中国語母語話者による日本語名詞句における「の」の脱落の要因
—コーパスから見る日本語名詞句の出現頻度と学習者の親密度からの考察—

王　婧瑶

1. 研究背景

名詞修飾としての日本語の「の」と中国語の「的」の用法は類似しているが、常に対応しているわけではないため、中国語を母語とする日本語学習者(以下、CLJ)にとっては、名詞修飾の「の」の習得は容易ではない。奥野(2005)は、CLJでは、「日本友達」、「車ドア」のように、「N1+の+N2(中国語:N1+N2)」における「の」の脱落が生じることがあると指摘している。その要因としては、松田他(2006)では、母語の影響の可能性が示されており、また、奥野・金(2010)では、母語の影響の可能性以外にも言語処理ストラテジーの可能性などが取り上げられている。

奥野・王(2019)は日本語の「の」と中国語の「的」における双方向の調査データを分析している。その結果、身近な名詞句であればあるほど、CLJ による「の」の脱落が少なくなり、また、身近な名詞句においても、名詞句によって CLJ による「の」の脱落状況も異なることが見られたため、日本語名詞句の出現頻度、CLJ による日本語名詞句の親密度が CLJ における「の」の脱落と関連することが示唆されている。

このように、奥野・王(2019)は、CLJによる「の」の脱落の要因について、今までの先行研究の観点と異なり、日本語名詞句の出現頻度、CLJによる日本語名詞句の親密度を考慮する必要性について指摘している。そこで本研究では実際に、CLJによる「の」の脱落は、この両者と本当に関係があるか、もし関係がある場合、どのような関係があるか、また、母語の影響の可能性をどのように捉えていけばよいかを、検討する。

2. 先行研究及び問題点

CLJ による「の」の脱落と、日本語名詞句の出現頻度、CLJ による日本語名詞句の親密度との関連性を明らかにするために、日本語の「の」と中国語の「的」の日中対照研究、CLJ による「の」の習得研究、言語心理学における英単語の頻度・親密度の研究及び CLJ による日本語親密度の先行研究を概観する。そこからまとめた問題点に基づき、本研究の位置づけを示す。

2.1　日本語の「の」と中国語の「的」の対照研究

張(2009)は修飾部と被修飾部の意味関係を12ケースに分類し、日本語では「の」が必要であるが、中国語では「的」が不要な名詞句を以下の5ケースに分けている。

① 修飾部がヘッドの擬似所有者であるケース(修飾部が代名詞のとき)

例 : 私の妹 ⇔ 我妹妹

② 修飾部がヘッドの「主」であるケース(修飾部が定指示ではないとき)

例 : 瓶の蓋 ⇔ 瓶盖

③ 修飾部がヘッドの素材であるケース

例：金のネックレス ⇔ 金项链

④ 修飾部がヘッドの職位・身分であるケース

例：弁護士の鈴木さん ⇔ 律师铃木

⑤ 修飾部がヘッドの内容や様式である場合[1]

例：数学の成績 ⇔ 数学成绩

以上のように、名詞句の修飾部と被修飾部の意味関係に注目して日本語の「の」と中国語の「的」の対応関係を分類しているが、分類しきれない部分があることが問題点だと思われる。例えば、「音楽の先生」、「音楽教師」である。張 (2009)では、この二つの名詞句は、修飾部がヘッドの内容や様式であるとされている。この説明に従うと、奥野・王(2019)で分析した「音楽の先生」と同様に、「音楽教師」は日本語の「の」が必要で、中国語の「的」が不要なケースに属しているはずであるが、この名詞句（音楽教師）の場合、「の」が入らなくても成立する。このように、ただ修飾部と被修飾部の意味関係という言語学の面だけから日本語の「の」と中国語の「的」の関係を検討するのは十分ではないと言える。

2.2 CLJ による名詞修飾の「の」の習得研究

CLJによる日本語名詞句における名詞修飾の「の」の習得状況に関する先行研究を以下のようにまとめる。

松田他(2006)は 7 ヵ国の日本語学習者と日本語母語話者による作文データを分析した。その結果、CLJ により「の」の脱落が起きることが示されており、母語影響の可能性、日本語の複合語生成ルールの不明等がその脱落の要因として取り上げられている。また、奥野・金(2010)では、『KY コーパス』における英・韓・中国語母語話者の発話における名詞句の使用実態を分析した結果、CLJ による「の」の脱落が生じる要因として、母語の影響の可能性のほかに、言語処理ストラテジーの可能性等を示している。しかし、CLJ はどのような名詞句で脱落を起こしやすいか、そしてその特徴が何かは、不明である。

蘇(2018)は日本語のレベルが異なる CLJ(上級以下)を対象に、連体修飾節における「の」の習得調査を行った。その結果、修飾部の名詞が所属、所有、生産地を表す場合、CLJ にとっては習得しやすいが、内容や場所指定等を表す場合、N2 になっても正答率が低かったことが示されている。その要因は母語の影響および修飾部と被修飾部の意味関係によって「の」の使用・不使用のルールが不明なことであると指摘している。しかし、これは上級以下の CLJ を対象とした結果に留まり、上級レベル以上の

[1] このケースは日本語では「の」を、中国語では「的」を使わないケース、日本語では「の」を使うが、中国語では「的」を使わないケースについての両方に属している。張(2009)では、このケースを前者のケースとして分類したが、後者について言及されていない。奥野・王(2019)はこのケースを後者のケースとして抽出した名詞句項目を対象に分析した。

CLJ による「の」の脱落状況や原因は解明していなかった。

奥野・金(2011)では、韓国語・中国語母語話者を対象に、中国語の「的」・韓国語の「의」と日本語の「の」の「一致、不一致、任意」の用法の名詞句を分析した。その結果、CLJ は日本語の「の」と中国語の「的」が「不一致」、「任意」の名詞句では、「の」の脱落を起こしやすいと指摘されているが、具体的にどのような名詞句では「の」の脱落をしやすいかは言及されていない。

また、奥野・王(2019)は日本語の「の」と中国語「的」の用法が対応しない名詞句(例：日本語の「数学の成績」と中国語の「数学成绩」)を対象に、CLJ、日本語を母語とする中国語学習者による翻訳データ(母語から目標言語への対訳)を分析した。その結果、身近な言葉であればあるほど、CLJ による「の」の脱落が少なくなり、また、身近な名詞句においても、名詞句によって CLJ による「の」の脱落状況が異なることが見られた。日本語名詞句の出現頻度、CLJ による日本語名詞句の親密度が CLJ における「の」の脱落と関連していることを示唆している。しかし、名詞句の出現頻度、CLJ による親密度が本当に作用しているかは不明である。

以上 4 つの先行研究から、日本語名詞句では、レベルを問わず、CLJ による「の」の脱落が起きること、また、その脱落の要因は母語の影響、学習者のストラテジー、複合語生成ルールの不明などの要因以外に、奥野・王(2019)では日本語名詞句の出現頻度、CLJ による日本語名詞句の親密度が関連しているのではないかと推測されていることが明らかとなった。では、実際に、言語習得研究では、単語の出現頻度、学習者の親密度が言語習得とどのように関連するとされているのか、2.3 で概観する。

2.3　単語の出現頻度、親密度に関する先行研究

単語の親密度、出現頻度は言語心理学の分野において言語認知の過程を解明するために重要な役割を果たしている。言語心理学研究において具体的にどのように作用するとされているのか、以下 2.3.1 でそれに関する先行研究を示す。

2.3.1　言語心理学における単語の出現頻度、親密度の研究

言語心理学で、単語の出現頻度、親密度が単語の認知や文の理解に与える影響を検討した研究として横川・薮内(2006)、Takahashi(2005)がある。横川・薮内(2006)は日本人英語学習者を対象に、親密度や頻度が英単語の認知に与える影響を検討した結果、親密度と頻度両方とも効果があるが、親密度のほうがより効果があったと指摘している。また、Takahashi(2005)では、日本人英語学習者を対象に、動詞の親密度及び頻度がガーデンパス文[2]の理解に与える影響を検討した結果、文の理解には親密度が大きく

[2] 横川・薮内(2006)では、「The horse raced past the barn fell」という文を処理した際に、「The horse raced past the barn」としていったん処理した構造が、動詞「fell」の出現によって「The horse 「raced past the barn」 fell」という構造への再分析が要求される文であると記述されている。

影響していると示された。

以上のように、頻度より親密度の方が学習者の単語の認知や文の理解に主に影響していることが分かった。しかし、英語学習者を対象とするものは多いが、日本語学習者を対象とするものは少なかった。では、日本語の出現頻度、日本語学習者(特にCLJ)にとっての親密度が日本語の習得にどのように影響しているかを以下、2.3.2に示す。

2.3.2 CLJを対象とする日本語の出現頻度、親密度の研究

CLJ を対象とする日本語の出現頻度、親密度の検討には陳(2014)、松島(2009)、姚(2018)がある。

陳(2014)はCLJを対象に、日本語単語の出現頻度およびCLJによる親密度の関係を検討した結果、単語の親密度の順位と頻度の順位の間の相関は中程度であることを指摘している。また、松島(2009)は、中級レベル以上の CLJ を対象に漢字二字熟語の親密度調査を行った。その結果、中国語と関係の強い日本語漢字語彙の方が、弱いものよりも親密度が高いことが示されている。更に、姚(2018)は上級レベルの CLJ を対象に、慣用句の理解度については、親密度、透明度との関連性を明らかにした結果、親密度と理解度の間に正の相関関係があったと示している。このように、日本語の出現頻度とCLJの親密度との関係及び、CLJの日本語の単語に対する親密度状況、CLJによる慣用句の親密度がその理解に与える影響を明らかにした。しかし、単語や慣用句以外のケースにおける日本語出現頻度、親密度の効果に関する検討はなされていない。

以上のことを踏まえ、本研究の目的と分析対象を以下のように示す。日本語では「の」が必要で、中国語では「的」が不要な日本語名詞句を対象に、CLJによる「の」の脱落は日本語名詞句の出現頻度、CLJによる日本語名詞句の親密度と関係があるか、もし関係がある場合、どのような関係があるか、以下の4パターン(頻度高×親密度高、頻度高×親密度、頻度低×親密度高、頻度低×親密度低)を検討していく。

3. 本研究で分析対象とする高頻度及び低頻度日本語名詞句の抽出

この章では、本研究の調査項目となる高頻度、低頻度の日本語名詞句(日本語：N1+の+N2 ⇔ 中国語:N1+N2)の抽出方法、高頻度、低頻度の分け方などを提示する。

建石(2018)は、研究者自身の内省が働かない部分に対する例文の補強や裏付けのためにコーパスを利用することが日中対照研究においても有効であると指摘している。それゆえ、本研究では、分析対象の日本語名詞句を抽出するために、日本語コーパス(BCCWJ)と中国語コーパス(CCL)の対照[3]を行った。それによって抽出した名詞句は日

[3] 日中コーパスを対照した際に、以下2つの作業をした。まず、『BCCWJ短単位語彙表(Version 1.1)』における12ジャンルの各ジャンルにおける順位は同じく上位5万位以内の日本語名詞(9588個)から、中国語(古代中国語も含む)の語形と意味と同様な名詞(603個)を抽出した。また、これらの名詞が修飾部(N1)、被修飾部(N2)とする場合の名詞句(N1+の+他名詞)、(他名詞+の+N2)は『BCCWJ』の中納言短単位検索機能を用いて抽出し、訳された中国語の形をCCLで検索した。中国語では「N1+他名詞」、「他名詞+N2」の形のみであれば、その日本語名詞句は本研究の調査項目として抽出する。以上の手

本語母語話者に「の」の、中国語母語話者に「的」の、使用・不使用の判断[4]をしてもらった。この二つの作業によって抽出した本研究の分析対象となる高頻度(12 個)、低頻度名詞句(16 個)[5]を表 1 に示す。

表 1　本研究の分析対象となる高頻度名詞句と低頻度名詞句

高頻度名詞句(頻度数)	低頻度名詞句(頻度数)
言論の自由(106)、家庭の主婦(71)、個人の問題(56)、公共の場所(45)、政治の中心(44)、生命の危機(42)、影響の程度(42)、地域の特色(41)、普通の家庭(40)、常識の範囲(37)、普通の女性(37)、法律の目的(32)	音楽の才能(25)、歴史の舞台(23)、人生の目標(21)、政治の実権(20)、最新の状態(17)、実際の年齢(22)、人物の性格(13)、梅雨の季節(12)、英語の成績(8)、生存の危機(7)、個人の感想(7)、今日の話題(6)、話題の人物(4)、地域の風景(2)、精神の支柱(1)、社会の話題(1)

4.　調査

4.1　調査協力者・時期・方法

2019 年 10 月下旬〜11 月中旬の間、日本国内の日本語学校、大学、大学院に在籍している上級レベルの CLJ を対象に、読み上げテスト(「の」の産出レベル)⇒文法性判断テスト(「の」の知識の有無レベル)⇒CLJ による日本語名詞句の親密度調査⇒フォローアップインタビュー調査(「の」の脱落の理由を確認する)の順で実施した。

4.2　調査内容

第 3 章で抽出した日本語の名詞句(28 個)を用いて、読み上げテスト、文法性判断テスト、CLJ による日本語名詞句の親密度調査の内容を作成し、以下のように示す。

読み上げテストでは、各問題(28 問)における空欄部分の形を含む全ての内容を声に出して読み上げさせて録音した。実施例は以下に提示する。

例：無駄なお金を使って塾に通わなくても、英語＿成績を上げる方法があるだろう。

文法性判断テストでは、調査文における下線の部分の形(正用判断項目・誤用判断項目)が正しいかどうかを判断してもらった。実施例は以下に提示する。

・正用判断項目例：無駄なお金を使って塾に通わなくても、<u>英語の成績</u>を上げる方法があるだろう。

・誤用判断項目例：数年間、校長先生をしていた彼は<u>言論自由</u>がなければ生徒の幸せもないと言った。

順に従って抽出したものは 123 個であった。

[4] 注 2 で示された 123 個の名詞句は日本語母語話者(6 名)に「の」の、中国語に訳したものを中国語母語話者(6 名)に「的」の、使用・不使用判断をしてもらった。ある名詞句に対して、日本語母語話者の 4 名以上が「の」が必須、中国語母語話者の 4 名以上が「的」が不要であると答えた場合、その名詞句は本研究での分析対象となる名詞句として抽出した。抽出したものは 31 個であった。

[5] 高頻度、低頻度名詞句を分けた際に、注 2 で抽出した 31 個の名詞句の平均頻度(26)を使用した。平均頻度より高いものを高頻度名詞句ケース、低いものを低頻度名詞句ケースに分けた。また、低頻度では、「数学の成績、体育の成績、音楽の成績」という 3 つの名詞句は修飾が同じであるため、本研究の意図とずれているので、分析対象から排除した。

また、日本語名詞句の親密度調査では、親密度を以下のように定義する。「ある日本語の名詞句を CLJ(30 名)が、どの程度知っているか、又はどの程度見聞きし、慣れているかと感じるか、その程度を 5 段階尺度(1 一度も聞いたり見たりしたことがない。2 ほとんど聞いた/見たことがない。3 たまに聞く/見る。4 よく聞く/見る。5 非常によく聞く/見る。)で評定した時の平均値」である。その実施例は以下に示す。

例：「英語の成績」　1　2　3　4　5

5. 調査結果

テストの結果を処理する前に、それらを数値化した。紙幅に限りがあるので、読み上げテスト、文法性判断テストにおける各名詞句の正答人数(正しく答えた CLJ の人数)、誤答人数(間違って答えた CLJ の人数)、各 CLJ による各日本語名詞句の親密度を数値化した結果は省略し、CLJ による日本語名詞句親密度の結果に基づいて抽出した高親密度名詞句、低親密度名詞句を提示する。

読み上げテストでは、CLJ による「個人の問題、個人の感想」の回答には本研究の意図とずれるもの[6]があったため、この 2 項目を分析対象外とした。また、第 4 章における本研究で扱う親密度の定義に従い、CLJ による各日本語名詞句の親密度(26 個)を計算し、これらの名詞句の中央値(3.6)を用いて高、低親密度名詞句に分けた。中央値に当たるものや本研究の分析対象外のもの[7]を除き、中央値より高いものを高親密度名詞句(10 個)、低いものを低親密度名詞句(9 個)に分けた。また、表 1 の結果を元に、本研究の分析対象となる 4 パターンを抽出し、表 2 に示す。

表 2　日本語名詞句の頻度高低・親密度高低の 4 パターン

頻度高・親密度高(頻度数、親密度値)	頻度高・親密度低(頻度数、親密度値)
影響の程度(42、3.9)、普通の女性(37、4.1)、地域の特色(41、3.9)、常識の範囲(37、3.7)	公共の場所(45、2.3)、法律の目的(32、3.2)、言論の自由(106、2)、家庭の主婦(71、2.9)、政治の中心(44、3.4)
頻度低・親密度高(頻度数、親密度値)	**頻度低・親密度低(頻度数、親密度値)**
今日の話題(6、4.3)、英語の成績(8、4.5)、人生の目標(21、4.2)、音楽の才能(25、4.3)、最新の状態(17、3.9)、人物の性格(13、3.8)	梅雨の季節(12、2.7)、歴史の舞台(23、3.1)、精神の支柱(1、2.2)、生存の危機(7、1.9)

5.1 「の」の有無の判断、「の」の脱落と日本語名詞句の頻度、CLJ による日本語名詞句の親密度との関連性

6 「個人の問題」を「個人的問題」、「個人の感想」を「個人的感想」に産出した CLJ が数名いた。
7 第 3 章で示した内容と同様に、同じケースにおける修飾部または被修飾部が異なる名詞とするが、高親密度名詞句には「普通の女性/家庭」、低親密度名詞句には「政治の中心/実権」、「生存の危機/生命の危機」があった。前者から「普通の家庭」、後者から「政治の実権」、「生命の危機」を外すと、両者それぞれの標準偏差がそれほど変わらないため、分析対象外とした。

日本語名詞句の出現頻度、CLJ による日本語名詞句の親密度は「の」の脱落と関連があるかを検討するために、この両者及び知識レベルの文法性判断テスト、産出レベルの読み上げテストの名詞句(19 個)における正答人数を用いてピアソンの積率相関係数で相関分析を行った。

文法性判断テストでは、日本語名詞句の出現頻度と正答人数を分析したところ、両者の間に弱い負の相関があり、相関係数は有意傾向があることが分かった(r=0.40, p<.10)。また、CLJ による日本語名詞句の親密度と正答人数を分析した結果、両者の間に強い正の相関があり、相関係数は有意であることが分かった(r=0.79, p<.01)。

以上の結果から文法性判断テストでは、日本語名詞句の出現頻度と正答人数に弱い相関関係が認められることが分かった。しかし、両者の間で本当に関連性があるか、この両者に関わる分布図を確認したところ、「言論の自由」、「家庭の主婦」、「公共の場所」は頻度は高いものの、正答人数が他の名詞句より低すぎるため、外れ値になっている。この 3 つの名詞句を外して、改めてピアソンの積率相関係数で相関分析を行った。その結果を以下に示す。

表 3　文法性判断テストの名詞句における正答人数、頻度及び親密度

	平均値	標準偏差
頻度	22.8	14.0
親密度	3.6	0.7
正答人数	20.5	7.7

日本語名詞句の頻度と正答人数を分析したところ、両者の間に相関がないことが分かった(r=0.31, *n.s.*)。また、CLJ による日本語名詞句の親密度と正答人数を分析した結果、両者の間に強い正の相関があり、相関係数は有意であることが分かった(r=0.72, p< .01)。以上の結果から、文法性判断テストでは、日本語名詞句の出現頻度ではなく、CLJ による日本語名詞句の親密度が正答人数と関連性を持っていると言えるだろう。

次に、読み上げテストでは、日本語名詞句の出現頻度と正答人数を分析した結果、両者の間に弱い負の相関があり、相関係数は有意傾向があることが分かった(r=0.41, p<.10)。また、CLJ による日本語名詞句の親密度と正答人数を分析したところ、両者の間に強い正の相関があり、相関係数は有意であることが分かった(r=0.85, p<.01)。

以上の結果から、読み上げテストでは、日本語名詞句の頻度と正答人数に弱い相関関係が認められた。しかし、両者の間で本当に関連性があるか、この両者に関わる分布図を確認したところ、「言論の自由」、「家庭の主婦」、「公共の場所」は頻度は高いものの、正答人数が他の名詞句より低すぎるため、外れ値になっている。この 3 つの名詞句を外して、改めてピアソンの積率相関係数で相関分析を行った。その結果を以下に示す。

表 4　読み上げテストの名詞句における正答人数、頻度及び親密度

	平均値	標準偏差
頻度	22.8	14.0
親密度	3.6	0.7
正答人数	20.4	7.4

日本語名詞句の出現頻度と正答人数を分析したところ、両者の間に相関がないことが分かった(r=0.27, *n.s.*)。また、CLJ による日本語名詞句の親密度と正答人数を分析した結果、両者の間に強い正の相関があり、相関係数は有意であることが分かった(r=0.84, p＜.01)。以上の結果から、読み上げテストでは、日本語名詞句の出現頻度は正答人数との間には関連がないのに対して、CLJ による日本語名詞句の親密度は正答人数の間に強い相関関係があることが分かった。

ここから、CLJ による「の」の有無の判断と「の」の脱落は CLJ による日本語名詞句の親密度のみと関連があることが分かった。では、実際に、親密度はどのように CLJ による「の」の有無の判断、「の」の脱落に影響を与えるかを次節で示す。

5.2 「の」の有無の判断、「の」の脱落における CLJ による日本語名詞句の親密度の影響

5.1 で、「言論の自由」、「家庭の主婦」、「公共の場所」を相関分析に入れても入れなくても、CLJ による日本語名詞句親密度は両テストにおける正答人数と関連性があることが見られたため、両テストにおける CLJ による日本語名詞句の親密度の影響を検討する際に、この 3 つの名詞句を入れて検討した。

文法性判断テストでは、高親密度名詞句(10 個)、低親密度名詞句(9 個)では、CLJ の正答延べ人数、誤答延べ人数を集計し、その結果を表 5 に示す。

表 5　高、低親密度名詞句における正答・誤答延べ人数

親密度	正答人数	誤答人数	合計
高	237▲	63▽	300
低	97▽	173▲	270
合計	334	236	570

▲は残差分析の結果有意に多いもの、▽は有意に少ないものを示す(p＜.01)。

カイ二乗検定の結果、正答人数、誤答人数は高親密度名詞句と低親密度名詞句の間では、有意差があった($\chi^2(1)$=106.91, p＜.01)。残差分析の結果、高親密度名詞句の「正答人数」と低親密度名詞句の「誤答人数」が有意に多く、高親密度名詞句の「誤答人数」と低親密度名詞句の「正答人数」が有意に少なかった。以上のことから、CLJ は高親密度名詞句では、「の」の有無を正しく判断できるのに対し、低親密度名詞句では、

「の」の有無を正しく判断できない傾向が強いことが明らかになった。

また、読み上げテストでは、高親密度名詞句(10 個)、低親密度名詞句(9 個)において、CLJ の回答人数における正答延べ人数、誤答延べ人数を集計した。結果を表 6 に示す。

表 6　高、低親密度名詞句における正答・誤答延べ人数

親密度	正答人数	誤答人数	合計
高	241▲	59▽	300
低	92▽	178▲	270
合計	333	237	570

▲は残差分析の結果有意に多いもの、▽は有意に少ないものを示す($p<.01$)。

カイ二乗検定の結果、回答状況の人数における正答人数・誤答人数は、高親密度名詞句と低親密度名詞句の間で有意差があった($\chi^2(1)=123.29, p<.01$)。残差分析の結果、読み上げテストでは、高親密度名詞句の「正答人数」と低親密度名詞句の「誤答人数」が有意に多く、高親密度名詞句の「誤答人数」と低親密度名詞句の「正答人数」が有意に少なかった。以上のことから、CLJ は高親密度名詞句では「の」を正しく使えるのに対して、低親密度名詞句では、「の」を脱落する傾向が強いことが明らかになった。

以上のように、CLJ が、高親密度名詞句では、「の」の使用の有無に対する判断をしやすく、「の」の脱落を起こしにくい一方、低親密度名詞句では、「の」の使用の有無に対する判断をしにくく、「の」の脱落も起こしやすいことから、CLJ による日本語名詞句の親密度の高低は「の」の有無の判断や脱落に影響していることが分かった。

6.　考察

第 5 章から分かったこと及びフォローアップインタビュー調査の結果を元に、6.1、6.2 では、「の」の脱落の要因を考察していく。

6.1　CLJ による日本語名詞句の親密度と「の」の脱落の関連

「の」の有無の判断及び脱落に影響を与えているのは、CLJ による日本語名詞句の親密度しか見られなかった。その要因はフォローアップ調査から分かったことに基づき、以下のように推測できる。

調査対象者の CLJ01、CLJ15 によると、「の」の有無を判断した際に、ある名詞句を見聞きしたことがあるのか否かをも参考しているという。また、CLJ10 によると、ある名詞句をよく見聞きする場合、その形式は自分の中の知識になり確実に判断できるが、それほど見聞きしない場合、名詞と名詞の間に「の」を繋ぐという文法や今まで積み重ねた語感、中国語での形式に基づいて判断するという。ここから、ある名詞句は出現頻度が高いにも関わらず、コーパスに客観的に存在しているものであるため、必ずしも CLJ の知識になっているわけではないということが言える。一方、見聞きし

たことがあり、CLJ による親密度があるものは、CLJ における暗示的な知識になり、産出レベルでは自動化できるため、より正しく産出できるのだろう。

6.2　CLJ による日本語名詞句の親密度の高低と「の」の脱落の関連

第 5 章では、高親密度名詞句では、CLJ による「の」の脱落が起きにくいが、低親密度名詞句では、CLJ による「の」の脱落が生じやすい傾向が見られた。しかし、この両ケースにおける各名詞句の詳細をみると、高親密度名詞句で CLJ による「の」の脱落が起きやすいもの(「最新の状態」、「常識の範囲」、「地域の特色」)[8]もあれば、低親密度名詞句で CLJ による「の」の脱落が起きにくいもの(「政治の中心」、「法律の目的」、「梅雨の季節」)[9]もあることが分かった。その要因は文法性判断テスト(知識面)と比較した結果及びフォローアップ調査の結果を元に、以下のように推測できる。

まず、高親密度名詞句における CLJ による「の」の脱落が起きやすいものがあるのは知識面の問題であるかを明らかにするため、文法性判断テスト、読み上げテストにおける上記の 3 つの名詞句の正答人数を比較し、その結果を表 7 に示す。

表 7　高親密度名詞句における文法性判断テストと読み上げテストの正答人数

名詞句	文法性判断テスト正答人数	読み上げテスト正答人数
地域の特色	12	17
最新の状態	21	17
常識の範囲	25	19

表 7 から、迫田(2002)で示されたことと同様に、言語知識と言語運用の間にギャップがあり、言語運用はそのまま言語知識を反映しているとは限らないことが分かった。その要因について、「最新の状態」、「常識の範囲」は、文法性判断テストにおいて、これらの名詞句における「の」の有無を正しく判断したが、読み上げテストでは「の」が脱落した CLJ15 から以下の意見が得られた。

> 確かに見たことがあるし、日本語では名詞と名詞を繋ぐ時に、「の」を入れるので、「の」を入れる必要があると思うが、書き言葉では、「の」を入れないと違和感を覚えるのに対して、話し言葉の場合、簡潔さも大事であり、「最新情報」、「最新作」、「出題範囲」などの名詞句も耳にしたことがあるので、「の」を入れずに話してもよいだろう。

ここから、CLJ15 は「の」の脱落可否を判断した時に、「の」の文法的用法ではなく、教科書における「の」の説明(名詞と名詞を繋ぐ際に使用する)に従って解釈したので、

---

8 この 3 つの名詞句の正答人数はこのケースの平均正答人数(24 人)より少なかった。
9 この 3 つの名詞句の正答人数はこのケースの平均正答人数(10 人)より多かった。

「の」に関する明示的知識を持っていないことが分かった。CLJ15 はインプットによって形成された暗示的知識を持っているが、「の」を脱落するかどうかを判断する際に、単にその名詞句に着目するだけではなく、文体(書き言葉か、話し言葉か)、また、「N1+の+N2」における N1、N2 が他の名詞と繋がる際の形(「N1+□」、「□+N2」)を参考にして「の」を意識的に脱落させようとしていることがあることが分かった。

「地域の特色」[10]では、両テストで正しく回答した CLJ01、文法性判断テストでは正しく回答したが、読み上げテストでは正しく回答しなかった CLJ09、文法性判断テストでは正しく回答しなかったが、読み上げテストでは正しく回答した CLJ16、両テストとも正しく答えなかった CLJ19 にインタビューした。

CLJ01 によると、「地域の特色」という言葉を見たことがあることと、教科書における名詞と名詞の間に「の」を入れる文法説明を参考し判断したという。また、CLJ09 から、判断した時に、「地域の特色」に特に違和感を感じなかったが、読んだ時に、前の文では「家族形態」の前に「の」と「や」があるので、「地域特色」と並列関係があるかと思い「の」を脱落して読んだという意見が得られた。また、CLJ16 の場合、文法性判断テストでは、前の文では「家族形態」には「の」や「や」があるので、「地域の特色」より「地域特色」のほうが並列関係を表せると思って判断したが、読んだ際に、「地域特色」というと、中国語のように聞こえるので、「の」を入れて読んだという。更に、CLJ19 によると、「の」と「家族形態」があるので、一つの文に「の」を複数入れると不自然な感じであり、「地域特色」にしたという。

文脈のことを考えない状態で、「地域の特色、地域特色」はどちらが正しいと思うかと CLJ09、CLJ16、CLJ19 に問いかけたところ、「地域の特色」だと回答した。日本語では、二つの名詞を繋ぐ際に、「の」を入れることをよく見ており、教科書ではこの用法の説明もあった(CLJ19)。「地域特色」は中国語では四字として使われているが、日本語では「地域の人々」や、「大学の特色」などがあるので、日本語では、「地域特色」だと、中国語のように聞こえるので不自然だと感じる(CLJ09、CLJ16)。

次に、なぜ、低親密度名詞句ケースでは CLJ による「の」の脱落が起きやすいか、また、なぜその中で「の」の脱落が起きにくいものもあったのか、これは知識面の問題であるかを明らかにするために、高親密度名詞句ケースと同様に、文法性判断テスト、読み上げテストにおける名詞句の正答人数を比較し、その結果を表 8 に示す。

表 8　低親密度名詞句における文法性判断テストと読み上げテストの正答人数

名詞句	文法性判断テスト正答人数	読み上げテスト正答人数
政治の中心	22	22
法律の目的	22	21
梅雨の季節	23	21

10 文法性判断テスト：そこで暮らす人々の家族形態や地域の特色があることを忘れてはならない。
読み上げテスト：そこで暮らす人々の家族形態や地域＿特色があることを忘れてはならない。

歴史の舞台	13	11
家庭の主婦	0	0
公共の場所	4	2
精神の支柱	4	6
言論の自由	2	4
生存の危機	7	5

表 8 の文法性判断テスト、読み上げテストにおける正答人数を比較し、両テストにおける「政治の中心」、「法律の目的」、「梅雨の季節」、「歴史の舞台」では正しく回答したが、他の名詞句では正しく答えなかった CLJ10、両テストにおいて、「梅雨の季節」しか正しく回答できなかった CLJ29 にインタビューした。その結果を以下に示す。

「法律の目的」、「梅雨の季節」は日本語では見たことがなく、「政治の中心」は見たことがある。「梅雨」や「目的」のような言葉は単独に使われる場合が多いが、他の名詞と共に使う時、「の」を付けることが多い気がする。例えば、「梅雨の時期」、「梅雨の頃」、「授業の目的」などである。また、「歴史の舞台」は日本語では見たことがないが、「政治家の舞台」、「サッカーの舞台」のような「. . . の舞台」の形式を見たことがある。それで、「の」を入れないと、違和感を覚えてしまう。また、日本語では、二つの漢字の間に「の」を入れる場合が多いのではないか。「家庭用品」、「集合場所」は日本語では見たことがあるが、「家庭の主婦」、「公共の場所」、「精神の支柱」、「言論の自由」、「生存の危機」は見たことがない。これらの言葉は中国語では、「家庭主妇」、「公共场所」、「精神支柱」、「言论自由」、「生存危机」のように、複合語のイメージが強いので、日本語でも同じではないかと思って判断した(CLJ10 からの意見)。

「都市の中心」、「仕事の目的」は見たことがあるから、「政治の中心」、「法律の目的」では、「の」を普通に入れてもいいと思う。しかし、「政治」、「中心」、「法律」、「目的」は普段あまり使わない言葉で、少し硬いと感じている。硬い語と硬い語と結ぶ際に、「の」を入れるとその名詞句自体は硬いニュアンスが伝わらないと思って、「の」を入れなかった。その他の名詞句は中国語では四字の形で使われているので、日本語でもそうではないかと思った(CLJ29 からの意見)。

以上のことを踏まえ、高親密度名詞句においても、低親密度名詞句においても、CLJ による「の」の脱落要因には以下(1)、(2)が見られた。

(1) 教科書における「の」の説明(名詞と名詞を繋ぐ際に使用する) やインプットによって自分の中で形成された暗示的知識を使用する傾向があるが、日本語名詞句における「の」の文法的用法に関する明示的知識を持っていない。

(2) 「N1+の+N2」におけるN1、N2とほかの名詞と繋いで名詞句になる際の形(「N1+の+□」、「□+の+N2」「N1+□」、「□+N2」)に引っ張られている。

次に、(2)に関して、高親密度名詞句、低親密度名詞句から分かったことを説明する。高親密度名詞句では以下の2点が分かった。① 文体や文脈を考慮する場合、「N1+の+N2」における「N1+□」、「□+N2」の形式も見たことがあるということを根拠に、「の」を意識的に脱落させて話す傾向がある。例えば、「最新の状態」、「常識の範囲」がある。「最新情報」、「出題範囲」を見聞きすることを根拠に、「最新の状態」を「最新状態」、「常識の範囲」を「常識範囲」として産出してしまう。② 文体や文脈を考慮せず、単に「N1+の+N2」を考える際には、「N1+の+N2」における「N1+の+□」、「□+の+N2」の形式を見たことがあることを根拠に、意識的に「の」を脱落させず日本語らしくすることもある。これは、CLJにとっては「N1+N2」の形式が中国語らしいと感じられるためである。例えば、「地域の人々」や「大学の特色」を見聞きしたことがあるので、この名詞句(地域の特色)では、「の」を付けると、日本語として自然だと認識し、「の」の脱落をしないようにする。

低親密度名詞句では、以下の2点が明らかとなった。まず、低親密度名詞句の「N1+の+N2」におけるN1、N2とほかの名詞と繋いで名詞句になる際の形(「N1+の+□」、「□+の+N2」)を見たことがある場合、「の」が脱落しないことが多い。例えば、「梅雨の時期」、「梅雨の頃」、「授業の目的」などを見聞きしたことがあるので、「梅雨の季節」、「法律の目的」では「の」が脱落しないことがある。また、「N1+の+N2」におけるN1、N2とほかの名詞と繋いで名詞句になる際の形(「N1+の+□」、「□+の+N2」)を見たことがない場合、中国語の影響を受けて「の」を脱落させてしまうことが多い。例えば、「家庭の□」、「□の主婦」、「公共の□」、「□の場所」の言語形式を見聞きせず、中国語では四字で使われるイメージが強いので、「家庭の主婦」を「家庭主婦」に、「公共の場所」を「公共場所」に産出してしまう。

7. まとめと今後の課題

本研究では、日本語名詞句におけるCLJの「の」の脱落は日本語名詞句の出現頻度ではなく、親密度の高低と関連があることが明らかになった。また、親密度と共に、教科書における「の」の用法説明、文体、文脈、「N1+の+N2」におけるN1、N2が他の名詞と繋いで名詞句になる際の形(N1+の+□、□+の+N2、N1+□、□+N2)などからも影響を受けていることが分かった。更に、高親密度名詞句では、「の」に関する暗示的知識ができており、「の」がないほうが、中国語らしく感じられるので、CLJは母語の影響を意識的に避けて「の」を脱落しないストラテジーを使用することが窺えた。一方、低親密度名詞句では、日本語の「の」の用法に関する知識不足などのため、CLJは母語の影響を無意識に受けることで「の」を脱落させてしまう傾向も見られた。

今回の調査では、CLJによる「N1+の+□」、「□+の+N2」の親密度調査を行わなかった。今後、この部分を含め、CLJは「N1+の+N2」における「の」の脱落では、「N1+の+□」、「□+の+N2」のどちらの影響を更に受けやすいかを検討していく。

参考文献

奥野由紀子(2005)「第二言語習得過程における言語転移の研究－日本語学習者による「の」の過剰使用を対象に－」風間書房.

奥野由紀子・金玄珠(2010)「日本語学習者の『の』の脱落に関する―考察―横断的発話資料に基づいて―」『留学生センター教育研究論集』17: 45-63.

奥野由紀子・金玄珠(2011).「漢字圏学習者の「の」の脱落における言語転移の様相―「の」「의」「的」の対応関係に着目して―」『国立国語研究所論集』2: 77-89.

奥野由紀子・王婧瑶(2019).「名詞句の日中言語双方向的検討による翻訳課題からの考察―親密度・定式度への考慮の必要性―」首都大学東京人文科学研究科「人文学報」515-7: 67-78.

建石始(2018)「対照言語学の分析」『コーパスから学ぶ日本語学 日本語教育への応用』pp. 125-125, 朝倉書店.

迫田久美子(2002)『日本語教育に生かす第二言語習得研究』アルク.

松田真希子・森篤嗣・金村久美・後藤寛樹(2006).「日本語学習者の名詞句誤用と言語転移―アジア 7 ヵ国による日本語作文データに基づく分析―」『留学生教育』11: 45-53.

松島弘枝(2009)「中国人日本語学習者を対象にした漢字二字熟語の親密度調査」*Sapientia* (43): 165-177, St.Thomas University.

横川博一・薮内智(2006)「親密度や頻度が英単語認知に及ぼす影響―語彙性判断課題によるパイロットスタ ディ―」横川博一編『日本人英語学習者の英単語親密度 文字編』pp.103-110, くろしお出版.

蘇雅玲(2018)「日本語の連体修飾節における『の』の習得研究」『日本言語文藝研究』18: 19-32.

張麟声(2009)「名詞にかかる連語的修飾構造の日中対照研究―「の」と“的”の使用の有無を中心に―」『言語文化研究』(言語情報編) 4: 22-36.

陳相州(2014)「台湾人日本語学習者を対象とした日本語単語親密度データベースの構築」『比較文化研究』111: 167-179.

姚新宇(2018)「中国人日本語学習者の慣用句の理解―親密度と透明度の影響を中心に―」『日本語教育学会秋季大会予稿集』, pp. 226-231.

Takahashi, A. (2005). Familiarity and frequency effects of English verbs in EFL readin g: Evidence from word recognition and processing of garden-pathsentences. Master' s thesis submitted to Graduate School, Kyoto Uniersity of Foreign Studies.

（おう　せいよう・首都大学東京大学院博士前期課程）