国立国語研究所学術情報リポジトリ

Comparison about Structures of Complex Sentence between Textbooks of Japanese Language and Composition written by High School Students : Quantitative Comparison about Structures of Subordinate and Connection forms

メタデータ	言語: jpn 出版者: 公開日: 2018-03-20 キーワード (Ja): キーワード (En): 作成者: 松本, 理美, MATSUMOTO, Satomi メールアドレス: 所属:
URL	https://doi.org/10.15084/00001452

国語教科書と高校生作文の複文構造比較
－従属節の構造と接続形式の量的比較－

松本 理美（立命館大学大学院）

Comparison about Structures of Complex Sentence between Textbooks of Japanese Language and Composition written by High School Students: Quantitative Comparison about Structures of Subordinate and Connection forms

Satomi Matsumoto (Ritsumeikan University)

要旨

コーパスという言語資源を活用した文体研究は、語彙、品詞、文法などに関するものなど、数多く見られるが、複文構造や従属節の研究への活用が十分であるとは言えない。これは、現時点で、解析器による従属節への情報付与技術の発展に対し、データ分析技術の普及が追いついていないことも起因していると考える。また、複文に着目した文体研究において、高校生作文や学校教科書を対象としたものは、管見の限りない。そこで、本研究では、文章中の従属節に着目し、各種学校の国語教科書と高校生作文における文体特徴を比較することを試み、文章カテゴリーごとに従属節の出現割合を求め、副詞節については、意味別に接続形式を出現頻度でランキングした。従属節の分析からは、国語教科書と高校生作文において、名詞修飾節と副詞節の出現割合に大きな差が見られ、副詞節の意味別接続形式ランキングからも文体特徴を捉えることができた。

1. はじめに

言語に関する研究において、昨今の多種多様なコーパス開発の恩恵に与り、先達が果てしない時間と労力をかけた語や用例等の分析が、大量の言語資源を活用して瞬時に行えるようになってきている。コーパス言語学の発展とともに、研究の対象は文字、語彙から文法まで、書き言葉から話し言葉まで、多種多様な領域に渡り、調査・分析が進められている。

解析器により付与された節情報を利用して連用節の出現を定量的に分析した丸山 (2014) は、コーパスを利用した調査が「母語話者の内省では知りえない言語事実を実証的に明らかに」（丸山 2014：402）できるという利点を挙げている。

一方で、石黒（2016）は、昨今のコーパス言語学の隆盛を歓迎しながらも、敢えて小規模な言語データベースにより、接続詞に着目した文体研究を行った。ここで石黒が指摘する通り、「コーパス・ネイティブ」（石黒 2016：161）の研究者の時代にして、地道な手作業による研究により、「見落とされている言語事実」（石黒 2016：161）が新たに発見されることもまた十分にあり得ると考える。

そこで、本研究では、コーパス開発の現段階では、まだ人手による作業に頼るところもある複文の従属節、特に副詞節の接続形式に着目して、文体特徴を捉えることを試みる。

そして、その対象として、日本語研究ではもとより、言語教育の分野でも谷間的存在であり、研究対象となることが少ない日本語母語、非母語の両高校生の作文を取り上げる。これは、以下の二つの理由からである。一つは、国際化が急進する現代において、日本語学習者としてではなく、生活者として来日を余儀なくされた年少者の日本語が、日本語のバリエーションの一つとして研究対象になり得るのではないかということである。もう一つは、日本人高校生を含め、「不完全な書きことば」から日本語を捉えなおすことで、新たな発見があるのではないかと考えたからである。

2. 研究目的

本研究では、文章中の複文割合や副詞節の接続形式の出現頻度が文章カテゴリーにより異なることを明らかにし、これらにより文体特徴を捉えることができること、またこれらが文章レベルの指標となり得ることを示すことを目的とする。

3. 研究方法

高校生の作文（日本語母語話者、日本語非母語話者[1]）、小学校教科書、中学校教科書、高校教科書の文章について、手作業により、従属節に機能分類と意味分類のタグ付けを行う。

本研究では、益岡・田窪 (1992) に基づき、述語を中心とする文節のひとまとまりを「節」と定義する。ただし、単独で述語の役割と名詞を修飾する役割を持っている形容詞については、補語を伴わずに名詞を修飾している際には節とみなさないものとする。以下に例文を挙げて説明する。

(1) 「赤いマフラーを巻いている少女は、私の妹だ。」

例文(1)の下線部「赤い」は形容詞であり、節ではない。

(2) 「ふさが赤いマフラーを巻いている少女は、私の妹だ。」

例文(2)の下線部「ふさが赤い」は「ふさが」という補語を伴うため、「マフラー」に係る名詞修飾節とする。

従属節については、益岡・田窪 (1992) を元に従属節を詳細に分類した池原 (2009) に従い、意味分類体系[2]1 段目の補足節・名詞修飾節（連体節）・副詞節・並列節の 4 つに分類する。また、副詞節については、意味分類体系 2 段目の 16 種まで分類する。

16 分類した副詞節のうち、本研究で着目するのは、時を表す副詞節、原因・理由を表す副詞節、条件・譲歩を表す副詞節、付帯状況・様態を表す副詞節、逆接を表す副詞節の 5 つの副詞節[3]である。本研究では、これらの節を、順に時間節、原因・理由節、条件・譲歩節、付帯状況・様態節、逆接節と呼ぶことにする[4]。

---

[1] 日本語母語話者の高校生を日本人高校生、日本語非母語話者の高校生を外国人高校生とする。

[2] 池原 (2009) の第 8 章付録「3.主節従属節間の意味分類体系」では、従属節を意味機能により 1 段から 4 段まで順に細かく 4 段階に分類しており、1 段目は 4 種、2 段目は 27 種、3 段目は 37 種、4 段目は 154 種に分類している。例えば、「このワインはおいしいだけに値段が高い。」という文の従属節「このワインはおいしいだけに」は、1 段目は副詞節、2 段目は因果関係、3 段目は原因、4 段目は特定原因に分類される。なお、1 段目は文法的、機能的な分類であり、2 段目以降は、意味や特徴に基づいた詳細分類である。

[3] 本研究では調査対象とした文章データの全てに出現した従属節に着目する。

[4] 本研究の意味分類の基準とした池原 (2009) では、時間節を時、原因・理由節を因果関係としているが、一般的呼称を採用し、このように定義する。

以上の従属節についてタグ付けを行い、文章カテゴリーごとに計量を行う。

対象とした文章は、本研究の調査協力校より回収した高校生作文 43 編（日本人 10 編、外国人 33 編）と、小学校教科書、中学校教科書、高校教科書から各 2 編である。

4. 結果と考察

4.1 従属節について

文章中の従属節を機能による４分類にタグ付けし、文章カテゴリーごとに各種従属節の数を計量したものを、表 1 に示す。

表 1 文章カテゴリーごとの各種従属節数 （ ）内は割合(%)

	補足節数	名詞修飾節数	副詞節数	並列節数	従属節の総数
外国人高校生	77 (18)	45 (10)	244 (56)	71 (16)	437
日本人高校生	55 (27)	23 (11)	90 (44)	37 (18)	205
小学校教科書	59 (29)	49 (24)	72 (36)	23 (11)	203
中学校教科書	88 (33)	87 (32)	77 (28)	19 (7)	271
高校教科書	96 (29)	134 (41)	80 (25)	15 (5)	325

文章カテゴリーにより、文の数、従属節の数が異なるため、表 1 に基づき、各種従属節が従属節全体に占める割合を求め、グラフにしたものを図 1 に示す。

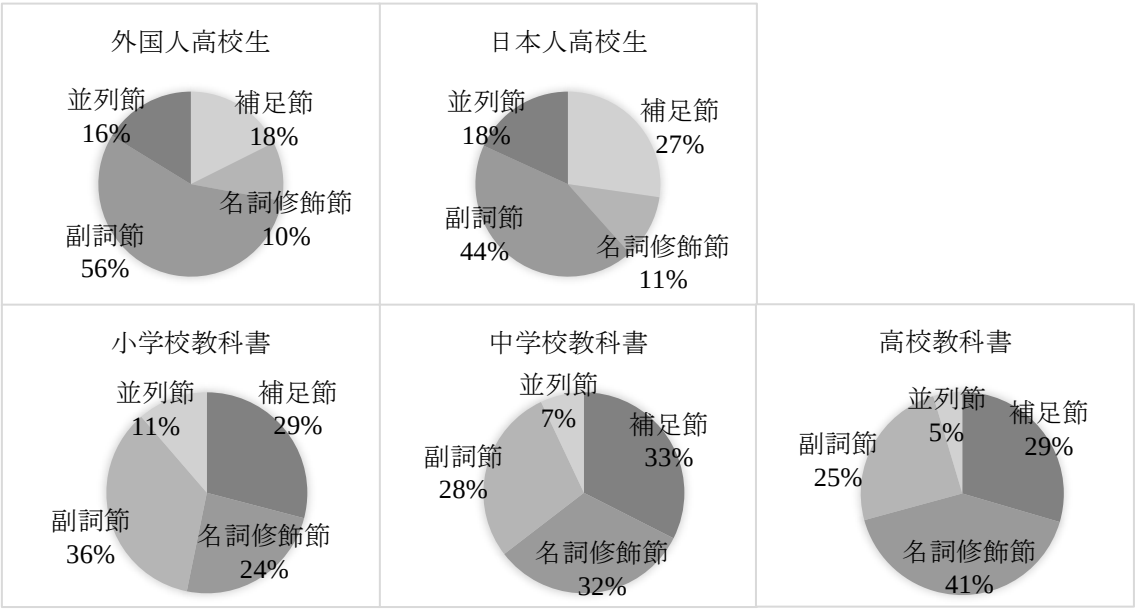

図 1 文章カテゴリーごとの各種従属節の割合

日本人高校生作文と外国人高校生作文を高校生作文群、小学校教科書、中学校教科書、高校教科書を教科書群とすると、高校生作文群と教科書群には、各種従属節の出現割合に大きな違いが認められる。特に大きな差は、補足節、名詞修飾節を併せた連体系の節と副

詞節、並列節を併せた連用系の節の間に、はっきりと現れている。高校生作文群は教科書群に比べ、連体系の節が少なく、連用系の節が多いことが明らかになった。

また、教科書群だけをみると、学年の上昇とともに名詞修飾節が増加し、副詞節、並列節が減少する傾向にあることも確認できた。

これらから、文章における従属節の割合が文章レベルを測る指標となる可能性が示唆できる。

4.2　副詞節のついて

16 種類に意味分類した副詞節のうち、本研究で分析対象とした全ての文章中に出現する副詞節に着目し、文章カテゴリーごとの意味別出現頻度を求めた。着目した副詞節は、時間節、原因理由節、条件・譲歩節、付帯状況・様態節、逆接節である。文章中に出現する頻度は以下のとおりである。

表 2　文章カテゴリーごとの副詞節の意味別出現頻度（　）内は割合(%)

	時	原因	条件	付帯	逆接	他	合計
外国人高校生作文	54 (22)	50 (20)	12 (5)	26 (11)	38 (16)	64 (26)	244
日本人高校生作文	12 (13)	33 (37)	8 (9)	17 (19)	10 (11)	10 (11)	90
小学校教科書	9 (12)	15 (21)	13 (18)	22 (31)	2 (3)	11 (15)	72
中学校教科書	9 (12)	12 (16)	17 (22)	14 (18)	7 (9)	18 (23)	77
高校教科書	12 (15)	7 (9)	20 (28)	16 (17)	4 (5)	21 (26)	80

表 2 で示した割合を、図 2 のグラフに示す。[5]

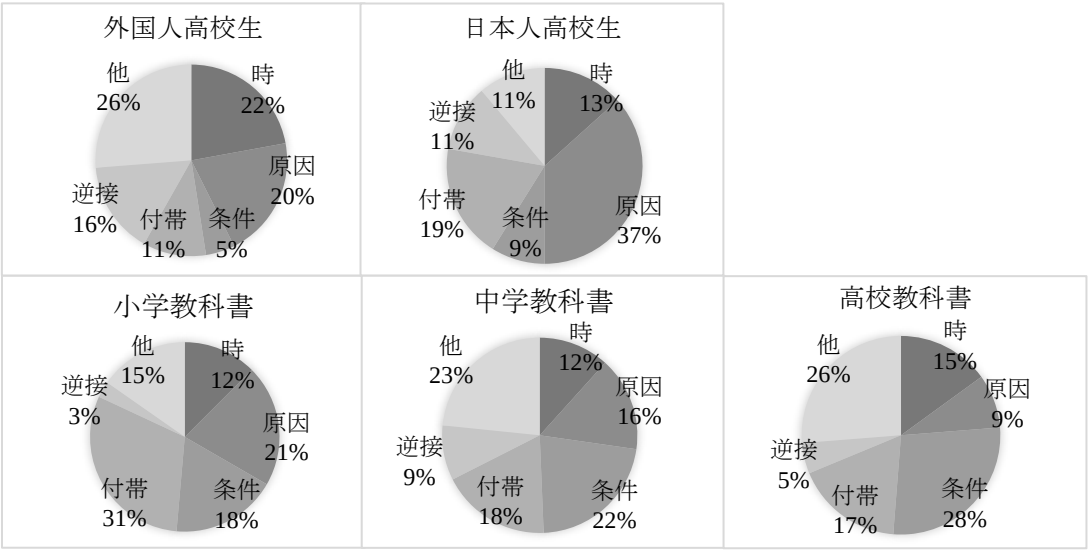

図 2　副詞節の意味別出現割合

[5] 表 2、図 2 では、時間節を「時」、原因・理由節を「原因」、条件・譲歩節を「条件」、付帯状況・様態節を「付帯」、逆接節を「逆接」と略して表記した。

ここで注目すべきは、条件節と逆接節である。高校生作文群において非常に少ない条件節が、教科書群では、学年上昇とともに増加していること、逆接節は、教科書群よりも高校生作文群に多く見られることが確認できた。

教科書群においては、学年上昇とともに、原因理由節、付帯状況・様態節の減少傾向が見られた。

このように、副詞節の意味分類によっても、文章カテゴリーの特徴が捉えられた。

4.3　副詞節の接続形式

本節では、文章カテゴリー別に、副詞節の接続形式について、意味に関係なく、節末の形式にのみ着目した分析を行う。4.3.1 では、石黒 (2016) を援用し、接続形式に見られる文体特徴を、文章カテゴリー間の共通点、相違点から探ることを意図し、文章カテゴリーごとに接続形式のランキングを行い、上位 3 位までを示す[6]。また、4.3.2 では、前項で、全文章カテゴリーにおいて頻度ランキング 1 位となった接続形式「て」に着目する。それぞれの文章カテゴリーにおいて、副詞節の接続形式「て」の意味をどのような割合で分担しているか、文章カテゴリーごとに意味分担割合を求め、その結果について考察する。

4.3.1　文章カテゴリー別接続形式の出現頻度ランキング

本項では文章カテゴリー別に接続形式の出現頻度に着目してランキングを行い、上位 3 位を示す。上位 3 位とした根拠は、データ数が多くないため、出現頻度が少ない接続形式も多く、下位になると出現頻度が「2」や「1」で同順位となる接続形式が多かったためである。なお、カッコ内には、カテゴリー中の副詞節全ての接続形式に占める割合を示す。

表 3　文章カテゴリー別　接続形式出現頻度ランキング　（カッコ内は割合）

外国人高校生	
て	60 (32%)
けど	29 (15%)
から	24 (13%)

日本人高校生	
て	32 (40%)
ので	9 (11%)
けど	6 (7%)

小学教科書	
て	13 (21%)
ように	9 (15%)
から	5 (8%)
と	5 (8%)
ながら	5 (8%)

中学教科書	
て	9 (15%)
が	6 (10%)
から	6 (10%)
と	6 (10%)

高校教科書	
て	8 (13%)
ば	7 (11%)
連用中止	6 (10%)

意味に関係なく出現頻度での節続形式ランキングの上位 3 位を表 3 に示した。全ての文章カテゴリーにおいて、接続形式「て」が 1 位を占めているが、2 位との差を見ると、教科書群と高校生作文群では、その出現傾向が明らかに異なっていることがわかる。

[6] 石黒 (2016) を援用し、節形式のランキングを行うことで、文章カテゴリーの文体特徴を捉える試みを行う。石黒 (2016) のように、同じ節形式について、それぞれの文章カテゴリーでの順位を確認したいと考えたが、データ数が少な過ぎたために、今回はランキングだけを行う。

また、教科書類ではほとんど出現しない接続形式「けれど」[7]や、教科書群では学年上昇に伴い出現頻度に減少が見られる接続形式「たら」「から」「ので」[8]が、高校生作文群ではいずれも上位にランキングされていることも確認できた。

この結果から、文章レベルと接続形式に相関がある可能性が示唆された。

4.3.2　副詞節の接続形式「て」について

次に、前項の接続形式の出現割合（頻度）ランキングで、いずれの文章カテゴリーにおいても 1 位であった接続形式「て」に着目して、文章カテゴリーごとに、接続形式「て」が分担する意味について分析を行う。文章カテゴリーごとに、接続形式「て」が分担する意味の割合を求め、考察を行う。

表 3 から、学年上昇に伴い、接続形式「て」の出現が減少していることが読み取れる。図 3 は、外国人高校生作文、日本人高校生作文、小学校教科書、中学校教科書、高校教科書の中で、接続形式「て」が時間節、原因理由節、付帯状況・様態節において、どのような割合でその意味を分担しているかをグラフにしたものである。

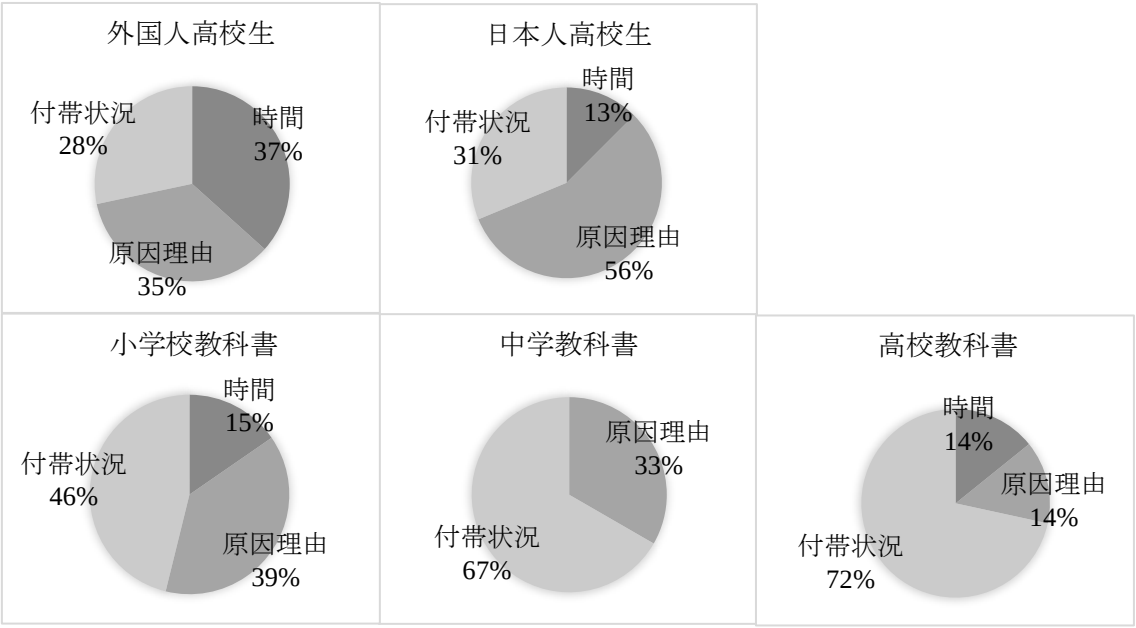

図 3　接続形式「て」が分担している意味

接続形式「て」の出現割合が高かった高校生作文群についてみると、外国人高校生の作文では、時間節、原因理由節、付帯状況節の三つの節において、ほぼ同じ割合で接続形式「て」を使用しているのに対し、日本人高校生の作文では、原因理由節での使用が多く、時間節としてはほとんど使用されていないことがわかる。

---

[7]「けれども」や、高校生作文群での「けど」「けども」を含む。

[8]「たら」は外国人高校生 4 位、日本人高校生 5 位、小学校教科書 9 位、中学校教科書 10 位、高校教科書出現なし、「から」は外国人高校生 3 位、日本人高校生 5 位、小学校教科書 3 位、中学校教科書 2 位、高校教科書 13 位、「ので」は外国人高校生 8 位、日本人高校生 2 位、小学校教科書 6 位、中学校教科書 7 位、高校教科書 13 位であった。

教科書の接続形式「て」を見ても、時間節での使用が極めて少なく、学年が上がるほどに付帯状況・様態節としての使用に比重が移っていくことが明らかになった。

このように、接続形式「て」が分担している意味が、文章カテゴリーによって異なることが確認され、これによっても、文体特徴がとらえられることが示された。

5. おわりに

本研究では、高校生作文群（外国人、日本人）と教科書群（小学校、中学校、高校）の複文構造や副詞節の意味、接続形式に着目し、それぞれの文体特徴を捉えることを試みた。

高校生作文群は教科書群に比べ、従属節の名詞修飾節の出現割合が少なく、副詞節、並列節の出現割合が多いことが明らかになった。教科書間の比較でも、学年上昇に伴い名詞修飾節が増加し、副詞節、並列節が減少する傾向が確認できた。

また、副詞節の意味別出現割合では、高校生作文群は教科書群と比較し、条件節の出現割合において少なく、逆接節において多いという結果が得られた。

そして、副詞節の接続形式に着目した分析において、高校生作文群で出現頻度ランキング上位の「けれど」「たら」「から」「ので」などは、教科書群での出現が少ないなど、副詞節の接続形式の出現頻度が文章カテゴリーによって異なることも明らかになった。

さらに、意義同形の副詞節接続形式が分担する意味が、高校生作文群と教科書群では異なることが確認された。特に、外国人高校生作文では、接続形式「て」が、時間節、原因理由節、付帯状況節において、ほぼ同じ割合で出現するのに対し、高校教科書では接続形式「て」の 72%が付帯状況を表す副詞節の接続形式であることが確認された。

以上の結果から、複文構造や副詞節の意味、接続形式の出現割合によって、文体特徴を捉えることが可能であり、これらには文章レベルを測る指標となる可能性があることが示唆される。

文構造の量的比較を文章レベルに関連づけた研究としては、柴崎 (2009) [9]や、石崎・伊佐原 (1988)[10] の研究などがあり、バトラー後藤 (2011) は、教科ごとの言語的な特徴を捉えることは、児童生徒への教科指導に直接役に立つばかりでなく、教科書の執筆者にも学年に応じたわかりやすい文章の執筆への配慮を促す際の、具体的な情報として有益であると述べている。

さらに、バトラーは、年少者のリテラシーや文法、とりわけ複文構造の問題については、教科書における言語表現のレベルの移行などを考えることにおいても、今後更なる研究が必要な課題であることを指摘している。

このように、多面的、多角的に文体を捉えることは、日本語学、日本語教育、国語教育など多くの分野の研究に資するものであり、本研究は、研究方法、研究対象ともに、意義のある研究であると考える。

---

[9] 石川他 (2010) は、「柴崎 (2009) は、小学校 1 年生から中学 3 年生までの国語教科書コーパスを用いて、６つの変数を手掛かりに、学年レベルの予測」を行ったことを報告し、その結果から「予測学年氏は実学年に充分近く、信頼性が高いために、小中学生を対象とした図書・国語教材の選択に役立つ」と述べている。

[10] 小学校、中学校、高等学校の教科書について、教科書の複文率が 75%になっていること、学齢が上がるほどに、用言数、埋め込み構造数、並列構造数などが増えていく傾向にあることを示している。

しかし、高校生作文、教科書の文章ともに分析したデータが少なく、筆者の経験、知識の不足からも十分な結果が得られたとはいい難い。引き続き、高校生作文データの集積を行い、教科書を含め、分析データを増進することを今後の課題とし、複文構造から文体を捉える研究を発展させたいと考える。

文　献

池原悟 (2009) 『非線形言語モデルによる自然言語処理』 岩波書店

石黒圭 (2016)「社会科学専門文献の接続詞の分野別文体特性」 庵功雄他編『日本語文法研究のフロンティア』pp.161-182. くろしお出版

石崎俊・伊佐原均 (1988)「日本語文の複雑さの定性的・定量的特徴抽出」『自然言語処理』67:6, pp.1-8.

柴崎秀子 (2009)「日本語リーダビリティー測定尺度の構築とソフトウエアの実用化」『科学研究費補助金研究成果報告書』基盤研究(B), 2007~2008, 研究番号 19300277

https://kaken.nii.ac.jp/ja/file/KAKENHI-PROJECT-19300277/19300277seika.pdf（2017 年 12 月 10 日確認)

バトラー後藤裕子 (2011) 『学習言語とは何か―教科学習に必要な言語能力―』三省堂

益岡隆志・田窪行則 (1992) 『基礎日本語文法』くろしお出版

丸山岳彦 (2014) 「現代日本語の連用節とモダリティ形式の分布－BCCWJ に基づく分析－」 益岡隆志他編『日本語複文構文の研究』pp.399-425. ひつじ書房

データとして使用した文献

一川誠 (2015) 「時計の時間と心の時間」 『国語 6 創造』 36-41 東京：光村図書出版株式会社

今道友信 (2016) 「温かいスープ」 『国語 3』 236-239 東京：光村図書出版株式会社

清水哲郎 (2014) 「死と向き合う」 『精選現代文 B』 406-411 東京：筑摩書房

中村桂子 (2015) 「生き物はつながりの中に」 『国語 6 創造』 226-229 東京：光村図書出版株式会社

鷲田清一 (2014) 「ふわふわ」 『精選現代文 B』 134-140 東京：筑摩書房

鷲田清一 (2016) 「誰かの代わりに」 『国語 3』 198-203 東京：光村図書出版株式会社