密には、このような検索式には二つの問題があります。一つは大文字の “A dog” のような事例を抽出できないという点で、もう一つは不定冠詞の異形である *an* を取得することができないという点です。ここでは議論を簡易にするため、そのような事情は考慮

しないことにします。

\*<sup>5</sup>(1) の正規表現の記述には冗長な部分が含まれてますが、ここでは度外視します。なお、(1a-c) の冒頭には半角スペースが含まれていることに注意してください。

\*<sup>6</sup> ここで使用した品詞パーサーは EngTagger で、<https://github.com/yohasebe/engtagger> から入手可能です。

\*<sup>7</sup> 生成 AI の発展によって、近年は事情が大きく変わってきていますが、理論言語学者が NLP に貢献できる一つの方向性としてはアノテーション方針の開発などが挙げられるでしょう。神原 (2021) は語の意味に関するアノテーション体系の開発・構築に関わる理論研究を事例研究とともに提示しています。なお、これに関する議論については黒田 (2012) なども参照してください。

\*<sup>8</sup> 国立国語研究所が提供しているコーパスである現代日本語書き言葉均衡コーパス (Balanced Corpus of Contemporary Written Japanese; BCCWJ) と中納言に関するならば、BCCWJ はコーパスデータで、「少納言」がそのデータにアクセスするためのインターフェイスとして機能していますが、神原 (2017) のような研究はこの区別がわかっていません。

\*<sup>9</sup> 私自身は Ruby や R を使えばある程度のことができますが、それまでには数年の時間がかかりました。

\*<sup>10</sup> とはいえ、生成 AI を上手に活用することで以前ほどの「プログラミング言語の流暢さ」が求められることはなくなるでしょう。しかし、出力されたコードのエラーを修正するためには基本的な知識が必要になるため、ある程度の学習は必要になることには注意しましょう。

\*<sup>11</sup> この辺りの話は神原の出身大学院の後輩のために執筆した神原ほか (2019) にて詳しく論じました。興味のある方は読んでみてください。

\*<sup>12</sup> 既存のコーパスを利用した分析に関して倫理的な問題が生じるとすれば、Web コーパスを利用した研究がその候補として挙げられるでしょう。Sketch Engine で利用可能な Web コーパスである TenTen コーパス (Jakubíček et al. 2013) などは大量のテキストデータを Web 上から取得し、コーパスとしてまとめたものです。多くの事例が利用可能となることは有り難い反面で、アダルトサイト等の一般的に好ましくないとされるソースの事例も多く含みます。そのような場で利用されることが多い表現の記述も重要であるとはいえ、その提示には細心の注意を払う必要があることには留意しましょう。

\*13 例えば, Kambara et al. (2023) は Open Science Framework (OSF) に利用したアノテーションの結果だけでなく, その処理のためのコードを公開しています。このように公開されたデータは様々な研究者の資料として利用できるため, 機会が許す限りにおいて積極的におこなうと良いでしょう。

\*14 例として「生き物」という内包をもつ集合を考えてみましょう。この集合には, この世に存在する多くの存在物がその集合の外延として含まれることとなります。ここで「生き物であり, かつ犬」という内包をもつ集合を考えると, 「生き物」という内包をもつ集合よりも少ない外延を含む集合が得られます。さらに「生き物であり, かつ犬, かつ日本に生息している」という内包をもつ集合であれば更に少ない外延を含む集合が得られます。このように集合の内包の豊富さとその外延の数はトレードオフの関係にあり, 検索条件の厳しさと検索結果の数の関係もおおむね同様の関係がみられる。

\*15 最適な一般化の述べ方については, 黒田 (2011a) を参照してください。

\*16 この「興味深さ」をどのように定式化するのか, という問題は非常に難しいため, 本稿ではこれ以上の考察をおこないません。

\*17 ここで意図されている言明を記号論理学における記法を使うと  $\{(P_1 \wedge P_2) \vee P_3\} \rightarrow Q$  となります。

\*18 近年では, このような複雑な一般化を統計モデルをつかって得ることが多いです。統計的なモデリングの解説は難しいため, 4. 節で紹介する文献を参照してください。

\*19 分析の実行可能性, という観点の他にもコーパスの事例の提示には倫理的な配慮が必要になる場合があります。詳細は注釈\*12 を参照してください。

\*20 これは (3) の分析でも用いたものです。

\*21 これは神原 (2021, Ch.4) で用いた CQL の一部に該当します。

\*22 共同研究などでアノテーションをおこなう際には, その基準を議論しながら設定していくことが多いため, 個人的には使わない機能です。

\*23 おそらく Sketch Engine 内では任意の規模の乱数の配列を管理することで, ランダム化によって同じ結果が得られるような工夫がなされていると思われます。

\*24 個人的には Python は簡単な読み書きしかできませんが, 心理言語学的な実験のプログラム作成や数値の処理にも便利なので一度学んでみるといいでしょう。

\*25 類似する議論としては Gries (2023b) が参考になるでしょう。

\*26 時折「数学や統計をどれだけ勉強する必要があるのか?」という質問を受けることが

あります。神原自身、数学がとてできるという自負は全くありませんが、個人的な意見としては、「ある程度の勉強は常に必要」と答えています。しかし、コーパス研究に関しては、カテゴリー変数と呼ばれる変数の分析を学べばそれである程度のことはできるようになります。更に先に進みたいのであれば更に多くのことが必要になりますが、どの本を読むべきかということよりも、自分が分析する対象がどのような特性を持つのかの事前の理解がより重要になるでしょう。また、共同研究によって自分の不足分を補ってもらうということも大切になるでしょう。例えば、Kambara & Chika (2023) では Gries (2019, 2023a) の議論をもとに、複数の共起指標の統合を試みましたが、それを数学者の共同研究者を迎えることで Kambara et al. (2024) のような論文を書くことをおこないません。このように自分の限界を知り、誰かに助けてもらうことは恥ずかしいことではないので、悩んだら周りに（いなければ私でも構いません）相談する癖を付けることが良いでしょう。

## References

- Baayen, R. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon*, 5(3):436–461.
- Cruse, A. D. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.
- Cruse, A. (2011). *Meaning in language: An introduction to semantics and pragmatics* (3rd edition). Oxford: Oxford University Press. (片岡宏仁 (訳) (2012). 『言語における意味: 意味論と語用論』東京: 東京電機大学出版局)
- Fellbaum, C. & Baker, C. F. (2013). Comparing and harmonizing different verb classifications in light of a semantic annotation task. *Linguistics*, 51(3), 707–727.
- Fillmore, C. J. (1990). “Corpus linguistics” vs. “computer-aided armchair linguistics”. In Svartvik, J. (ed.), *Directions in corpus linguistics: Proceedings from a 1991 Nobel symposium on corpus linguistics* (pp. 35–66). Berlin: Mouton de Gruyter.
- Glynn, D. & Fischer, K. (Eds.), (2010). *Quantitative methods in cognitive semantics: Corpus-driven approaches*. Berlin: Mouton de Gruyter.
- Glynn, D. & Robinson, J. A. (Eds.), (2014). *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*. Amsterdam: John Benjamins.

- Gries, S. T. (2010). Corpus linguistics and theoretical linguistics: A love-hate relationship? Not necessarily... *International Journal of Corpus Linguistics*, 15(3), 327–343.
- Gries, S. T. (2017). *Quantitative corpus linguistics with R: A practical introduction* (2nd edition). London: Routledge.
- Gries, S. T. (2019). 15 years of collocations: Some long overdue additions/corrections (to/of actually all sorts of corpus-linguistics measures). *International Journal of Corpus Linguistics*, 24(3), 385–412.
- Gries, S. T. (2021). *Statistics for linguistics with R: A practical introduction* (3rd edition). Berlin: Mouton de Gruyter.
- Gries, S. T. (2023). Overhauling collocational analysis: Towards more descriptive simplicity and more explanatory adequacy. *Cognitive Semantics*, 9(3), 351–386.
- Gries, S. T. (2023). Quantitative corpus methods of cognitive semantics/linguistics. In Li, F. T. (ed.) *Handbook of cognitive semantics* (pp. 328–350). Leiden: Brill.
- Gries, S. T. & Divjak, D. (2010). Quantitative approaches in usage-based cognitive semantics: Myths, erroneous assumptions, and a proposal. In Glynn, D. & Fischer, K., (Eds.), *Quantitative methods in cognitive semantics: Corpus-driven approaches* (pp. 333–353). Berlin: Mouton de Gruyter.
- Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen Corpus Family. In *7th International Corpus Linguistics Conference (CL 2013)*, (pp. 125–127).
- Kambara, K. & Chika, T. (2023). Toward a corpus-based identification of nominal relationality and uniqueness: A constructionist approach. In *Proceedings of the 37th Pacific Asia Conference on Language, Information and Computation* (pp. 47–73).
- Kambara, K., Chika, T., & Takahashi, N. (2024). Conflating directional association measures: A case study on NP constructions. *Journal of Corpus-based Lexicology Studies*, 6, 95–110.
- Kambara, K., Nozawa, H., & Takahashi, T. (2023). Differentiating valence patterns: A quantitative analysis based on formal and semantic attributes. *Constructions*,

15(2), DOI: <https://doi.org/10.24338/cons-571>.

- Kambara, K. & Yamanaka, T. (2023). Philosophy of data science for corpus linguistics: A pragmatistic point of view. *Annals of the Japan Association for Philosophy of Science*, 32, 47–73.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovvář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 7–36.
- Kilgarriff, A., Husák, M., McAdam, K., Rundell, M., & Rychlý, P. (2008). GDEX: Automatically finding good dictionary examples in a corpus. In *Proceedings of the XIII EURALEX international congress* (pp. 425–432).
- Kilgarriff, A., Rychlý, P., Smrž, P., & Tugwell, D. (2004). The Sketch Engine. In Williams, G. & Vessier, S. (Eds.), *Proceedings of XI EURALEX International Congress* (pp. 105–116).
- Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.
- McEnery, T. & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press. (石川慎一郎 (訳) (2014). 『概説コーパス言語学: 手法・理論・実践』 東京: ひつじ書房)
- McEnery, T. & Brezina, V. (2012). *Fundamental principles of corpus linguistics*. Cambridge: Cambridge University Press.
- Murphy, M. L. (2003). *Semantic relations and the lexicon: Antonymy, synonymy, and other paradigm*. Cambridge: Cambridge University Press.
- Murphy, M. L. (2010). *Lexical meaning*. Cambridge: Cambridge University Press.
- Pustejovsky, J. & Stubbs, A. (2013). *Natural language annotation for machine learning*. California: O'Reilly Media, Inc.
- R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Stefanowitsch, A. (2020). *Corpus linguistics: A guide to the methodology*. Berlin: Language Science Press.
- Stefanowitsch, A. & Gries, S. T. (2003). *Collostructions: Investigating the interaction*

- of words and constructions. *International Journal of Corpus Linguistics*, 8(2), 209–243.
- Wickham, H. & Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data*. California: O’Reilly Media, Inc.
- Winter, B. (2019). *Statistics for linguists: An introduction using R*. London: Routledge.
- 浅尾仁彦・李在鎬. (2013). 言語研究のためのプログラミング入門: *Python* を活用したテキスト処理. 東京: 開拓社.
- 石川慎一郎. (2021). ベーシックコーパス言語学 (第二版). 東京: ひつじ書房.
- 石川慎一郎・前田忠彦・山崎誠 (編) (2010). 言語研究のための統計入門. 東京: くろしお出版.
- 神原一帆. (2017). 日本語における「ノ構文」の振る舞い: 少納言コーパスを用いた属性パターンに基づく調査と「進撃の巨人」. *日本認知言語学会論文集*, 17, 390–401.
- 神原一帆. (2021). フレーム意味論にもとづく名詞の意味分析. 博士論文, 京都大学大学院人間・環境学研究科, 京都.
- 神原一帆・春日悠生・田中悠介. (2019). 発表の構成について: 分かりやすい発表をするための留意点. (最終更新: 2022年8月14日).
- 神原一帆・野澤元・高橋武志. (2022). 事態認知の焦点化パターンに対するコーパス基盤アプローチ: 動詞 STAB を例とした意味役割の分析. *Journal of Corpus-based Lexicology Studies*, 4, 14–28.
- 神原一帆・野澤元・高橋武志. (2024). 事態の焦点化と構文選択に関する量的分析: 動詞 *stab* の事例分析をとおして. *日本認知言語学会論文集*, 24, xxx–xxx.
- 久保拓弥. (2012). データ解析のための統計モデリング入門: 一般化線形モデル・階層ベイズモデル・MCMC. 東京: 岩波書店.
- 黒田航. (2003). 認知形態論. 吉村公宏 (編) 認知音韻・形態論 (pp. 79–154). 東京: 大修館書店.
- 黒田航. (2011a). 一般化の述べ方について: いかにも“過小”般化と“過大”般化を避けて最適な一般化を達成するか. (最終更新: 2011年9月14日).
- 黒田航. (2011b). 自作例を使った研究の基礎. 辻幸夫 (監修) 中本敬子・李在鎬 (編), 認知言語学の方法: 内省・コーパス・実験 (pp. 29–63). 東京: ひつじ書房.



- 黒田航. (2012). 言語学と言語処理の共生は可能か?: 統計基盤の言語処理の限界はどこにあるか? それは知識基盤の言語処理で克服できるか? 人工知能学会誌, 27(3), 326–332.
- 黒田航. (2017). Sketch Engine を使う; Regular Expressions を学ぶ. (スケッチエンジン・正規表現講習会 発表資料) .
- 竹内理・水本篤 (2023). 外国語教育研究ハンドブック: 研究手法のより良い理解のために (増補版) . 東京: 松柏社.
- 田野村忠温. (2012). *Ruby* によるテキストデータ処理. 東京: 明治書院.
- 長谷部陽一郎. (2020). English-corpora.org を用いた言語データの採取. (最終更新: 2020年5月29日) .
- 福田純也・矢野雅貴・田村 祐 (編). (2023). 第二言語研究の思考法: 認知システムの研究には何が必要か 東京: くろしお出版.
- 松村優哉・湯谷啓明・紀ノ定保礼・前田和寛. (2021). *R ユーザのための RStudio* [実践入門: *Tidyverse* によるモダンな分析フローの世界 (第二版)]. 東京: 技術評論社.
- 水本篤・竹内理. (2010). 効果量と検定力分析入門: 統計的検定を正しく使うために. より良い外国語教育研究のための方法 (pp. 47–73).
- 守田貴弘. (2013). 意味的分類の科学的妥当性. 言語研究, 144, 29–53.

## 付録 A 練習問題

“\*”をつけたものは難易度が高いです。これらの問題については本稿で扱った内容以外のことを考慮する必要がありますので、現時点でできなくても落ち込まないでください。用意した問題は40問近くありますが、特に解答は用意していません。回答に際してサポートが必要な場合は [kazy0324@pep-rg.jp](mailto:kazy0324@pep-rg.jp) までご連絡いただければ個別にフィードバックさせていただきます。

### A.1 正規表現

- A.1-1. 指定するファイルの中から、不定冠詞 *a* に後続する語の総数をテキストエディターの検索機能を使って調べなさい (e.g., “a dog”, “a beautiful”, ...).
- A.1-2. 指定するファイルの中から、不定冠詞 *a*, または定冠詞 *the* に後続する語の総数

をテキストエディターの検索機能を使って調べなさい (e.g., “a dog”, “the tall”, ...).

A.1-3. \*dogs.txt の中から、単数形の dog で出現する回数をテキストエディターの検索機能を使って調べなさい。

♣ dogs.txt は無作為に dog と dogs という表現を 10,000 件ランダムに連ねたものです。複数形を検索し、その差分を引き算によって求めることでも同じ結果が得られますが、折角なので正規表現を使いましょう。

A.1-4. \*hyphens.txt の中から、ハイフン“-”が 6 回連続で用いられている回数をテキストエディターの検索機能を使って調べなさい。

♣ hyphens.txt は“-”が最大 50 回連続する行を 10,000 件連ねたものです。間違っても手動で数えようなどとは思わないこと。

A.1-5. \* 指定するファイルの中から、五文字からなる英単語の総数をテキストエディターの検索機能を使って調べなさい (e.g., “whose”, “blond”).

A.1-6. \* 指定するファイルの中から、不定冠詞 a, または定冠詞 the に後続する語の総数五文字からなる英単語の総数をテキストエディターの検索機能を使って調べなさい (e.g., “whose”, “blond”).

A.1-7. \*(1) で挙げた正規表現には、英語における名詞の全てを列挙する必要がある、ということに加えて決定的な問題がもう一つあります。K さんのタスクとの関係からこれがどのようなものなのかを論じなさい。

## A.2 Sketch Engine

### 1.2.1 Word Sketch

1.2.1-1. BNC (tagged by CLAWS) のデータを利用し、horse の Word Sketch を取得し、その結果の一部を可視化しなさい。

1.2.1-2. BNC (tagged by CLAWS) のデータを利用し、horse と共起する語のリスト全てを csv としてダウンロードしなさい。

1.2.1-3. BNC (tagged by CLAWS) のデータを利用し、horse を目的語としてとる事例の全てを csv としてダウンロードしなさい。

1.2.1-4. \*BNC (tagged by CLAWS) のデータを利用し、動詞として用いられる *dog* の Word Sketch を取得し、その結果の一部を可視化しなさい。

1.2.1-5. \*BNC (tagged by CLAWS) のデータを利用し、動詞として用いられる *dog* と共起する語のリスト全てを csv としてダウンロードしなさい。

## 1.2.2 Word Sketch Difference

1.2.2-1. BNC (tagged by CLAWS) のデータを利用し、*begin* と *start* の Word Sketch を取得し、その結果の一部を可視化しなさい。

1.2.2-2. BNC (tagged by CLAWS) のデータを利用し、*end* と *finish* の Word Sketch を取得し、その結果の一部を可視化しなさい。

1.2.2-3. BNC (tagged by CLAWS) のデータを利用し、*stallion* と *mare* の Word Sketch を取得し、その結果の一部を可視化しなさい。

## 1.2.3 Thesaurus

1.2.3-1. BNC (tagged by CLAWS) のデータを利用し、*run* という動詞のレマの thesaurus のリストを.csv 形式でダウンロードしなさい。

1.2.3-2. BNC (tagged by CLAWS) のデータを利用し、*run* という名詞のレマの thesaurus のリストを.csv 形式でダウンロードしなさい。

1.2.3-3. BNC (tagged by CLAWS) のデータを利用し、*run* という動詞のレマの thesaurus のリストを可視化したものをダウンロードしなさい。

♣ Number of Collocates を 20 とし、Bubble Chart と Word Cloud の両方をダウンロードすること。

1.2.3-4. BNC (tagged by CLAWS) のデータを利用し、*run* という名詞のレマの thesaurus のリストを.csv 形式でダウンロードしなさい。

♣ Number of Collocates を 20 とし、Bubble Chart と Word Cloud の両方をダウンロードすること。

## 1.2.4 Concordance

- 1.2.4-1. BNC (tagged by CLAWS) のデータから CQL を利用し、レマ *read* のコンコーダンスを取得し、その結果を.csv 形式でダウンロードしなさい。
- 1.2.4-2. BNC (tagged by CLAWS) のデータから CQL を利用し、動詞のレマ *read* のコンコーダンスを取得し、その結果を.csv 形式でダウンロードしなさい。
- 1.2.4-3. \*BNC (tagged by CLAWS) から CQL を利用し、動詞 *read* の直後に定冠詞がくる事例のコンコーダンスを取得し、その結果を.csv 形式でダウンロードしなさい。
- 1.2.4-4. \*BNC (tagged by CLAWS) から CQL を利用し、定冠詞の直後に名詞が後続する事例のコンコーダンスを取得し、その結果を.csv 形式でダウンロードしなさい。
- 1.2.4-5. \*BNC (tagged by CLAWS) のデータを利用し、動詞 *read* の直後に定冠詞がくる事例のコンコーダンスを取得し、その結果を.csv 形式でダウンロードしなさい。
- 1.2.4-6. \*BNC (tagged by CLAWS) のデータを利用し、動詞 *stab* の目的語のコンコーダンスを取得し、その結果を.csv 形式でダウンロードしなさい。
  - ♣ 神原ほか (2022, 2024) では動詞 *stab* を分析していますが、動詞ではない事例を省くためにこの CQL をもちいてデータを抽出しました。

## 1.2.5 Wordlist

- 1.2.5-1. re から始まる動詞のリストを抽出し、その結果を.csv 形式でダウンロードし、この結果が *retake* のような事例ばかりでないことを確認しなさい。
- 1.2.5-2. tion でおわる名詞のリストを抽出し、その結果を.csv 形式でダウンロードしなさい。
- 1.2.5-3. ly でおわる形容詞のリストを抽出し、その結果を.csv 形式でダウンロードしなさい。
- 1.2.5-4. exam という形式を含むレマのリストを抽出し、その結果を.csv 形式でダウン

ロードしなさい。

- 1.2.5-5. re から始まる動詞のリストを抽出し、その結果を.csv 形式でダウンロードし、この結果が *retake* のような事例ばかりでないことを確認しなさい。

## 1.2.6 N-grams

- 1.2.6-1. BNC (tagged by CLAWS) のデータを利用し、word 単位で、最低頻度を 10 とした 3-gram のリストを.csv 形式でダウンロードしなさい。
- 1.2.6-2. BNC (tagged by CLAWS) のデータを利用し、lemma 単位で、最低頻度を 10 とした 3-gram のリストを.csv 形式でダウンロードしなさい。
- 1.2.6-3. \*BNC (tagged by CLAWS) のデータを利用し、word 単位で、{*I, you, we, they, it*}を除く、最低頻度を 10 とした 4-gram のリストを.csv 形式でダウンロードしなさい。
- 1.2.6-4. \*BNC (tagged by CLAWS) のデータを利用し、lemma 単位で、{*I, you, we, they, it*}を除く、最低頻度を 10 とした 4-gram のリストを.csv 形式でダウンロードしなさい。

## 1.2.7 Keywords

- 1.2.7-1. BNC (tagged by CLAWS) における話し言葉と書き言葉の Single-words を比較し、その絶対頻度と相対頻度を含む結果をダウンロードしなさい。
- 1.2.7-2. BNC (tagged by CLAWS) における話し言葉と書き言葉の N-grams を比較し、その絶対頻度と相対頻度を含む結果をダウンロードしなさい。

## 1.2.8 応用編

ここでは実際の研究等で利用する可能性のあるタスクの一部を挙げています。ここでみるものは複数の機能を利用する必要があるため、一筋縄ではいかないと思われませんが、是非挑戦してみてください。

- 1.2.8-1. ★CQL を利用し、BNC (tagged by CLAWS) のデータから *be made from* と *be made of* の両方を含むコンコーダンスを.csv 形式でダウンロードしなさい。
- 1.2.8-2. ★BNC (tagged by CLAWS) に含まれる名詞のレマ (e.g., *test*) のリストを入手し、その結果をもとに動詞形 (e.g., “*tested*”) のレマの (i) コンコーダンスと (ii) 頻度表の両方を.csv 形式でダウンロードしなさい。
- 1.2.8-3. ★BNC (tagged by CLAWS) に含まれる動詞のレマ (e.g., *test*) のリストを入手し、その結果をもとに “re” からはじまる動詞形 (e.g., “*retested*”) のレマの (i) コンコーダンスと (ii) 頻度表の両方を.csv 形式でダウンロードしなさい。
- 1.2.8-4. ★2.2 節で論じたような検索条件の厳しさに応じて頻度が影響を受けるさまを任意の名詞五語を例に確認し、その結果を別のアプリケーションを用いて可視化し、図 2 と類似の結果が得られることを確認しなさい。
- ♣ 神原はこの結果の可視化に R Core Team (2022) を利用しましたが、適宜 Excel 等の表計算アプリケーションを利用して問題ありません。

我が国の **motivational strategies** の研究動向の概観：  
CiNii データベースを用いたシステマティックレビュー

田中博晃  
近畿大学

**An Overview of Research Trends in Motivational Strategies in Japan:  
A Systematic Review Using the CiNii Database**

Tanaka, Hiroaki  
*Kindai University*

---

**Abstract**

本研究の目的は、システマティックレビューの手法を用いて我が国の英語教育分野での **motivational strategies** の研究動向を把握した上で、それらの研究成果を統合する試みを行うことである。国立情報学研究所の CiNii データベースから抽出・スクリーニングした過去 10 年（2013 年～2023 年）に発行された英語教育分野の **motivational strategies** に関する論文を 11 編を分析した結果、我が国で最も多く行われた **motivational strategies** の研究は「実践の省察と提言」で、研究対象はすべて大学生だった。**motivational strategies** の有効性を評価した研究は 3 編で、それらすべてで Dörnyei (2001) の **motivational strategies** が用いられていた。これらの研究結果を統合した結果、日本人大学生にはグループやクラスの雰囲気作りという基礎的な環境を創出する **motivational strategies**、教育課程、学習事項、教材、タスク、アクティビティなどを学習者の内在的な目標などの実用価値と関連づけることで、学習者の興味関心をひくという **motivational strategies** の有効性が高いと考えられる。

**Keywords:** *motivational strategies, systematic review, CiNii*

---

**1. はじめに**

近年の外国語教育分野での動機づけ研究は大きく分けて 6 つの方向性で議論されている (Dörnyei, 2020)。それが、(1)「動機づけ理論の概念整備」、(2)「ダイナミックス」、(3)「教育場面への理論の応用」、(4)「無意識/意識下」、(5)「L2 motivation self system に代表されるビジョン」、(6)「長期的な持続性」、である。Boo, Dörnyei, and Ryan (2015) は、外国語教育研究分野での動機づけ研究の動向をより詳細に調査するため、2005 年から 2014 年までに出版された 415 本の主要な動機づけ研究の論文を統括レビューした。

その結果、415本のうち約3分の1に該当する論文が「動機づけ概念」(motivation)ではなく、学習者を「動機づけること」(motivating)を扱っていたことが明らかになった。つまり、上述の6つの研究分野の内、(3)の「教育場面への理論の応用」が外国語教育研究における動機づけ研究の主要なテーマの中でも活発に研究されていると言えよう。特に理論を実践に応用する具体的な研究テーマとして着目されているのが、「授業で教師が学習者の動機づけを高めるための具体的な授業テクニック」(Dörnyei, 2001)と定義される *motivational strategies* である。

そこで本研究では、*motivational strategies* に関する我が国の研究動向についてシステマティックレビューを用いて概観することで、現状での我が国の *motivational strategies* の研究成果がどのように集約できるのか議論したい。

## 2. 研究の背景

### 2.1 *motivational strategies* について

まず *motivational strategies* の訳語について、*motivational strategies* を外国語教育研究の分野に導入した Dörnyei (2001)の翻訳書では「動機づけを高める英語指導ストラテジー」(ドルニエイ, 2005)とされている。しかしストラテジーという表現は「学習ストラテジー」(*learning strategies*)を想起させることもあり、訳語が統一されているわけではない。他にも「動機づけを高める方略」(例えば、田中, 2010; 2013)などもあり、近年は「動機づけ方略」が多用されている(例えば、廣森 2011; 2014; 竹内, 2016)。のちに詳述するが、本研究はシステマティックレビューの手法を用いて CiNii に登録されている論文を抽出するため、*motivational strategies* の多様な訳語を検索語として使用した。よって、この論文内で *motivational strategies* について記述する際は、検索語との混乱を避けるために、英語表現の *motivational strategies* をそのまま使用する。

*motivational strategies* を最も体系的に研究した文献として Dörnyei (2001)が挙げられる。Dörnyei (2001)は Dörnyei & Csizér (1998)が実証的な研究手法を使って提示した10個の *motivational strategies* をさらに拡張させ、心理学で提唱された動機づけ理論をベースに35個の *motivational strategies* (Appendix 参照)を創出した。その上で、それら35個の *motivational strategies* を授業実践場面を想定した4つのカテゴリーに分類した *process oriented model* を提示した。4つのカテゴリーとは、「動機づけの基礎的な環境の創造」(*creating the basic motivational conditions*)、「学習開始時の動機づけの喚起」(*generating initial motivation*)、「動機づけの維持と保護」(*maintaining and protecting motivation*)、そして「肯定的な回顧的自己評価の促進」(*encouraging positive retrospective self-evaluation*)である。Dörnyei (2001)によると、*motivational strategies* には場面カテゴリーに準じた特性があり、*process oriented model* は35個の *motivational strategies* を適切な使用場面に応じて4つの段階に分類した点が特徴である(図1参照)。



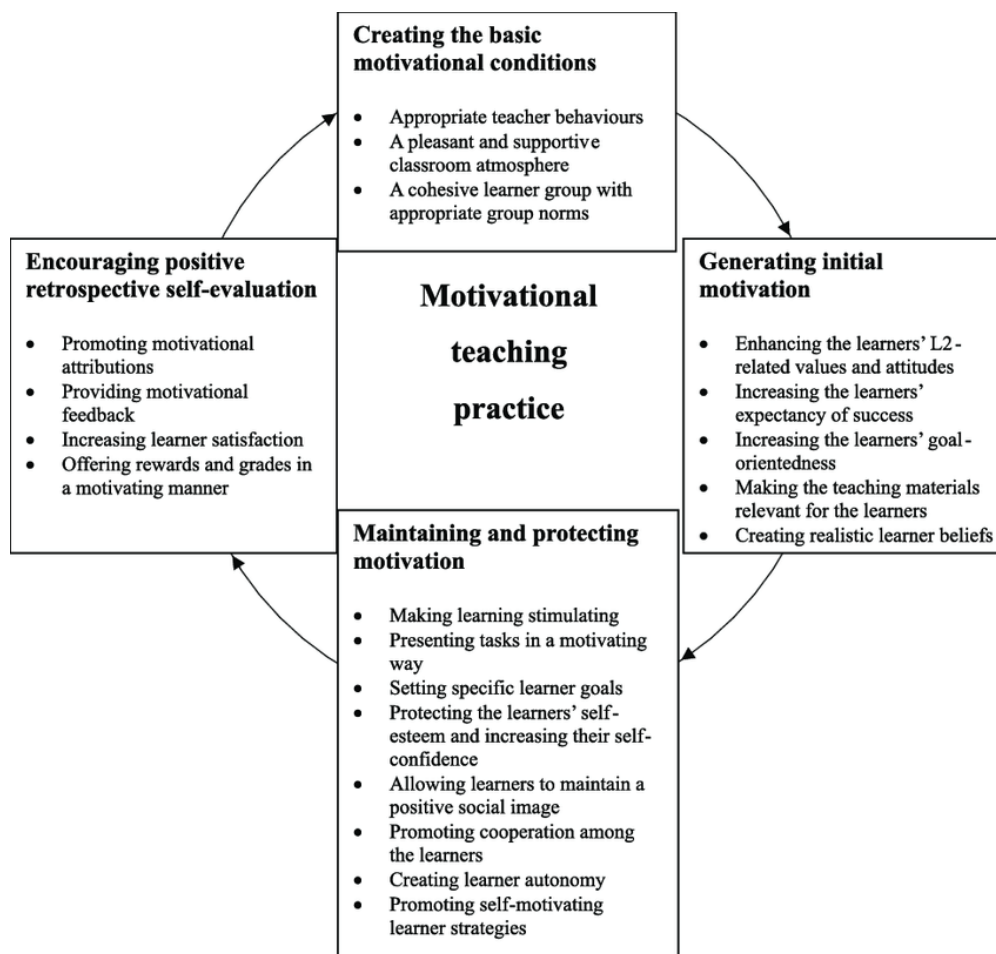


図 1. Dörnyei (2001)による process oriented model

Dörnyei & Csizér (1998)および Dörnyei (2001)の motivational strategies は、その後、多くの研究で引用され、国内外で研究が進められてきた。近年の国外の motivational strategies の研究動向については、Li (2020)が体系的なレビューを行っている。Li (2020)は 2000 年から 2019 年に Social Sciences Citation Index (SSCI) database に所収された研究論文から 10 編の論文 (Alrabai, 2014; Chen & Dörnyei, 2007; Dörnyei & Csizer, 1998; Guilloteaux & Dörnyei, 2008, Lee & Lin, 2019; Lee, Gardenr & Lin, 2019; Papi & Abdollazadeh, 2012; Sugita McEown & Takeuchi, 2014; Wong, 2014) を抽出し、それらの研究成果と今後の motivational strategies の研究課題を提示した。Li (2020)は motivational strategies に関する論文を「実験研究」、「調査研究」、「開発研究」の 3 つに分類し、その中でも「調査研究」が最も多くなされている点を指摘した。「調査研究」の中でも最も多く見られた研究は、motivational strategies の使用と動機づけの高さや学習成果などの想定される効果変数との相関関係を扱った質問紙調査 (例えば、Guilloteaux & Dörnyei, 2008) であった。また motivational strategies の使用実態とその評価を行う「調査研究」も見られた (例えば、Cheng & Dörnyei, 2007)。一方、motivational strategies の効果を直接測定する「実験研究」では、ランダムサンプリングを伴う実験群と対照群による厳密

な実験研究は見られず、ランダムサンプリングを伴わない準実験研究がわずかに見られたのみだった（例えば、Albrabi, 2014）。

一方、国内の研究においても Dörnyei (2020) の(3)「教育場面への理論の応用」に関する研究は多く見られるが（例えば、Yamamoto, 2022）、*motivational strategies* について体系的に網羅したレビュー研究は著者が把握する限りではまだ行われていないようだ。

## 2.2 システマティックレビュー

文献調査を網羅的に行い整理する手法として近年英語教育学部分野で注目されている手法の 1 つにシステマティックレビューがある。「システマティックレビュー」(*systematic review*) は 1990 年代から医療分野で用いられてきた文献調査の手法で、The Cochrane Collaboration など代表的な定義が複数存在するが、本研究では「あるテーマに関するすべての研究の収集、批判的吟味、統合におけるバイアスを制限するための戦略の応用」(唐 他, 2014) とする。特に「統合」のプロセスを量的に行っているシステマティックレビューは「量的システマティックレビュー」(*quantitative systematic review*) と呼ばれ、複数の研究結果を量的に統合する「メタ分析」(*meta analysis*) はそのオプションの 1 つと理解されている。一方、「統合」プロセスを質的に行っている場合は「質的システマティックレビュー」(*qualitative systematic review*) と呼ばれる。

システマティックレビューを用いた英語教育分野の研究としては、Honda *et al.* (2018), Mizumoto, Urano and Maeda (2014), Stapleton and Collett (2011), 本田 他 (2020) などが挙げられる。Honda *et al.* (2018) は関東甲信越英語教育学会の機関誌 *KATE Journal* に所収されている 272 編の論文 (1985 年～2017 年) を対象に、本田 他 (2020) は小学校英語教育学会の機関誌 *JES Journal* の 189 編の論文 (2003 年～2019 年) を対象に、Mizumoto *et al.* (2014) は全国英語教育学会の機関誌 *ARELE* に所収されている 450 編の論文 (1990 年～2013 年) を対象に、Stapleton and Collett (2011) は全国語学教育学会の機関誌 *JALT Journal* に所収されている 297 編の論文 (1979 年～2008 年) を対象に、各学会誌での研究動向の調査を行っている。これらの研究は、いずれも特定の学会の機関誌を対象に、そこに所収されている論文から当該学会での研究動向を概観している。

そこで本研究は日本国内の *motivational strategies* の研究動向を探るため、システマティックレビューの手法を用いる。その際、分析範囲を特定の学会誌に絞らずに、CiNii に登録されている我が国の論文まで広げることで、我が国での *motivational strategies* の研究動向を概観する。

## 3. 目的

本研究の目的は我が国の英語教育分野での *motivational strategies* の研究動向の把握である。特に CiNii のデータベースを利用し、そこから過去 10 年間で所収されている論文を整理したうえで、それらの研究動向を分析した上で、先行研究の *motivational*

strategies の研究成果を統合する試みを行う。

## 4. 調査

### 4.1 対象文献の抽出手順

本研究では、国立情報学研究所 (NII) の CiNii (NII 学術情報ナビゲータ) を使用し、過去 10 年 (2023 年～2014 年) に発行された英語教育分野の motivational strategies に関する論文の抽出を行った。システムティックレビューの手法に則り、3 段階のデータ抽出を行った (表 1 参照)。

第 1 次抽出として、CiNii のデータベースにて motivational strategies とそれに関連するキーワードで検索を行った。英語で記載された論文と日本語で記載された論文の両方を抽出するために、英語と日本語のキーワードで検索を行った。特に日本語検索の際は motivational strategies の訳語の多様性を考慮し、できるだけ多くの検索ワードを準備した。具体的な検索ワードは「motivational strategies, English」、「motivational strategies, EFL」、「motivational strategies, L2」、「動機づけ方略, 英語」、「動機付け方略, 英語」、「動機づけを高める方略, 英語」、「動機づけを高める, 方略, 英語」、「動機づけを高める, ストラテジー, 英語」、「動機づけを高める英語指導ストラテジー」である。その結果、21 件の論文が抽出された。

表 1. 文献抽出の手順

段階	方法	結果
第 1 次抽出	CiNii のデータベースでの検索	21 編
第 2 次抽出	文献を以下の基準で第 1 次スクリーニング ・学会出版物あるいは大学紀要である。 ・本文がオンラインで公開されている	14 編
第 3 次抽出	文献を以下の基準で第 2 次スクリーニング ・論文のタイトル, 要旨, および本文から研究目的が motivational strategies であるか	11 編

次に第 2 次抽出として、検索結果から得られた文献のスクリーニングを行った。CiNii の検索結果には、学会発表の要旨、発表論文集、英検などの助成団体からの補助金への成果報告書も含まれていることから、以下の 2 つの条件を満たしているかどうかで絞り込みを行った。

- (1) 学会出版物（日本学術会議にて学術研究団体として登録された学会が発行する出版物に収録されている文献）あるいは、大学紀要（大学が発行する研究紀要及び報告書に掲載されている文献）である。
- (2) 論文の本文がオンラインで公開されている文献である。

以上の結果、14編の論文が抽出された。

次に第3次抽出として、抽出された論文のタイトル、要旨、および本文から *motivational strategies* をメインの目的として扱っている論文かどうか判別するスクリーニングを行った。その結果、1件は *teacher motivation* を維持・高めるための *motivational strategies* の研究で、2件は要旨に *motivational strategies* の言及があるが、論文自体の目的に *motivational strategies* が関連していない研究であったので、この3編は分析対象から除外した。その結果、11編の論文（アスコー, 2016; 林田, 2019; Kobayashi, 2014; 古賀, 2022; Leung, 2019; Millington, 2014; Sugiwa, 2019; 吉住, 2014; 竹内, 2016; Warchulski, 2014; Yamamoto, 2022）が分析対象として抽出された。

## 4.2 分析プロトコル

次に抽出された11編の論文の内容を精査し、システマティックレビューの先行研究（Honda et al., 2018; 本田 他, 2020; Mizumoto *et al.*, 2014; Stapleton and Collett, 2011）および *motivational strategies* のレビュー研究（Li, 2020）を参照し、4つの特性（「形式」、「研究対象」、「データタイプ」、「教育的介入の有無」）と、特性ごとのカテゴリーを用いて *motivational strategies* の研究動向の全体像の把握を行った。

まず形式はリサーチタイプと目的の観点から Li (2020) に基づいて「実験研究」、「調査研究」、「開発研究」の3つの種類で分類を行った。

「実験研究」とは、*motivational strategies* の効果を検討する目的で、実験研究の手法を使って、プレ測定とポスト測定を使ながら動機づけの変動を統計的に検討する研究である。

「調査研究」とは、*motivational strategies* の使用実体や有用性の評価を目的に質問紙調査などのデータから調査を行う研究である。

「開発研究」は動機づけ理論に基づきながら *motivational strategies* の開発を目的にした研究で、先行研究での研究成果や動機づけ理論に基づく考察が中心になる。

しかし各論文を上記の基準で精査した結果、Li (2020) の分類に当てはまらない論文が多数見られた。そこで本研究では Li (2020) の分類に加え、抽出された論文からボトムアップ的に新たなカテゴリーを創出した。それらは「概説」と「実践の省察と提言」である。「概説」とは応用言語学研究の入門者や中高の英語教員を対象にした先行研究の概略を解説した論文である。「実践の省察と提言」とは著者自身の教育実践を *motivational strategies* の観点から授業改善しようとする研究である。

また 11 編の論文の内、著者自身の教育活動の一環として学習者の動機づけの変動をプレ測定とポスト測定で検討している研究もあった。介入による動機づけの変動をプレ測定とポスト測定で統計的に検討している点では「実験研究」の要素もあり、また自身の教育活動に対する学生の評価という点では「実践の省察と提言」の要素もあった。そこで、本研究では該当研究のデータ収集デザインに着目し、上記を区別した。まず測定する質問紙に着目し、それが心理尺度かどうかで「実験研究」と「実践の省察と提言」を区別した。つまり動機づけという心理概念である潜在変数を観測変数から測定する場合、1つの概念を最小でも3項目で測定する必要がある (Costello & Osborne, 2005)。この基準に合致しない測定器具を用いている場合は「実験研究」に該当しないと考え、「実践の省察と提言」と見なした。「実験研究」に必要なデータ収集デザインには、他にもサンプリングや要因の事前統制など多様な条件があるが、本研究ではまずは尺度の観点からスクリーニングを行った。その結果、対象となったすべての研究で「実験研究」と「実践の省察と提言」が判別可能であった。

次に「研究対象」「データタイプ」「教育的介入の有無」については、システマティックレビューの先行研究 (Honda et al., 2018; 本田 他, 2020; Mizumoto et al., 2014; Stapleton and Collett, 2011) に基づいて各カテゴリーを設定した。「研究対象」のカテゴリーは「中学生」「高校生」「大学生」「その他」とし、「データタイプ」のカテゴリーは「量的」「質的」「混合」「その他・なし」とし、「教育的介入の有無」のカテゴリーは「有」「無」とした。以上の分析プロトコルは以下の表 2 のとおりである。

表 2. 分析プロトコル

特性	カテゴリー
形式	「実験研究」, 「開発研究」, 「調査研究」, 「概説」, 「実践の省察と提言」
研究対象	「中学生」, 「高校生」, 「大学生」, 「その他」
データタイプ	「量的」, 「質的」, 「混合」, 「その他・なし」
教育的介入の有無	「有」, 「無」

## 5. 結果

### 5.1 全体傾向の把握

motivational strategies の研究の全体傾向を把握するため、抽出された 11 編の論文を「リサーチタイプと目的」の観点から分類を行った。その結果、表 3 の結果が得られた。

過去 10 年で最も多く行われた motivational strategies の研究は「実践の省察と提言」の 5 件であった。次に多かったのが「調査研究」および「概説」で、各 3 件だった。一方、「実験研究」と「開発研究」は 0 件であった。なお国際誌の調査を行った Li (2020) では「調査研究」が最も多く 6 件で、「実験研究」と「開発研究」が同数の 2 件だった。

国際誌という性質上、「概説」と「実践の省察と提言」は見られなかった。

以上の点から、我が国の英語教育分野での *motivational strategies* の研究は、研究成果を著者自身が自分の教育実践に取り入れ、授業改善を行う応用研究が多く行われていることが示された。一方で、国際誌でみられる動機づけ理論から *motivational strategies* を開発する研究や開発された *motivational strategies* の効果を実験研究で検証する、いわば基礎研究に該当する研究は見られなかった。

表 3. 「リサーチタイプと目的」による分類

	実験研究	調査研究	開発研究	概説	実践の省察と提言
本研究	0	3	0	3	5
Li (2020)	2	6	2	-	-

次に「研究対象」、「データタイプ」、「教育的介入の有無」の3つの特性から *motivational strategies* の11編の研究の分類を行った。なおこれら3つ特性は *ARELE* 所収論文のシステマティックレビューを行った Mizumoto *et al.* (2004)に準拠したプロトコルを用いて、本研究との比較を行うこととした。なお Mizumoto *et al.* (2004)のターゲットは *motivational strategies* という特定の研究領域に特化せずに、英語教育全体の傾向を把握することだが、この比較を通して *motivational strategies* の研究動向が英語教育全体でどのように位置づけられるかが把握できると考えられる。

システマティックレビューの結果、表4の結果が得られた。11編の *motivational strategies* の研究のうちデータ収集が行われた研究は6編であった。6編すべての研究で大学生を対象にデータ収集を行っており、そのうち2編は教員にもデータ収集を行っていた。次にデータタイプについては、量的が3編、質的が1編、混合が2編であった。ただ *motivational strategies* の研究の多くは「実践の省察と提言」であったため、データ収集の方法は心理測定尺度よりは授業評価アンケートの形式が多くを占めていた。介入の有無については、介入有が2編、無が4編だった。

表 4. 「研究対象」、「データタイプ」、「教育的介入の有無」による分類

	対象				タイプ			介入		
	中	高	大	他	量	質	混	他	有	無
本研究	0	0	7	2	2	2	3	-	2	5
Mizumoto <i>et al.</i> (2004)	45	82	211	-	341	13	31	28	84	329

対象について Mizumoto *et al.* (2004)と比較すると、*motivational strategies* の研究は多

くが大学生が占めているのに対し、*ARELE* 所収論文では大学生が多いものの、その他の年齢層の学習者にも研究の範囲が広がられていた。タイプについては *motivational strategies* の研究は量的、質的、混合がまんべんなく用いられているのに対して、Mizumoto *et al.* (2004)では量的が圧倒的に多いという結果であった。介入の有無については、両研究とも介入無が介入有よりも多い結果だった。

## 5.2 *motivational strategies* の有効性

次に 11 編の論文のうち、*motivational strategies* の有効性を調査した論文に焦点を当て、日本人英語学習者に有効だと考えられる *motivational strategies* の抽出を行った。*motivational strategies* の有効性を調査した論文は「調査研究」で 3 編、「実践の省察と提言」で 1 編であった。これら 4 つ論文ではすべて Dörnyei (2001)の 35 の *motivational strategies* を使用していた。

ただし、そのうちの 1 編の論文は 5 件法の質問紙調査で *motivational strategies* の評価を行ったが、1~2 が否定的回答、3~5 が肯定的回答となっている点、各質問項目が Dörnyei (2001)のどの *motivational strategies* に該当するのか明記されていなかった点を考慮し、分析から除外とした。その結果、3 つの論文 (アスコー, 2016; Kobayashi, 2014; Yamamoto, 2022) を分析対象とした。

対象になった 3 つの論文の内、2 件は *motivational strategies* の有効性を大学生対象に質問紙調査を行った研究であった。もう 1 件は動機づけが高まった授業を行った教員に対して半構造化インタビューを行い、得られた *motivational strategies* の有効性に関する質的データを Dörnyei (2001)に基づいて分類した研究だった。

3 つの論文で得られた結果を統合した結果、35 個の *motivational strategies* の内 20 個の *motivational strategies* の有効性が支持されたことが明らかになった (表 5 参照)。具体的には、3 つの研究すべてで有効性が支持された *motivational strategies* は 1 個のみで、2 つの研究で有効性が支持されたのは 10 個、1 つの研究で有効性が支持されたのは 9 個であった。

まず、3 つの研究すべてで有効性が支持された *motivational strategies* は 1 つのみで、それは基礎的な環境を作る *motivational strategies* の中の C6「集団の結束を強化する」だった。

次に 2 つの研究で有効性が支持されたのは、基礎的な環境を作る *motivational strategies* では、C1「扱う教材に対する自分の熱意と、それが自分に個人的にどんな影響を及ぼしているかについて、実例を挙げて説明する」、C3「生徒と個人的な関係を築く」、および C5「教室に楽しく、支援的な雰囲気を作る」だった。授業開始時の *motivational strategies* では、G10「L2 学習過程に対する学習者の内在的な関心を高める」と G15「教育課程と教材を、学習者に関連の深いものにする」であった。動機づけを維持する *motivational strategies* では、M17「教室内での活動が単調にならないように、学

習をより興味深く楽しくする」と M23「学習者に定期的な成功体験を与える」であった。最後の肯定的な自己評価を促進するための motivational strategies では、E32「肯定的情報フィードバックを学習者に与える」、E33「学習者の満足感を高める」、および E35「動機づけを高める方法で評点を用いる。評点の持つ動機づけを失わせる衝撃をできる限り少なくする」であった。

表 5. motivational strategies の有効性の評価の比較

process-oriented model	motivational strategies	Yamamoto (2022)	アスコー (2016)	Kobayashi (2014)	計
C	6	○	○	○	3
C	1	○	○		2
C	3	○	○		2
C	5	○	○		2
G	10	○	○		2
G	15		○	○	2
M	17	○	○		2
M	23	○	○		2
E	32	○		○	2
E	33	○	○		2
E	35	○	○		2
C	2	○			1
G	9		○		1
G	11		○		1
G	13	○			1
M	18			○	1
M	20	○			1
M	21	○			1
M	25		○		1
M	27		○		1

Note. 各 motivational strategies の通し番号とその内容の詳細は Appendix 参照

最後に 3 つのうちいずれか 1 つの研究で有効性が支持されたのは、基礎的な環境を作る motivational strategies では C2「生徒の学習を真剣に受けとめる」のみだった。授業開始時の motivational strategies では、G9「仲間のお手本を見せることで、言語に関連する価値観を高める」、G11「L2 とその使用者、また外国人らしさ全般に対する肯定的



で開放的な気質を育てることで、統合的価値観を高める」、および G13「特定の課題および学習全般に関する生徒の成功期待感を高める」であった。動機づけを維持する *motivational strategies* では、M18「タスクの魅力を増すことで、学習を学習者にとって興味深く楽しいものにする」、M20「動機づけを高める方法でタスクを提示し、実施する」、M21「教室で目標設定の手法を用いる」、M25「学習環境において不安を誘発する要素を取り除き、あるいは緩和することによって、言語不安を軽減することを支援する」および M27「学習者が学習課題に取り組んでいるときに、肯定的な社会的心象を保持することを可能にする」であった。最後の肯定的な自己評価を促進するための *motivational strategies* での該当はなかった。

## 6. 考察

本研究はシステマティックレビューという手法を用いて我が国の過去 10 年での *motivational strategies* の研究動向を概観した。

まず我が国の過去 10 年で最も多く行われた *motivational strategies* の研究は「実践の省察と提言」で、研究者が授業実践者として *motivational strategies* を自身の授業実践の中に取り入れる方法を検討し、その成果を授業評価アンケートなどから省察する内容だった。これは CiNii で抽出された論文の多くが大学紀要に所収されていたことに関連していると考えられる。

また Li (2020)の国際誌のレビュー研究の結果でも見られた *motivational strategies* の有効性の評価を行う研究が我が国の研究でも見られた。具体的には 716 名の大学生を対象にした大規模な質問紙調査から *motivational strategies* の評価を行った研究 (アスコ、2016)、動機づけを高める効果が得られた授業の実践者に対して半構造化インタビューを行い、質的データから有効な *motivational strategies* を抽出した研究 (Yamamoto, 2022)、そして自らの授業実践に *motivational strategies* を取り入れて授業評価アンケートから有効性を検討した研究 (Kobayashi, 2014) が挙げられる。

これらの研究で得られた結果を質的システマティックレビューで統合したところ、Dörnyei (2001)が提示する 35 の *motivational strategies* の内、日本人大学生の英語学習者に有効と評価できる *motivational strategies* は 20 個であった。中でも 3 編すべてで評価が高かった *motivational strategies* は「集団の結束を強化する」(C6) だった。

近年の大学教育ではアクティブラーニング型授業が重視されており、授業内でペアワークやグループワークを使いながら、プロジェクト型学習、協調学習、協同学習などが行われている。グループでの学習において、学習者同士が充実した学習体験を得るには、グループに学習者がうまく溶け込み、活発なコミュニケーションを行うことで、学習者同士が学びあえる雰囲気作りが大切である。システマティックレビューの結果、同様のグループダイナミクスに関連した *motivational strategies* として、「教室に楽しく、支援的な雰囲気を作る」(C5) も有効性の評価が高い *motivational strategies* として本研究

から得られた。以上の結果から、グループやクラスの雰囲気作りという基礎的な環境を創出することが、日本人英語学習者に有効な *motivational strategies* と考えられる。

また有効性評価が高かった他の *motivational strategies* に着目すると、学習者の内発的動機づけに関連する *motivational strategies* も見られた。具体的には「教室内での活動が単調にならないように、学習をより興味深く楽しいものにする」(M17)、「タスクの魅力を増すことで、学習を学習者にとって興味深く楽しいものにする」(M18)、「動機づけを高める方法でタスクを提示し、実施する」(M20)である。学習者の興味や関心をひき、内発的動機づけを高めることの重要性は広く認知されている一方で、どのように興味や関心をひいて内発的動機づけを高めるか、という具体的な方法については多くの教員の悩みでもあろう。

内発的動機づけを高める具体的方法のヒントになる *motivational strategies* は、システマティックレビューで得られた「教育課程と教材を、学習者に関連の深いものにする」(G15)である。この *motivational strategies* の理論背景になっているのは「期待×価値理論」(*expectancy × value theory*, Wigfield and Eccles, 2000)である。「期待×価値理論」は学習者が学習に対して見いだす「価値」(*value*)の質に着目した動機づけ理論で、その中で提唱されている「実用価値」(*utility value*)は「現在の学習が将来のために役立つと位置付ける認知」(鈴木・櫻井, 2001)と定義される。この実用価値が学習者の自己に取り入れられ、十分に内在化されていれば、学習者は内発的に動機づけられた状態にあると見なされる (Bradshaw, 2023; Vansteenkiste, Lens, & Deci, 2006)。つまり学習者が持つ目標などの実用価値に沿った教材、アクティビティやタスクを取り入れることで、学習者を内発的動機づけに導くことができると考えられる。

以上の結果から、教育課程、学習事項、教材、タスク、アクティビティなどを学習者の内発的な目標などの実用価値と関連づけることで、学習者の興味関心をひくという *motivational strategies* が日本人英語学習者に有効と考えられる。

## 7. 今後の課題

最後に本研究の限界点について述べておく。第1に本研究はシステマティックレビューの手法を用いて *motivational strategies* を扱っている研究を CiNii から網羅的に抽出した。第1次抽出はできるだけ多くの論文を対象にできるように多様なキーワードを用いて検索した。しかし *motivational strategies* を中心に扱いながらも、タイトル、アブストラクト、そして論文キーワードに *motivational strategies*、あるいはその訳語を含めていない論文は CiNii の検索システムのアルゴリズムでは抽出できなかった。今後はハンドサーチなども含めながら、より網羅的に論文を抽出する必要がある。第2に本研究は *motivational strategies* の有効性の評価として最大公約数の発想を用いながら質的な結果の統合を行った。今後はメタ分析なども用いてながら、統合の方法をより多様化する必要がある。

## 謝辞

本研究は科研費 22K00834 の支援を受けた研究である。

## 参考文献一覧

- Alrabai, F. (2016). The effects of teachers' in-class motivational intervention on learners' EFL achievement. *Applied Linguistics*, 37, 307–333.
- アスコー明子 (2016) . 「小学校教員免許履修学生が考える英語学習動機付け—動機付けストラテジー質問用紙からの考察—」 . 『国際経営・文化研究』, 21, 229-254.
- Boo, Z., Dörnyei, Z., & Ryan, S. (2015). L2 motivation research 2005-2014: Understanding a publication surge and a changing landscape. *System*, 55, 145–157.
- Bradshaw, E. L. (2023). Causes, costs, and caveats: Reflections and future directions for goal contents theory. In R. M. Ryan (Ed.), *The Oxford handbook of self-determination theory* (pp. 139–159). Oxford University Press.
- Cheng, H.-F. & Dörnyei, Z. (2007). The use of motivational strategies in language instruction: The case of EFL teaching in Taiwan. *Innovation in Language Learning and Teaching 11*, 153–174.
- Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, 10, 1–9.
- Dörnyei Z. (2001). *Motivational strategies in the language classroom*. Cambridge University Press. (ドルニェイ, Z. 米山朝二・関昭典 (訳) (2005) . 動機づけを高める英語指導ストラテジー35. 大修館書店) .
- Dörnyei, Z. (2020). *Innovations and challenges in language learning motivation*. Routledge.
- Dörnyei Z. & Csizér, K (1998). Ten commandments for motivating language learners: Results of an empirical study. *Language Teaching Research*, 23, 203–229.
- Guilloteaux & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly*, 42, 55–77.
- 林田朋子 (2019) . 「授業力向上のための英語指導実践の省察—動機づけを高めるための英語指導ストラテジーを用いて—」 . 『現代社会学部紀要』, 18, 105–116.
- Honda, K., Hoshika, M., Aoyama, T., Someya, F., & Yamamoto, T. (2018). A systematic review of articles in KATE 1–31: Changing trends in the field of English education. *KATE Journal*, 32, 85–98.
- 本田勝久・田所貴大・星加真実・染谷藤重 (2020) . 「小学校英語における研究動向—JES Journal のシステマティックレビュー—」 . *JES Journal*, 20, 351-366.

- 唐文涛・小島原典子・河合富士美・津谷喜一郎 (2014) . 「診療ガイドラインとシステムティック・レビュー」. 『薬理と治療』, 42, 189-197.
- Kobayashi, K. (2014). Student reactions to content-based instruction in EFL reading classes: A survey on motivational strategies. 『大阪女学院短期大学紀要』, 44, 31-44.
- 古賀功(2022). 「基本的欲求と動機づけ:教師が使用する動機づけ方略の考察」. 『KELES ジャーナル』, 7, 46-53.
- Lee, T. S., Gardner, D., & Lau, K. (2019). The effects of L2 motivational strategies: Within and beyond the L2 classroom. *Innovation in Language Learning and Teaching*, 14, 1-15.
- Lee, T. S., & Lin, S. Y. (2019). English teachers' uses of motivational strategies beyond an established framework. *Educational Research*, 61, 1-18.
- Leung, R. C. Y. (2019). Teacher-orientated pedagogical practices for reducing demotivation in English language learning contexts in Japan. 『実践女子大学 CLEIP ジャーナル』, 5, 9-20.
- Li, H. (2020). Motivational strategies in English as foreign language teaching: A research review. *INTESOL Journal* 17, 137-154.
- Mizumoto, A., Urano, K, & Maeda, H. (2014). A systematic review of published articles in ARELE 1-24: Focusing on their themes, methods, and outcomes. *ARELE*, 25, 33-48.
- Millington, N. T. (2014). Using personalized narratives as motivational strategies. 『国際情報学部研究紀要』, 14, 97-105.
- Papi, M. & Abdollahzadeh, E. (2012). Teacher motivational practice, student motivation, and possible L2 selves: An examination in the Iranian EFL context. *Language Learning*, 62, 571-594.
- Stapleton, P., & Collett, P. (2010). JALT Journal turns 30: A retrospective look at the first three decades. *JALT Journal*, 32, 75-90.
- Sugita McEown, M., & Takeuchi, O. (2014). Motivational strategies in EFL classrooms: How do teachers impact students' motivation? *Innovation in Language Learning and Teaching*, 8, 20-38.
- Sugiura, E (2019). What motivational strategies can teachers use to encourage reading in an L2 for pleasure? 『名古屋短期大学研究紀要』, 57, 81-92.
- 吉住香織 (2014) . 「Motivational strategies と生徒の英語学習意欲 : 学習者はどのような指導を動機づけに効果的と考えるか」. 『國學院大學教育開発推進機構紀要』, 5, 37-60.
- 竹内理 (2016) . 「英語学習の動機を高め、維持するには : 動機づけ要因と動機づけ方略の観点から」. *JACET Kansai Journal*, 18, 36-51.
- 田中博晃 (2010) . 「英語の授業で内発的動機づけを高める研究」. *JACET Journal*, 50, 63-80.

- 廣森友人 (2014) . 「学習者の動機づけと英語熟達度が動機づけ方略への認識に与える影響」 . 『立命館言語文化研究』 , 22, 159–167.
- Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic versus extrinsic goal contents in self-determination theory: Another look at the quality of academic motivation. *Educational Psychologist, 41*, 19–31.
- Warchulski, D. (2014). Accounting for learners' positive behavior using a motivational strategies framework. *New Directions in Teaching and Learning English Discussion, 2*, 187–193.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81.
- Wong, R. M. (2014). An investigation of strategies for student motivation in the Chinese EFL context. *Innovation in Language Learning and Teaching, 8*, 132–154.
- Yamamoto, J. (2022). A Comparison of two different types of motivational strategies: A case study of two EFL instructors. *International Journal of Curriculum and Development, 25*, 43–56.

## Appendix. Dörnyei (2001)による 35 の motivational strategies

### C. 動機づけの基礎的な環境を作る motivational strategies

1. 教材に対する教員の熱意と、それが自分に与えた影響を実例で説明し解説する。
2. 生徒の学習を真剣に受け止める。
3. 生徒と個人的な関係を築く
4. 生徒の親たちと協力関係を築く。
5. 教室に楽しく、支持的な雰囲気を作る。
6. 集団の結束強化を促進する。
7. 明示的な集団規範を作成して、生徒と話し合い、認めてもらう。
8. 集団規範をしっかりと守らせるようにする。

### G. 学習開始時の動機づけの喚起する motivational strategies

9. 仲間のお手本を見せることで、言語に関連する価値観を高める。
10. L2 学習過程に対する学習者の内在的な関心を高める。
11. L2 とその使用者、また外国らしさ全般に対する肯定的で開放的な気質を育てることで、「統合的」価値観を高める。
12. L2 知識と結びついた道具的価値観に対する生徒の理解を高める。
13. 特定の課題および学習全般に関する生徒の成功期待感を高める。
14. 生徒が認める明確な教室目標を立てることで目標志向性を高める。
15. 教育課程と教材を学習者に関連の深いものにする。
16. 現実的な学習者信念を作る手助けをする。

### M. 動機づけを維持し保護する motivational strategies

17. 教室内での活動が単調にならないように、学習をより興味深く楽しくする。
18. タスクの魅力を増すことで、学習を学習者にとって興味深く楽しくする。
19. 学習者をタスクへの積極的な参加者になるように求めることで、学習を興味深く楽しくする。
20. 動機づけを高める方法でタスクを提示し、実施する。
21. 教室で目標設定の手法を用いる。
22. 生徒の目標に向けた情熱を形式化するため、生徒と契約手法を用いる。
23. 学習者に定期的な成功経験を与える。
24. 定期的に励ましを与えることにより、学習者の自信を育む。
25. 学習環境において不安を誘発する要素を除き、あるいは緩和することによって、言語不安を軽減することを支援する。
26. 学習者に多様な学習ストラテジーを教えることにより、自己の学習能力に対する自

信を構築する。

27. 学習者が学習課題に取り組む時に、肯定的な社会的心象を保持することを可能にする。
28. 学習者間の協力を促進することにより、生徒の動機づけを高める。
29. 学習者の自律性を積極的に促進することで生徒の動機づけを強化する。
30. 学習者の自己動機づけ能力を強化する。

E. 肯定的な自己評価を促進する **motivational strategies**

31. 学習者の中に努力帰属を高める。
32. 肯定的情報フィードバックを学習者に与える。
33. 学習者の満足感を高める。
34. 動機づけを高めるように報酬を与える。
35. 動機づけを高める方法で評価を用いる。評点の持つ動機づけを失わせる衝撃をできる限り少なくする。

# Methods of Research on Depth of Processing for Written Corrective Feedback: A Scoping Review

Yoko Fujisawa

*Graduate Student, Kansai University*

Atsushi Doi

*Graduate Student, Kansai University*

Natsuko Shintani

*Kansai University*

---

## Abstract

This review employed a scoping review approach to synthesize studies published between 2015 and 2024, focusing on the depth of processing (DoP) of written corrective feedback (WCF) in second language (L2) writing. The objective was to identify trends and issues in the research methods employed in this area. The publications were extracted from Google Scholar using three keywords. After screening with coding criteria, 15 studies were identified from an initial pool of 1,850 published studies. Our review suggests that most of the selected studies were conducted in university settings and compared different types of WCF as independent variables, with DoP and accuracy as dependent variables. To measure DoP, the two primary methods adopted were think-aloud protocols (TA) and written languaging (WL). The levels of DoP varied across the studies, ranging from two to five. Several gaps in the existing DoP studies were identified and discussed, providing recommendations for future research.

**Keywords:** depth of processing, written corrective feedback, scoping review

---

## 1. Introduction

A scoping review is a type of literature review that aims to map out the existing research on a particular topic or field to examine the breadth and nature of studies and research in the area of interest, and to identify gaps and potential areas for future research within the existing literature (Hanzawa et al., 2024). Unlike systematic reviews, which aim to synthesize empirical evidence to answer a specific research question, scoping



reviews have a broader focus on identifying the nature, extent, and range of research within a given area. Several studies in the field of applied linguistics have employed scoping reviews to explore diverse topics, such as language learning strategies (Chong & Reinders, 2022), and task-based language teaching (Tullock & Ortega, 2017). This review employed a scoping review approach to explore the research methods in research on DoP of WCF in L2 writing.

Research on WCF has gained significant momentum, with previous studies demonstrating its effectiveness in facilitating language learning and improving accuracy in revised writing (e.g., Bitchener & Storch, 2016; Shintani et al., 2014; Van Beuningen et al., 2012). The majority of these studies have focused on the effects of different types of WCF, particularly on revising discrete linguistic forms in revised drafts or new pieces of writing. These studies have reported small to moderate beneficial effects of WCF, although many variables mediate these effects (Kang & Han, 2015).

In addition to product-oriented research, recent years have seen growing attention to process-oriented research examining how learners process WCF (e.g., Shintani & Ellis, 2015; Simard et al., 2015). Investigating WCF processing is an intriguing aspect for both researchers and educators. Theoretically, L2 learning theories, such as the Noticing Hypothesis and the Interaction Hypothesis, assume that WCF can lead to L2 learning through noticing and cognitive comparison during processing. Pedagogically, as Leow (2015) rightly points out, while WCF is an external intervention, it is crucial to elucidate the internal perspective, i.e., how learners utilize the WCF.

## **2. Theoretical background**

Schmidt's Noticing Hypothesis (Schmidt, 1990) is one of the most influential theories in second language acquisition (SLA). To provide a definition of *noticing*, Schmidt introduces *consciousness* as a broad category of human cognition, divided into awareness, intention, and knowledge. Consciousness as awareness has three levels: perception, noticing, and understanding. Noticing refers to focal awareness where one pays attention to a stimulus as a private experience that can be verbally reported. While noticing alone does not lead to acquisition, the Noticing Hypothesis claims that learners can acquire linguistic input that they notice, or nothing is learned unless it has been noticed (Lightbown & Spada, 2013).

The effects of WCF on learners' noticing have attracted researchers, leading to extensive research on this topic (e.g., Coyle & Cánovas Guirao, 2019; Hanaoka, 2007; Hanaoka & Izumi, 2012; Izumi & Bigelow, 2000; Lázaro-Ibarrola, 2021; Lindgren & Sullivan, 2003; Qi & Lapkin, 2001). More recent studies (e.g., Suzuki, 2012, 2017) have

categorized these levels of linguistic processing during writing and WCF processing as lower (noticing) or higher (understanding). These levels correspond to Schmidt's (1990) notions of awareness at the level of noticing versus understanding. Studies classified awareness levels based on presence or absence of linguistic reasoning in TA and reflections.

Leow (2015) criticizes that noticing feedback does not guarantee deeper processing or restructuring of prior knowledge. He argues that the Noticing Hypothesis does not account for how prior knowledge interacts with noticed feedback and DoP needed to restructure interlanguage and thus, noticing is a low level of awareness. Leow (2015) proposed the concept of DoP, which refers to the cognitive effort, analysis, and elaboration that learners use when encoding input, including activating prior knowledge, testing hypotheses, and forming rules (Leow, 2015). Examining DoP in L2 writing is particularly important because factors like the slower pace of writing/ processing and visibility of written WCF might facilitate deeper processing.

Some researchers reviewed studies on DoP in L2 writing. Cerezo et al., (2019), for example, summarized seven studies that examined the levels of DoP or levels of awareness. Cerezo et al. included studies informed by three different theoretical underpinnings: four out of seven studies were based on the Noticing Hypothesis (Kuiken & Vedder, 2002; Qi & Lapkin, 2001; Sachs & Polio, 2007; Suzuki, 2012), two based on the notion of engagement (Storch, 2008; Wigglesworth & Storch, 2012), and one based on DoP (Caras, 2019). Although the review is informative for our understanding of research on DoP, there is a lack of focus on theoretical underpinnings. Considering the fact that more studies have been done on DoP in recent years and there is still no paper reviewing the methods measuring DoP among the studies, it is important to investigate updated trends in methodology for investigating DoP in L2 writing. To address this issue, the current study aims to synthesize the research methods used to analyze DoP, focusing on studies examining WCF on L2 writing. Our review focuses entirely on research based on Leow's (2015) DoP, synthesizes the research methods used in such studies, and explores the unresolved issues to offer recommendations for future research.

### **3. Research questions**

In this review, we identify patterns in the research designs and data collection methods used in studies exploring how deeply students process different error correction techniques in L2 writing. We also aim to highlight the issues with the research methods used in these studies. Two research questions are set.

1. What are the trends in research methods used in studies investigating DoP

elicited by WCF?

2. What are the issues with the research methods used in studies investigating DoP?

## **4. Methods**

### **4.1 Selection of studies**

To address the research questions, we focused on published studies that measured L2 learners' DoP of WCF on their writing. The literature search was carried out by consulting an electronic database, Google Scholar. We searched for studies published from 2015, which is the year Leow (2015) was published, to February 2024, when the literature search for this review started. We combined three keywords ("depth of processing," "second language," and "writing") and used them when searching literature on the database. Initially, through this database search, a total of 1,850 reports were identified.

The retrieved studies were screened based on the following four criteria to determine which to include in this review.

1. Only experimental studies that examined learners' DoP while they received WCF on L2 writing tasks (e.g., essays and blog entries) and predefined operational definitions of DoP were included. Studies employing reformulation and model text were also included if these tools were used to provide WCF on L2 writing.
2. Studies that examined cognitive processing of WCF based on the Noticing Hypothesis were excluded (e.g., Suzuki, 2017).
3. Studies that did not focus on L2 learning were excluded. For example, DeRobles (2019) was excluded because all participants were heritage language learners, not L2 learners, who had grown up in bilingual households either since birth or before the age of six.
4. In cases where the same study was published in separate papers, the paper with the most detailed description of research methods was included, while the other was excluded. For example, a dissertation (Caras, 2017) was included, while a book chapter (Caras, 2019) was excluded.

As a result, a total of 15 studies remained which satisfied all of the criteria. The studies comprised seven journal articles, five book chapters, three dissertations.

### **4.2 Coding criteria and procedures**

The retrieved literature was first coded for the study identifier (i.e., authors and year), and publication type (journal article, book chapter, or dissertation). Furthermore,

eight features were also coded: type of institution, number of participants, independent variables, dependent variables, control group, writing task, type of WCF, data for DoP, levels of DoP (presented in Table 1).

Type of institution represented the type of institution from which the participant's data was collected. It was coded as either university, secondary (i.e., senior high school and junior high school), primary (i.e., primary school), online L2 course, and not mentioned (i.e., no information provided). Number of participants showed the total number of participants in each study.

For the writing task category (i.e., the type of writing tasks employed), six subcategories were identified: essay, decision-making task, blog entry, story retelling task (i.e., reading a story and retelling it without consulting the original text), story based on scenario (i.e., writing a story according to a provided written scenario), and story based on a picture prompt (i.e., recounting a story with a picture prompt depicting the story). Essay refers to a writing task in which participants were instructed to produce a piece of writing about a prompt with a topic. Picture-based and problem-solving tasks, which were originally developed by Gilabert (2007), were categorized into decision-making task.

We classified type of WCF (i.e., the types of feedback used in treatment) into five categories: direct WCF, metalinguistic WCF, indirect WCF, reformulation, and computer-generated WCF. The coding was conducted according to what was described in each study. For instance, "indirect corrective feedback" in Park and Kim (2019) was coded as indirect WCF and "metalinguistic error coding" in Bowles and Gastañaga (2022) was coded as metalinguistic WCF. However, there were three exceptions. A type of WCF named "underlining" (i.e., underlining errors) in Bowles and Gastañaga (2022) was coded as indirect WCF against the following explanation of indirect WCF "the teacher indicates that an error exists but does not provide the correction" (Ellis, 2009, p. 98). "Teacher feedback" in Kaivanpanah et al. (2020) was coded as metalinguistic WCF because it involved indicating what kinds of errors students had made (e.g., grammar, wrong word, and spelling) by using codes. This was in line with the following explanation of metalinguistic WCF: "The teacher provides some kind of metalinguistic clue as to the nature of the error" (Ellis, 2009, p. 98). "Indirect WCF with metalinguistic explanations" in Ma (2020) was also coded as metalinguistic WCF because it pointed out error types and offered metalinguistic explanations of how to correct the errors (but did not provide explicit corrections). Direct WCF, metalinguistic WCF, indirect WCF, and computer-generated WCF were further coded as either focused or unfocused.

The independent variables in the included studies were categorized based on the

interventions conducted in the experiments. Eight types of independent variables were identified: (1) type of WCF (i.e., comparing the impacts of two or more types of WCF on dependent variables); (2) teacher vs. computer WCF (i.e., comparing teacher-written WCF with computer-generated WCF); (3) +/- TA (TA vs. silent conditions); (4) TA, WL, and TA+WL; (5) TA in L1 vs. L2 (i.e., comparing L1 with L2); (6) individual vs. collaborative writing (i.e., comparing individual writing with collaborative one); (7) pen-and-paper vs. computer-mediated environment; and (8) proficiency levels (elementary vs. intermediate).

The dependent variables of each study were first identified using the statements in the research questions. As this review focused on DoP, the terms, which were mentioned in the RQs, “process” (Caras, 2017; Ma, 2020), “cognitively process” (Shao & Liu, 2022), and “levels of processing” (Adrada-Rafael & Filgueras-Gómez, 2019; Park & Kim, 2019) were also coded as DoP. Bowles and Gastañaga (2022) worded accuracy as “accurate error revision,” and Manchón et al. (2020) worded “linguistic accuracy of the rewritten texts,” while Ma (2020) worded “accuracy with le and classifiers.” Therefore, each study was scrutinized, and those that measured the accuracy of the rewrite were coded as revision accuracy, while those that measured the accuracy regarding grammatical items were coded as accuracy of target items.

Additionally, two wording in the RQs from two studies (McBride, 2023; Shao & Liu, 2022) were modified when coded. McBride’s (2023) research question, “How does writing in a traditional pen-and-paper environment versus writing in a computer-mediated environment affect L2 written production in terms of CAF measures?” was coded as CAF in writing. Similarly, a research question in Shao and Liu (2022), “To what extent does WCF foster the acquisition of regular past tense by L2 learners?” was coded as Grammar acquisition.

Data for DoP represented data collection instruments of DoP, which were coded as either TA or WL. Although McBride (2023) and McBride and Manchón (2023) both employed three types of data instruments: TA, WL, and TA + WL, only TA was coded. It was because the data of WL, TA + WL were analyzed based on other theoretical approaches (i.e., the Noticing Hypothesis and engagement). Levels of DoP represented how many levels of DoP were set in each study. Control group was coded either as yes or no to indicate the presence and absence of a control group in each study.

The first and second authors completed the coding. One author was responsible for half of the 15 studies, and the other one conducted the remaining half. We then checked each other’s coding results and resolved all disparities or confusion through discussion.

## **5. Results and discussion**

To answer Research Question 1 (the trends in research methods used in studies investigating DoP), we summarized the research methods of the 15 studies included in this review as shown in Table 1. In the following sections, we examine the results and seek for the answer to Research Question 2, that asked the issues with the research methods used in studies investigating DoP of WCF.

### **5.1 General research methods in DoP studies**

The literature search targeted publications from 2015 to 2024, but as shown in Table 1, all studies were from 2019 or later, except for Caras (2017), suggesting a recent increase in interest in DoP. These included seven journal articles, five book chapters, and three dissertations. Three out of five book chapters (Adrada-Rafael & Filgueras-Gómez, 2019; Cerezo et al., 2019; Park & Kim, 2019) were published in a single handbook, *The Routledge Handbook of Second Language Research in Classroom Learning* (Leow, 2019).

**Table 1***Included Studies*

Study	Pub. type	Institution	<i>n</i>	Independent variables	Dependent variables	Control group	Writing task	Type of WCF	Data for DoP
Abdi Tabari et al. (2023)	Journal	University	80	Type of WCF, +/- TA	Revision quality, DoP	No	Essay	Direct, indirect WCF (all unfocused)	TA
Adrada-Rafael and Filgueras-Gómez (2019)	Book chapter	University	29	TA in L1 vs. L2	Revision accuracy, DoP	No	Story based on a picture prompt	Reformulation	TA
Bowles and Gastañaga (2022)	Journal	University	35	Type of WCF	Revision accuracy, DoP	No	Essay	Direct, metalinguistic, indirect WCF (all unfocused)	TA
Caras (2017)	Dissertation	University	61	Type of WCF	Accuracy of target items, DoP	Yes	Blog entry	Direct, metalinguistic, Indirect WCF (all unfocused)	TA
Cerezo et al. (2019)	Book chapter	University	46	Type of WCF	Revision accuracy, error noticing, DoP	Yes	Essay	Direct, indirect WCF (all unfocused)	WL
Kaivanpanah et al. (2020)	Journal	Not mentioned	60	Teacher vs. computer WCF +/- TA	Quality of writing, DoP	No	Essay	Metalinguistic, computer-generated WCF (all unfocused)	TA
Kim and Bowles (2019)	Journal	University	22	Type of WCF	DoP, error types	No	Essay	Direct WCF (unfocused), reformulation	TA
Leow et al. (2022)	Journal	University	10	Type of WCF	Subsequent compositions, curricular tests, final exam, targeted linguistic feature, over time, DoP	No	Essay	Direct, metalinguistic WCF (all unfocused)	TA
Ma (2020)	Dissertation	Online L2 course	38	Type of WCF	Accuracy of target items, engagement, DoP	No	Story based on scenario	Direct, metalinguistic WCF (all unfocused)	TA
Manchón et al. (2020)	Book chapter	University	118	Individual vs. collaborative writing	Revision accuracy, error detection, DoP	Yes	Decision-making task	Direct WCF (unfocused)	WL
McBride and Manchón (2023)	Book chapter	University	18	TA, WL, and TA+WL	Engagement, Affordance for investigating DoP	No	Decision-making task	Direct WCF (unfocused)	TA
McBride (2023)	Dissertation	University	36	TA, WL, and TA+WL Pen-and-paper vs. computer-mediated environment	CAF in writing, DoP	No	Decision-making task	Direct WCF (unfocused)	TA
Nicolás-Conesa et al. (2019)	Journal	University	46	Type of WCF	Revision accuracy, error types	Yes	Essay	Direct, indirect WCF (unfocused)	WL
Park and Kim (2019)	Book chapter	University	24	Proficiency levels (elementary vs. intermediate)	Revision accuracy, DoP	No	Essay	Indirect WCF (unfocused)	TA
Shao and Liu (2022)	Journal	Secondary	99	Type of WCF	Grammar acquisition, DoP	Yes	Story retelling task	Direct, metalinguistic, indirect WCF (all focused)	TA

As Table 1 illustrates, the majority of the included DoP studies were conducted at universities, with three exceptions. Ma (2020) conducted research in an online Mandarin Chinese course designed specifically for professionals who would live in China. The mean ages of her two experimental groups were 38 and 41, respectively. Shao and Liu (2022) investigated secondary school students. Kaivanpanah et al. (2020) did not provide information regarding the institutional setting. The imbalance in research settings being skewed towards universities has been pointed out in many previous meta-analyses. The studies on DoP reviewed here exhibit this same tendency of being predominantly conducted in university settings. Since DoP is a cognitive process, it is presumed to be influenced by the learner's cognitive levels, calling for research investigating young learners and comparative research on DoP between children and adults.

The number of participants ranged from 10 to 118. Some studies employed a within-group design, in which a single group experienced two different treatments, and their performance in the two treatments was compared. For example, in Kim and Bowles (2019), all 20 participants completed two different essay writing tasks, receiving either direct WCF or reformulation, and the researchers compared the learners' DoP when receiving different types of WCF. Similarly, Bowles and Gastañaga (2022) had 35 participants who wrote three different essays, receiving direct WCF, metalinguistic error coding, and underlining, respectively. Although in all these studies, the essay tasks were counterbalanced to control for task difficulty, the small sample sizes may limit the generalizability of the findings to the broader population.

Other studies had multiple experimental groups. Manchón et al. (2020) had 118 participants divided into two experimental groups, an individual and a collaborative writing condition, as well as a control group. They examined differences in error detection, accuracy of the rewritten writing, and DoP between the conditions. Caras (2017) had three experimental groups: unfocused direct WCF ( $n = 15$ ), unfocused metalinguistic WCF ( $n = 16$ ), and unfocused indirect WCF ( $n = 15$ ), as well as a control group ( $n = 15$ ), and examined the learners' process of WCF.

Regarding writing tasks, eight studies employed essay writing with three variations: argumentative (Kaivanpanah et al., 2020; Kim & Bowles, 2019), narrative (Ma, 2020; Nicolás-Conesa et al., 2019; Tabari et al., 2023), and others (Bowles & Gastañaga, 2022; Cerezo et al., 2019; Park & Kim, 2019). The other three studies (Manchón et al., 2020; McBride, 2023; McBride & Manchón, 2023) employed a decision-making task in which the participants explained a solution to rescue as many people as possible based on a picture of a building on fire. Adrada-Rafael and Filgueras-



Gómez (2019) asked participants to write a story based on a picture prompt. Ma (2020) had them write a story based on a scenario. Caras (2017) adopted a blog entry task in which participants wrote a blog entry on a Spanish-language site and received contextual instructions and a checklist to address specific questions. Shao and Liu (2022) adopted a story-retelling task in which the participants read a story about a past event and reproduced it in their writing without referencing the original story.

## 5.2 Independent variables

Nine out of 15 studies compared the effectiveness of two or more types of WCF, while the remaining six employed one type of WCF to observe DoP when receiving WCF. Among the nine comparative studies, six compared two types, direct and indirect WCF, while three compared three types, including metalinguistic WCF.

Twelve studies included direct WCF in an independent variable. While most of them only employed direct corrections on the erroneous forms, Abdi Tabari et al. (2023) provided explicit correction with metalinguistic WCF. Seven studies included indirect WCF, which was provided in various forms: highlighting (e.g., Abdi Tabari et al., 2023), underlining (e.g., Caras, 2017), and indicated by a circle (e.g., Shao & Liu, 2022). Two of the seven studies, Cerezo et al. (2019) and Nicolás-Conesa et al. (2019), provided metalinguistic codes (i.e., codes representing error category) as a part of indirect WCF. Six studies included metalinguistic WCF, five indicated errors by using error coding. Among them, four studies (Caras, 2017; Leow et al., 2022; Ma, 2020; Shao & Liu, 2022) further complemented error codes with metalinguistic explanations or samples of rules. The remained study, Shao and Liu (2022) provided a brief explanation of the target linguistic rule with direct error corrections.

Two studies (Adrada-Rafael & Filgueras-Gómez, 2019; Kim & Bowles, 2019) employed reformulation, which refers to a native-speaker rewriting students' original compositions to make them sound native-like with bearing in mind to keep the students' ideas much as possible (Ellis, 2009). While Adrada-Rafael and Filgueras-Gómez (2019) provided the participants only with reformulation, Kim and Bowles (2019) compared it with direct WCF. Only one study, Kaivanpanah et al. (2020), employed computer-generated WCF which compared it with teacher-written metalinguistic WCF.

The overview of type of WCF suggests that measuring DoP of WCF was attempted in relatively limited situations: when learners received direct, indirect, and metalinguistic WCF. Research on DoP when given computer-generated WCF or when comparing reformulated writing or model texts with learners' own writing is limited.

All studies except one employed unfocused WCF. Shao and Liu (2022) was the

only study that used focused WCF. This is assumingly because the researchers were interested in the learners' DoP, which might differ according to targeted linguistic features. Caras (2017), for example, examined the levels of DoP for the feedback targeting different grammatical features (i.e., two target dichotomies, which were *ser* versus *estar* and the preterit versus imperfect past tense aspects) and suggested that the levels of DoP differ according to the linguistic features.

Besides impacts of WCF on dependent variables, other interventions such as writing environment (i.e., pen-and-paper vs. computer-mediated environment), how a task is completed (i.e., individual vs. collaborative writing), and proficiency levels (elementary vs. intermediate) were investigated, even though there was only one study for each of the three aspects.

10 out of the 15 studies did not include a control group. As two of them did not investigate the effects of WCF on the improvement of writing and focused entirely on DoP (Kim & Bowles, 2019) or engagement and affordances for inspecting DoP (McBride & Manchón, 2023), they did not need to include a control group. However, the other studies examined the improvements as a result of providing WCF by measuring revision accuracy and grammar acquisition. In such research, including a control group would allow stronger implications (Ellis et al., 2019). Five studies had a control group. In four of these studies, the control groups completed the writing tasks without receiving any WCF (Cerezo et al., 2019; Manchón et al., 2020; Nicolás-Conesa et al., 2019; Shao & Liu, 2022), and in the other study, the control group received commentary on the overall organization and content of their writing (Caras, 2017).

### **5.3 Dependent variables**

For a dependent variable, eight studies measured accuracy. We identified two types of accuracy measures in the studies: six studies investigated revision accuracy, and two assessed the accuracy of specific target items. The former studies include Nicolás-Conesa et al. (2019), which quantified errors within each task and group to assess accuracy. Park and Kim (2019) employed three categories: accurate correction, inaccurate correction, and no correction to investigate participants' self-corrections. Conversely, the latter two studies evaluated accuracy on specific target items. Caras (2017) concentrated on copular verbs, gender agreement in adjectives, and past tense aspect in Spanish, while Ma (2020) focused on *le* and classifiers, two fundamental structures frequently used in Chinese. Only two studies explored the development of L2 proficiency beyond evaluating the accuracy of revisions. These studies assessed the complexity, accuracy, and fluency of learners' new writings (McBride, 2023) or utilized

evaluative tests administered by the educational institution where the study was conducted (Leow et al., 2022) to ascertain whether writing tasks with WCF contributed to enhancements in L2 proficiency.

All the studies included in this review examined DoP as a dependent variable, except for Nicolás-Conesa et al. (2019), which measured DoP but did not specify how the data was utilized. Concerning the measurement of DoP, 10 out of 15 studies categorized DoP into three levels: high, medium (intermediate), and low, following Leow's (2015) framework. Two studies (Kim & Bowles, 2019; Leow et al., 2022) modified Leow's (2015) coding scheme; Kim and Bowles (2019) classified DoP into two levels: high or low, while Leow et al. (2022) employed four levels of DoP: High DoP + Understanding, High DoP – Understanding, Medium DoP + Reporting, and Low DoP + Noticing. The remaining three studies (Cerezo et al., 2019; Manchón et al., 2020; Nicolás-Conesa et al., 2019) classified DoP into five levels. As previously noted, the categorization of DoP varied across studies.

Regarding the data collection method for identifying DoP, 12 out of the 15 studies employed TA, while the other three used WL. These two data collection instruments allow researchers to obtain concurrent introspective data, which refers to real-time self-reflective information of ongoing experiences. Unlike retrospective data, which is gathered after the fact and often relies on memory or existing records to reconstruct past events (e.g., interviews and stimulated recall), concurrent introspective data is more likely to provide real-time insights into cognitive processes (Sasaki, 2013). However, such concurrent introspective data collection has potential issues of reactivity and veridicality; the TA or WL themselves might interfere with the learner's natural behavior or serve as an intervention for the learners (Zhang & Zhang, 2019). In response to these concerns, Abdi Tabari et al. (2023), in the current review, treated TA as an intervention to facilitate learning and made it one of the independent variables in their study. The results suggested that the data collection mode (TA, WL) impacts learners' cognitive processes and contributes to learning. Taking further steps, McBride and Manchón (2023) involved three experimental groups (i.e., TA, WL, and TA+WL) to examine how learners' engagement and DoP differ in each data collection mode. These studies suggest that researchers should be mindful of the potential influence of TA or WL on data of learners' cognitive processes when employing these data collection methods. TA language (i.e., whether L1 or L2) was also expected to have impacts on learners' DoP in one study (Adrada-Rafael & Filgueras-Gómez, 2019) in this review, however, no other study investigated this aspect.

The methods employed for TA practice tasks varied across studies. Four studies

utilized a math calculation task (Adrada-Rafael & Filgueras-Gómez, 2019; Caras, 2017; Leow et al., 2022; Park & Kim, 2019), while three studies (Bowles & Gastañaga, 2022; Kim & Bowles, 2019; Tabari et al., 2023) incorporated a practice task similar to the actual task in which participants engaged in TA while receiving WCF. Considering the actual task participants were expected to undertake during TA, the more closely the practice task resembles the actual task, the more effectively participants may approach the latter. This aligns with Bowles's (2010) observation that although arithmetic tasks have the advantage of being non-verbal and thus not interfering with the practice itself, they may be challenging to apply in an actual task. Instead, he suggested that verbal warm-up tasks that resemble the actual task enable participants to verbalize more readily while engaging in the actual task. The remaining studies, such as Kaivanpanah et al. (2020), Ma (2020), McBride (2023), and McBride and Manchón (2023), provided instructions on TA without practice, while Shao and Liu (2022) employed a reading comprehension task as a TA practice alongside instructions.

WL has also been used in previous research to inspect cognitive activity while processing WCF (e.g., Suzuki, 2012, 2017) and was mentioned as a good option for collecting data in a real classroom without disrupting the class dynamics (Nicolás-Conesa et al., 2019). In this review, three studies (i.e., Cerezo et al., 2019; Manchón et al., 2020; Nicolás-Conesa et al., 2019) employed it as a data collection instrument of DoP. In their WL phase, the three studies asked participants to correct errors or reflect on errors already provided on their compositions, and to provide a linguistic explanation for each error and corresponding correction in a form. Two of the three studies, Cerezo et al. (2019) and Nicolás-Conesa et al. (2019) provided the following prompt (adopted from Suzuki, 2012); "Have a look at your essay and identify your errors. Why is each linguistic form incorrect? Explain it." Languages used in WL activities also varied among studies. Manchón et al. (2020) instructed them to do it in L2 (English) but permitted them to use Spanish (most of the participants' L1) if necessary, and Cerezo et al. (2019) also employed participants' L2 (i.e., English). However, no information of language in WL was provided in Nicolás-Conesa et al. (2019).

## **6. Conclusion**

In this scoping review, we focused on the research methods employed in studies examining DoP. The review provided an overview of the existing studies investigating learners' DoP when receiving WCF. It highlighted some research gaps in the existing literature, such as targeting learners other than university students, computer-mediated writing contexts, or writing for specific purposes. The majority of the studies included

in this review employed commonly used comparisons in the field (i.e., direct, indirect, and metalinguistic WCF) and lacked exploration of other types, such as reformulation and computer-generated WCF. Research using WL is still scarce.

This review also identified some limitations in the research methods used in the DoP experiments, such as reliance on within-group designs, small sample sizes, and a lack of control groups. Another issue highlighted in this review is the potential influence of TA when examining DoP. As Shao and Liu (2022) argue, TA may serve not only as a means of obtaining introspective data but also as a potential educational intervention.

As this is a preliminary review, we acknowledge that there are more aspects to analyze in the literature to fulfill the purpose of the scoping review. The most important aspect to review is the way the existing literature differentiates the levels of DoP. As shown in the current review, studies vary in terms of how many levels they set when examining learners' DoP. The conceptual underpinnings of the DoP levels can be explored by conducting an in-depth analysis of the DoP categories in individual studies. Such analysis would provide a conceptual scope of the notion of DoP. We intend to further pursue such analysis.

While limited, we believe this review demonstrated the usefulness of scoping reviews in providing an overview of the existing literature, highlighting potential gaps, and future research directions.

## References

- Adrada-Rafael, S., & Filgueras-Gómez, M. (2019). Reactivity, language of think-aloud protocol, and depth of processing in the processing of reformulated feedback. In R. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 199–211). Routledge.
- Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development* (Vol. 96). Multilingual Matters.
- Bowles, M. A. (2010). *The think-aloud controversy in second language research*. Routledge. <https://www.routledge.com/The-Think-Aloud-Controversy-in-Second-Language-Research/Bowles/p/book/9780415994842>
- Bowles, M. A., & Gastañaga, K. (2022). Heritage, second and third language learner processing of written corrective feedback: Evidence from think-alouds. *Studies in Second Language Learning and Teaching*, 12(4), 675–696. <https://doi.org/10.14746/ssllt.2022.12.4.7>
- Caras A. (2019). Written corrective feedback in compositions and the role of depth of processing. In R. Leow (Ed.), *The Routledge handbook of second language research*

- in classroom learning* (pp. 186–198). Routledge.
- Caras, A. M. (2017). “*Why are these underlined?*” *Depth of processing and type of written corrective feedback in L2 Spanish compositions* (Publication No. 10601613) [Doctoral dissertation, Georgetown University]. ProQuest Dissertations and Theses Global.
- Cerezo, L., Manchón, R. M., & Nicolás-Conesa, F. (2019). What do learners notice while processing written corrective feedback? A look at depth of processing via written languaging. In R. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 171–185). Routledge.
- Chong, S. W., & Reinders, H. (2022). Autonomy of English language learners: A scoping review of research and practice. *Language Teaching Research*, 0(0). <https://doi.org/10.1177/13621688221075812>
- Coyle, Y., & Cánovas Guirao, J. (2019). Learning to write in a second language: The role of guided interaction in promoting children’s noticing from model texts. *CLIL Journal of Innovation and Research in Plurilingual and Pluricultural Education*, 2(1), 21. <https://doi.org/10.5565/rev/clil.22>
- DeRobles, G. (2019). *The effects of type of written corrective feedback and level of proficiency on processing and accuracy in heritage language learners of Spanish* (Publication No. 13899079) [Doctoral dissertation, Georgetown University]. ProQuest Dissertations and Theses Global.
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63(2), 97–107. <https://doi.org/10.1093/elt/ccn023>
- Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). *Task-Based Language Teaching: Theory and Practice*. Cambridge: Cambridge University Press.
- Hanaoka, O. (2007). Output, noticing, and learning: an investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 11(4), 459–479. <https://doi.org/10.1177/1362168807080963>
- Hanaoka, O., & Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing*, 21(4), 332–347. <https://doi.org/10.1016/j.jslw.2012.09.008>
- Hanzawa, K., Suzuki, Y., Yanagisawa, A., & Fukuta, J. (2024, March 15–19). *A scoping review on task repetition studies focusing on speaking ability*. [Paper presentation]. American Association for Applied Linguistic Conference, Houston, Texas, United States. <https://www.xcdsystem.com/aaal/program/SqDM5Li/index.cfm>
- Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34(2), 239. <https://doi.org/10.2307/3587952>

- Kaivanpanah, S., Alavi, M., & Meschi, R. (2020). L2 writers' processing of teacher vs. computer-generated feedback. *Journal of English Language Teaching and Learning*, 12(26), 175–215. <https://doi.org/10.22034/elt.2020.11472>
- Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, 99(1), 1–18. <https://doi.org/10.1111/modl.12189>
- Kim, H. R., & Bowles, M. (2019). How deeply do second language learners process written corrective feedback? Insights gained from think-alouds. *TESOL Quarterly*, 53(4), 913–938. <https://doi.org/10.1002/tesq.522>
- Lázaro-Ibarrola A., Hidalgo M. Á. (2021). Give me a second chance: Task repetition and collaborative writing with child EFL learners. *Language Teaching for Young Learners*, 3, 275–299. <https://doi.org/10.1075/ltyl.20009.laz>
- Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. Routledge.
- Leow, R. P., Thinglum, A., & Leow, S. A. (2022). WCF processing in the L2 curriculum: a look at type of WCF, type of linguistic item, and L2 performance. *Studies in Second Language Learning and Teaching*, 12(4), 651–673. <https://doi.org/10.14746/ssllt.2022.12.4.6>
- Lightbown, P. M., & Spada, N. (2013). *How languages are learned* (4th ed.). Oxford University Press.
- Lindgren, E., & Sullivan, K. P. H. (2003). Stimulated recall as a trigger for increasing noticing and language awareness in the L2 writing classroom: A case study of two young female writers. *Language Awareness*, 12(3–4), 172–186. <https://doi.org/10.1080/09658410308667075>
- Ma, X. (2020). *Writing in a task-based individualized curriculum: Effectiveness of direct and indirect written corrective feedback* (Publication No. 28089503) [Doctoral dissertation, Georgetown University]. ProQuest Dissertations and Theses Global.
- Manchón, R. M., Nicolás-Conesa, F., Cerezo, L., & Criado, R. (2020). L2 writers' processing of written corrective feedback. In R. M. Manchón (Ed.), *Languaging in language learning and teaching: A collection of empirical studies* (pp. 241–263). John Benjamins Publishing Company.
- McBride, S. (2023). *The effects of written corrective feedback processing conditions on L2 learners' written texts and levels of depth of processing* [Doctoral dissertation, University of Murcia]. <https://doi.org/10.13140/RG.2.2.10891.46881>
- McBride, S., & Manchón, R. M. (2023). Analysing L2 writers' processing of written corrective feedback via written languaging and think-aloud protocols. In R. M.

- Manchón & J. R. de Larios (Eds.), *Research methods in the study of L2 writing processes* (Vol. 5, pp. 337–514). John Benjamins Publishing Company.
- Nicolás-Conesa, F., Manchón, R. M., & Cerezo, L. (2019). The effect of unfocused direct and indirect written corrective feedback on rewritten texts and new Texts: looking into feedback for accuracy and feedback for acquisition. *The Modern Language Journal*, 103(4), 848–873. <https://doi.org/10.1111/modl.12592>
- Park, E. S., & Kim, O. Y. (2019). Learners' engagement with indirect written corrective feedback: Depth of processing and self-correction. In R. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 212–226). Routledge.
- Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing*, 10(4), 277–303. [https://doi.org/10.1016/S1060-3743\(01\)00046-7](https://doi.org/10.1016/S1060-3743(01)00046-7)
- Sasaki, M. (2013). Introspective Methods. In A. J. Kunnan (Ed.), *The Companion to Language Assessment, First Edition*. (pp. 1340–1357). Wiley.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158. <https://doi.org/10.1093/applin/11.2.129>
- Shao, J., & Liu, Y. (2022). Written corrective feedback, learner-internal cognitive processes, and the acquisition of regular past tense by Chinese L2 learners of English. *Applied Linguistics Review*, 13(6), 1005–1028. <https://doi.org/10.1515/applirev-2019-0131>
- Shintani, N., & Ellis, R. (2015). Does language analytical ability mediate the effect of written feedback on grammatical accuracy in second language writing? *System*, 49, 110–119. <https://doi.org/10.1016/j.system.2015.01.006>
- Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning*, 64(1), 103–131. <https://doi.org/10.1111/lang.12029>
- Simard, D., Guénette, D., & Bergeron, A. (2015). L2 learners' interpretation and understanding of written corrective feedback: Insights from their metalinguistic reflections. *Language Awareness*, 24(3), 233–254. <https://doi.org/10.1080/09658416.2015.1076432>
- Suzuki, W. (2012). Written languaging, direct correction, and second language writing revision. *Language Learning*, 62(4), 1110–1133. <https://doi.org/10.1111/j.1467-9922.2012.00720.x>
- Suzuki, W. (2017). The effect of quality of written languaging on second language learning. *Writing and Pedagogy*, 8(3), 461–482. <https://doi.org/10.1558/wap.27291>



- Tabari, M. A., Sato, M., & Wang, Y. (2023). Engagement with written corrective feedback: Examination of feedback types and think-aloud protocol as pedagogical interventions. *Language Teaching Research*, 0(0). <https://doi.org/10.1177/13621688231202574>
- Tulloch, B., & Ortega, L. (2017). Fluency and multilingualism in study abroad: Lessons from a scoping review. *System*, 71, 7–21. <https://doi.org/10.1016/j.system.2017.09.019>
- Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62(1), 1–41. <https://doi.org/10.1111/j.1467-9922.2011.00674.x>
- Zhang, L. J., & Zhang, D. (2019). Think-aloud protocols. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 302–312). Routledge.