The University of Osaka
Institutional Knowledge Archive

Title	日本人中国語学習者の作文における接続詞の使用実態：中国語母語話者による作文との比較
Author(s)	徐，勤
Citation	大阪大学言語文化学. 2022, 31, p. 123-139
Version Type	VoR
URL	https://doi.org/10.18910/87498
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

日本人中国語学習者の作文における接続詞の使用実態[1]

―中国語母語話者による作文との比較―*

徐　勤**

キーワード：日本人中国語学習者、中国語作文、接続詞

本研究基于自建语料库，通过与汉语母语者的记叙文语料进行对比，从连词的使用频度、词汇多样性、高频连词、连词的出现位置等方面考察了日本中级、高级汉语学习者记叙文写作中连词的使用情况。结果表明：中级、高级汉语学习者的记叙文作文中连词的出现频度均显著高于汉语母语者，而中级、高级汉语学习者之间不存在显著性差异；中级、高级汉语学习者记叙文中连词的词汇多样性均显著低于汉语母语者，且中级汉语学习者作文的连词词汇多样性显著低于高级汉语学习者。表明学习者的汉语写作存在连词使用频率过高、但连词种类偏少的问题，且随着汉语水平的提高，高级汉语学习者作文中连词的种类虽有所增加，但其连词词汇多样性仍未达到汉语母语者的水平。据此，本文使用 R 语言绘制词云图和对应分析图进行可视化分析，发现中级、高级汉语学习者记叙文作文中连词的使用情况较为相似，但与汉语母语者存在较大差异。

进一步考察各个连词的出现频率发现，中级、高级汉语学习者的作文语料库中，前 10 个高频连词的总出现频度约占连词总频度的 80%，而汉语母语者作文语料库中前 10 个高频连词的总出现频度的占比仅为 60%，表明学习者作文中高频连词的比例远高于汉语母语者的作文。同时，学习者的两个作文语料库中前 9 个高频连词的种类一致，而汉语母语者使用频率最高的前 9 个连词中的“而”和“与”并未出现在学习者的高频连词列表里。此外，中级、高级汉语学习者作文语料库中前 3 个高频连词的种类和出现顺序一致，均为“和”、“所以”、“但是”，这 3 个高频连词的总出现频率分别占各语料库中连词总频度的约 53% 和 49%，而汉语母语者作文语料库中前 3 个高频连词“和”、“而”、“但”的总出现频率仅约占该语料库中连词总频度的 38%。单因素方差分析的结果表明，学习者作文中过度使用了难度较低的“和”、“所以”、“但是”，但对难度较高的“而”、“与”则使用不足。另外，本文还基于 CasualConc 考察了连词在句子中的出现位置，结果表明，中级、高级汉语学习者在句首使用连词的比例高于汉语母语者。本研究的发现或可为二语汉语写作相关

[1] 本研究は中国語の接続詞のみを対象とした研究であり、接続助詞については議論しない。また、例文についての和訳は主に逐語の訳である。中国語の接続詞を日本語に訳する際には、便宜上、日本語の接続（助）詞を使用している。

* 日本人汉语学习者中文写作连词使用情况研究—以汉语母语者的作文作为对比—（徐 勤（XU Qin））
** 大阪大学大学院言語文化研究科博士後期課程

的连词教学和研究提供借鉴和参考。

1 はじめに

接続詞は「前後の文脈の関係を表すもの」（石黒, 2008: 44）として、文、節、句、語などを繋ぐ機能を持っており、文章を作る上で非常に重要な役割を果たしている。接続詞を適切に使用することで、文章の理解を促進することができる（Altenberg & Tapper, 1998: 80）。しかし、日本人中国語学習者が書いた中国語作文を読んでみると、接続詞が過剰に用いられている印象を受けることが以下のように少なくない。

（1）我妈妈很会做菜，**所以**我想像我妈妈。我妈妈的得意菜是干烧蝦仁，**所以**我想一起做。最近我也喜欢做点心。**但是**，因为失败了，**所以**我想再做。我的第四个爱好是去时尚的咖啡馆。我在那里喝红茶，吃蛋糕。**但是**，因为蛋糕的卡路里很高，**所以**我不得不注意不能吃多。（JPM_16）
（訳：私の母は料理が得意です。**それゆえ**、私は私の母のような人間になりたいです。私の母の得意な料理はエビチリです。**それゆえ**、私は一緒に作りたいです。最近私はお菓子を作るのも好きです。**しかし**、失敗した**ので**、**それゆえ**、私はまた作りたいです。私の４番目の趣味はおしゃれなカフェに行くことです。私はそこで紅茶を飲んだり、ケーキを食べたりします。**しかし**、ケーキはカロリーが高く**て**、**それゆえ**、私は食べ過ぎないように気をつけなければなりません。）

（2）第二天，我们去故宫博物院，这座博物院非常有名，**所以**我从以前想访问**和**看见白菜**和**烧鱼块。我会看见白菜，**但是**当时烧鱼块在中国，**所以**我不会看见它。白菜比我想象更小，**但是**它很漂亮**和**细腻，我对它感动了。博物院非常大，**所以**我们花了很长时间。**然后**，我们一边喝珍珠奶茶，一边去各地。在鼎泰丰，我们吃晚饭。（JPH_75）
（訳：２日目、私たちは故宮博物院に行きました。この故宮博物院はとても有名なところです。**それゆえ**、私は昔から翠玉白菜**と**肉形石を見てみたいと思っていました。翠玉白菜を見ました**が**、その時肉形石は中国にありました。**それゆえ**、見ることができませんでした。翠玉白菜は思ったよりも小さかったです。**しかし**、とても繊細で美しくて、私は感動しました。故宮博物院はとても大きいです。**それゆえ**、とても時間がかかりました。**それで**、私たちはタピオカミルクティーを飲み**ながら**回りました。私たちは鼎泰豐で夕食を食べました。）

例（1）は「中級中国語作文」の授業を受講している大学２年生の日本人が書いた文

であり、例（2）は「上級中国語作文」の授業を受講している大学3年生の日本人が書いた文である。上の例文を見ると、“但是”（しかし、～が）、“因为”（～ので）、“所以”（それゆえ）などの接続詞が多く使われている。第二言語（L2）ライティングの研究分野において、L2学習者が書いたエッセイにおける接続詞の過少使用（underuse）、過剰使用（overuse）などの研究が注目に値する（Bolton *et al.*, 2002; Chen, 2006; Crewe, 1990等）。本研究の目的は、中国語を専攻としている日本人大学生の中国語作文における接続詞の使用実態を、中国語母語話者の作文と対照しつつ明らかにし、そこから日本語母語話者に対する中国語教育の問題点を探り、教材の改善に資することである。

2 先行研究

コーパス言語学の分野では、出現頻度に着目した接続詞使用実態に関する研究の進展が著しい。それらを大別するならば、まず、異なるレジスターや文体における接続詞の出現頻度に関する研究が顕著であると言える（崔、張, 2002; 石黒ら, 2009; 髙野、上村, 2017等）。例えば、髙野・上村（2017）はBCCWJコーパスを使って、レジスターによる接続詞の出現頻度の偏りを考察するために、χ^2検定を行うことで、接続詞の出現とレジスターに関連があることを確認しており、書籍・教科書・国会会議録では接続詞が多く、新聞・ブログ・知恵袋・白書・広報誌などのようなジャンルでは接続詞が少ないことを報告している。崔・張（2002）は中国語の文語体と口語体における接続詞の使用状況を調査した結果、異なる文体で用いられる接続詞の使用頻度には大きな差があることを明らかにした。

一方、外国語教育に関する研究では、英語学習者が産出した書き言葉における接続詞の使用実態についての研究が多く行われてきた（Appel, 2020; Chen, 2006; Heino, 2010; Ishikawa, 2011; Nakayama, 2021; Narita *et al*., 2004; 今尾, 2018等）。EFLやESL学習者の接続詞使用に関する研究によると、学習者は書き言葉の中で接続詞を過剰使用したり、過少使用したりする場合がある（Altenberg & Tapper, 1998）。例えば、Appel（2020）は、中国、インドネシア、韓国の英語学習者が書いた英語意見文におけるlinking adverbialsの使用状況を比較した。Appel（2020）によると、韓国語母語話者の英作文では結果（例、*therefore*）を表す接続表現の使用が多く、中国語母語話者の英作文では時間を表す接続表現（例、*nowadays*）、インドネシア語母語話者の英作文では対照関係（例、*however*）を表す接続詞が多いなど母語によって異なる干渉が観察されるということを示した。Narita *et al*.（2004）は、日本人EFL上級学習者が書いた英語意見文におけるlogical connectorsという接続表現（25個）の使用頻度と文ごとの出現位置について、英語母語話者の英作文との比較を行った。日本人EFL上級学習者の英作文では、英語母語話

者の作文に比べ、文頭において logical connectors が著しく多く使用されており、*for example, of course, first* などの接続詞が過剰使用されている一方、*then, yet, instead* などの接続詞は過少にしか使用されていないと述べられている（Narita *et al*., 2004）。今尾（2018）は、日本人が書いた日本語作文と英語母語話者の英作文とを比較対象とし、日本人英語学習者による作文における接続表現の使用頻度や出現位置を検討した結果、英語母語話者の英作文に比べ、日本人大学生が書いた英作文と日本語作文では、文頭で接続表現が用いられる割合が多いと述べている。

また、日本語と中国語ライティングでの接続詞の使用頻度に関する研究もある。蘇（2002）は中国人が書いた中国語の論文（A）・日本語の論文（B）及び日本人が書いた日本語の論文（C）における接続詞について考察した。その結果、接続詞の使用率は A < B < C の順になっており、中国人より日本人の方が、文章を書く際に接続詞の使用率が高いことを報告している。しかし、中国語教育の分野における日本人中国語学習者の接続詞の使用については、主に作文から集めた誤用例に基づいて接続詞の誤用を分析した研究（盧 , 2008）は見られるものの、接続詞の使用頻度に関する量的研究は少ない。陳（2013）は日本人留学生の HSK 作文[2] を分析対象とし、中国人と比べ、日本人留学生が接続詞をより頻繁に使用していることに言及しているが、主に誤用分析を中心としており、接続詞の使用実態に関する量的考察は十分になされていない。そこで、本研究では、筆者が編纂した中国語作文コーパスを典拠として、日本人中国語学習者（以下は「学習者」とする）が書いた中国語作文における接続詞の使用実態について考察を試みる。また、中国語母語話者（以下は「母語話者」とする）によって書かれた中国語作文を比較対象として分析を行い、その違いを明らかにする。

3 研究手法

3.1 データ

本研究では、学習者が書いた叙述文を研究対象として取り上げる。具体的には、大阪大学外国語学部の中国語専攻に在籍しており、「中級中国語作文」の授業を受講している大学 2 年生の日本人と、「上級中国語作文」の授業を受講している大学 3 年生の日本人が書いた中国語作文を収集して電子化し、それぞれを学習者の中級作文コーパス（以下は「JPM」とする）と上級作文コーパス（以下は「JPH」とする）にしたものを分析データとする。JPM と JPH には、それぞれ 95 篇の叙述文が含まれている。さらに、学習者の接続詞の使用と比較対照するために、本研究では母語話者が書いた叙述文（95 篇）

[2] HSK 作文は、1992 年から 2005 年までの間に受験者が HSK 上級試験のために書いた作文である。

も収集し、母語話者作文コーパス（以下は「CHN」とする）を比較対象とする。表1はコーパスのサイズ及び各コーパスにおける接続詞の種類（Types）と総頻度（Tokens）の情報である。

表1　各コーパスの基本情報

Corpus	Sub-corpus	総篇数	総語数	接続詞	
				Types	Tokens
学習者作文コーパス	JPM	95	27,559	52	990
	JPH	95	49,839	62	1,810
母語話者作文コーパス	CHN	95	65,539	81	1,149

3.2 リサーチクエスチョン

本研究では、以下のリサーチクエスチョンを立てて考察を行う。

（1）接続詞の出現頻度と語彙多様性について、学習者と母語話者の間に違いはあるか。違いがあるとすれば、どのような違いがあるか。
（2）高頻度で用いられる接続詞は、学習者の作文においてどのような傾向が見られるか。
（3）接続詞の出現位置は、学習者と母語話者の間で違いはあるか。違いがあるとすれば、どのような違いがあるか。

3.3 データの処理

本研究では、データの処理は次のような手順で行う。

（1）各コーパス内の作文テキストに対し、NLPIR[3] で単語分割とタグ付けを行う。
（2）テキストごとの接続詞の出現頻度と種類について、Python で統計処理を行う。
（3）コーパス間に接続詞の出現頻度の差があるかどうかについて、One-way ANOVA（一元配置分散分析）を行う（有意水準 $\alpha \leq 0.05$）。本研究では、主に One-way ANOVA 分析の多重比較（Post-Hoc Test）[4] の結果を報告する。
（4）R と CasualConc を用い、データを可視化する。

---

[3] NLPIR は中国語のテキストを分かち書きすることができるツールである。URL: http://kgb.lingjoin.com/nlpir/〔最終アクセス日：2021 年 11 月 15 日〕
[4]「UNDERSTANDING THE ONE-WAY ANOVA」を参考にした。URL: http://oak.ucc.nau.edu/rh232/courses/EPS525/Handouts/Understanding%20the%20One-way%20ANOVA.pdf〔最終アクセス日：2021 年 9 月 10 日〕

4 結果と考察

4.1 接続詞の使用頻度と語彙多様性

表 1 の接続詞の総頻度（Tokens）を見ると、JPM より JPH と CHN の方が接続詞を多く使用しているように見えるが、各コーパスのサイズが異なるので、コーパスから得られた粗頻度をそのままでは比較できない。そのため、本研究は各テキストにおける接続詞の頻度を 1,000 語当たりの相対頻度（頻度÷各テキストの総語数× 1000）で比較する。また、Guiraud が提唱した R 値（R 値 = Type ／√ Token ）を接続詞の語彙多様性の指標として採用する。表 2 は One-way ANOVA 分析を行った多重比較の結果であり、図 1 は接続詞の相対頻度と語彙多様性の箱ひげ図（Boxplot）[5] である。

表 2 によると、JPM と JPH において接続詞の相対頻度（約 35 回、36 回）は CHN（約 17 回）の約 2 倍ほどであるが、接続詞の語彙多様性（R 値）は CHN より低いことが分かった。多重比較の結果によれば、接続詞の出現頻度は、JPM と JPH の間に有意差は認められないものの（$p = .799 > 0.05$）、学習者（JPM、JPH）と母語話者の間に有意差があり（$p = .000 < 0.05$）、学習者の方が接続詞を過剰に使用することが明らかになった。接続詞の語彙多様性は JPM < JPH < CHN の順で、学習者（JPM、JPH）と母語話者の間（$p = .000 < 0.05$; $p = .034 < 0.05$）及び学習者（JPM と JPH）の間（$p = .000 < 0.05$）において両方とも有意差があった。また、図 1 から明らかなように、学習者の方が母語話者より全体的に接続詞の使用頻度が高く、上級になると接続詞の語彙多様性が増える傾向は見られるものの、学習者の語彙多様性は母語話者より低い。これにより、学習者の作文では限られた種類の接続詞が繰り返し使用されていることが確認された。

表 2　接続詞の相対頻度と語彙多様性についての多重比較 [6]

項　目	Mean（平均値）			多重比較（Sig）		
	JPM	JPH	CHN	JPM vs. CHN	JPH vs. CHN	JPM vs. JPH
相対頻度	35.31	36.42	17.36	.000***	.000***	.799
R 値	1.76	2.07	2.22	.000***	.034*	.000***

注：$^{*}p < 0.05$、$^{**}p < 0.01$、$^{***}p \leq 0.001$（以下同じ）。

[5] 箱ひげ図では、データのばらつきを分かりやすく表現することができる。点（dots）は個別の作文の統計量を表し、横軸にコーパス名、縦軸に比較項目の統計量が配置されている。箱ひげ図により、比較項目におけるコーパスごとの違いが見られる。

[6] 相対頻度の Mean は作文テキストごとの接続詞の出現頻度を 1,000 語当たりの頻度に調整し、その合計を篇数で割った平均値である。語彙多様性の Mean は作文テキストごとの R 値を算出し、その合計を篇数で割った平均値である。

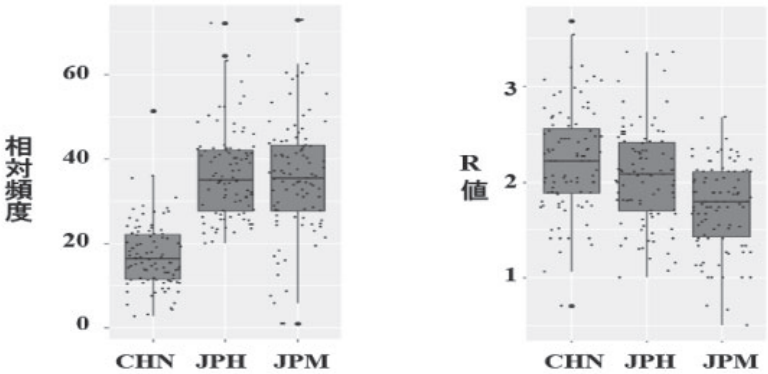

図１　相対頻度と語彙多様性の箱ひげ図

学習者は母語話者より語彙多様性が低いということは予想の範囲内ではあるが、学習者の方が接続詞の使用頻度が高いため、その原因を解明する必要がある。接続詞の過剰使用や過少使用は、必ずしも誤用を意味するものではない（Heino, 2010; Narita *et al.*, 2004）という指摘がある。しかし、接続詞を適切に使わないと、文章のまとまりや読みやすさが損なわれてしまう（Larsen-Walker, 2017）場合は少なくない。「中国語の複文の構造では、接続詞類がなくても文を繋ぐことが可能」（王, 2009）であるため、母語話者の作文では、通常接続詞を使用しなくとも文脈の前後関係を理解できる。それに対して、学習者の作文では、接続詞の誤用（例文５）のほか、接続表現が必要ではない文脈（例文３・４）においても接続詞が出現しているケースが散見される。

（3）现在已经开始了上网课，**所以**我也努力学习汉语。（JPH_02）
（訳：現在オンライン授業が始まりました。**だから**、私も頑張って中国語を勉強します。）

（4）我跟爸爸**和**妈妈一起讨论留学的问题。（JPM_27）
（訳：留学のことについて私は父**と**母と話し合いました。）[7]

（5）今天我吃了酸辣汤面、炒饭、青椒肉丝、**和**小笼包，都很好吃。（JPM_24）
（訳：今日私は酸辣湯麺（サンラータンメン）、炒飯、チンジャオロース、**と**小籠包を食べました。全部美味しかったです。）[8]

---

[7] 中国語の“爸爸妈妈”は「お父さんとお母さん」の意味であるが、例（4）の文脈において、“爸爸妈妈”の真ん中に“和”（と）を入れる必要がない。
[8] “和”の前に句読点を使用しないというのが文法のルールである。三つ以上の単語を並列に出現する場合、“A、B、C”或いは“A、B 和 C”の表現が限定であり、“A、B、和 C”は正しくない表現である。

上記の場合、接続詞を全て削除しても、文脈の意味を十分に伝えることができる。しかし、接続詞の正しい用法を習得できていない（例えば、例文5）学習者は、接続詞の省略を習得する段階には至っていないことが推察される。また、もう一つの原因として、母語である日本語の干渉が強く疑われる。中国語作文を書く時、学習者は日本語において頻繁に用いる表現（例えば、だから）を中国語（例えば、**所以**）に直訳し、作文内で用いることで、学習者が母語話者より全体的に接続詞を多く使用している事例が数多く発見されるからである。

4.2 接続詞使用実態の全体像

各サブコーパスの接続詞の全体像を把握するために、Word Cloud で可視化した図2を提示する。Word Cloud では、テキストに高い頻度で現れている単語が大きいフォントで表示されている（小林, 2017: 119）。

図2　接続詞の Word Cloud

図2を見ると、JPM と JPH には、“**和**”（和訳は表4を参照）、“**所以**”、“**但是**”などの単語が高頻度で現れている一方、CHN には“**和**”、“**而**”、“**但**”という単語が高頻度で出現している。学習者（JPM、JPH）の Word Cloud は比較的近似しているが、学習者と母語話者（CHN）の間には、全体的に大きな違いが見られる。

次に、各サブコーパスにおけるテキストと接続詞の関係を直感的に把握するため、対応分析（Correspondence analysis）を行い、その結果として得られた3次元散布図（図3）を示す。図3において、左の図は各テキスト（計285個）の分布であり、星マーク（★）は CHN、四角マーク（■）は JPH、三角マーク（▲）は JPM の作文テキストである。「頻度パターンの近いテキスト同士が近く布置され、頻度パターンの異なるテキスト同士は遠く布置され」（小林 , 2017:171）という指摘のように、3つのサブコーパスのうち、JPM と JPH はマークが近く布置され、重なっている箇所も多いため、接続詞の頻度パターンがよく似ていることが分かった。一方、右の図はテキストの分布と対応している語彙別接続詞の分布である。例えば、右の図における一番右下に位置する接続詞“**接着**”

（引き続いて）は、左の図における一番右下に位置するCHN_71に対応しており、CHN_71に特徴的であることを意味している。他のテキストに比べ、CHN_71において“**接着**”の相対頻度（5.01回）が著しく高いからである。

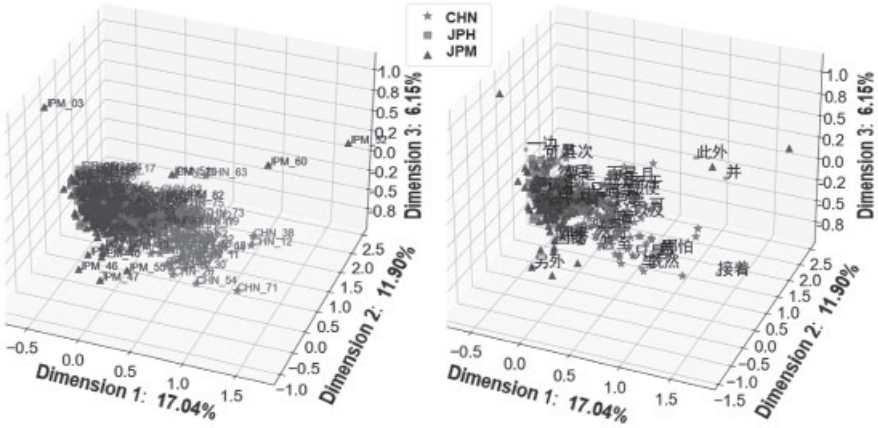

図3　テキストと接続詞の対応分析（上位50語）

また、各軸のラベルに示されている百分率は、次元（Dimension）の寄与率（contribution rate）を表す指標である。3次元の寄与率はそれぞれ17.04%、11.90%、6.15%である。CHNのテキストは主にDimension1（+）、Dimension2（-）、Dimension3（-）側に分布しているのに対し、JPMとJPHのテキストは主にDimension1（-）、Dimension2（-）、Dimension3（-）側に分布しているため、第1次元は学習者と母語話者の関係性を17.04%説明できる。第2次元と第3次元は、主にJPMにおける外れ値の影響で、例えば、JPM_46の高頻度語“**另外**”（13.27回）、JPM_60の高頻度語“**此外**”（5.85回）、JPM_52の高頻度語“**并**”（25.25回）など[9]、データが散らばっていると考えられる。一部の外れ値を除き、JPMとJPHにおける接続詞の使用実態が全体的に近似しており、学習者同士（JPM・JPH）と母語話者（CHN）の間にほぼ一貫した相異が存在していることが分かった。

4.3 高頻度接続詞

文章における語彙は、一般的に、少数の高頻度語と多数の低頻度語から構成されている（石川, 2012:152）。学習者と母語話者の作文コーパスに出現している高頻度接続詞が

[9] 中国語“**另外**”、“**此外**”、“**并**”の意味は、それぞれ「そのほか」、「これ以外に」、「かつ」である。

全体に占める比率は、中国語レベルの低い学習者作文コーパスの方が高いという仮説を立てることができる。この仮説を検証するため、各コーパスの接続詞において、頻度別の構成比を調べ、頻度が極めて高いとみなされる上位1語(T1、TはTopの略)、5語(T5)、10語（T10)、また、中低位語とみなされる15語（T15)、20語（T20)、25語（T25)、30語（T30)、35語（T35）の累計構成比を求めた。接続詞における頻度別構成分析の結果は図4の通りである。

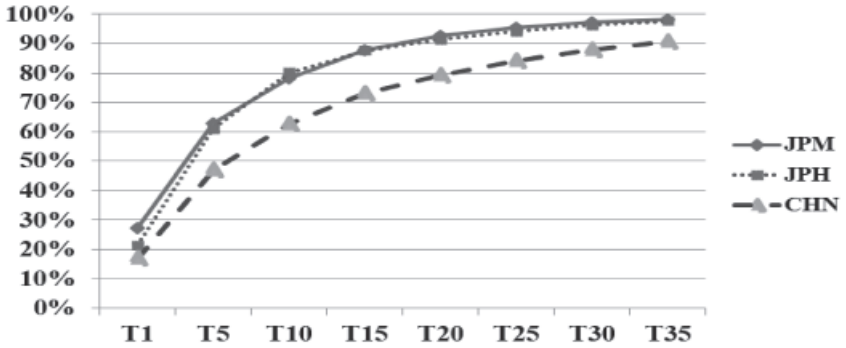

図4　累計構成比の変化

図4を観察すると、JPMとJPHにおいて、接続詞の累計構成比の変化が近似している。また、T1語からの構成比率は、JPMとJPHが一貫してCHNを上回ることが確認され、T10語以上になると差は次第に縮小する。JPMとJPHには、高頻度とみなされる上位10語の接続詞が総頻度の約80%を占めているのに対し、CHNの上位10語の接続詞は全体の約60%しか占めていない。したがって、学習者の作文で使われる接続詞の高頻度語比率が高いことが確認された。

各サブコーパスの語彙別接続詞（主に上位10語）の頻度と割合を示した結果を表3に示す。表3における「頻度」列について、括弧外は粗頻度、括弧内は相対頻度の平均値である。JPMとJPHにおける上位9位までの接続詞の順位は少し異なっているが、接続詞の種類は同じであることが明らかである。それに対して、JPM・JPHの上位9語のうち7種の語彙がCHNの接続詞の上位9語に含まれているが、ほかの2語“**而**”、“**与**”はJPM・JPHの上位9語に含まれていない。興味深いのは、使用頻度の最も高い3つの接続詞である。JPMとJPHにおける上位3語の接続詞の種類と順位が同じであり、“**和**”、“**所以**”、“**但是**”の総使用率が接続詞総頻度の約53%と49%を占めているのに対し、CHNでは“**和**”、“**而**”、“**但**”の順で総頻度の約38%を占めているにすぎない。また、括弧内の相対頻度を見ると、JPMとCHNでは上位3語はそれら以外の接続詞の

使用頻度と比べておよそ2倍以上であり、JPHでは上位3語の接続詞はそれ以外の接続詞と比べて約1.5倍以上の頻度で使用されていることが分かった。

表3　語彙別の接続詞の頻度と割合（降順）

JPM				JPH				CHN			
	語彙	頻度	割合		語彙	頻度	割合		語彙	頻度	割合
1	和	268(9.65)	27.1%	1	和	387(7.72)	21.4%	1	和	193(2.98)	16.8%
2	所以	157(5.22)	15.9%	2	所以	345(6.95)	19.1%	2	而	146(2.09)	12.7%
3	但是	100(3.61)	10.1%	3	但是	157(3.17)	8.7%	3	但	99(1.57)	8.6%
4	因为	52(1.81)	5.3%	4	可是	117(2.36)	6.5%	4	与	54(0.82)	4.7%
5	而且	45(1.60)	4.5%	5	而且	97(1.97)	5.4%	5	因为	47(0.76)	4.1%
6	然后	41(1.28)	4.1%	6	不过	82(1.67)	4.5%	6	但是	43(0.64)	3.7%
7	但	30(1.20)	3.0%	7	因为	80(1.59)	4.4%	7	然后	42(0.63)	3.7%
8	可是	30(0.86)	3.0%	8	然后	75(1.48)	4.1%	8	可是	40(0.64)	3.5%
9	不过	26(0.77)	2.6%	9	但	62(1.25)	3.4%	9	所以	29(0.39)	2.5%
10	如果	25(0.97)	2.5%	10	虽然	47(0.95)	2.6%	10	并	29(0.46)	2.5%
…	…	…	…	…	…	…	…	…	…	…	…
52	要是	1（略）	0.1%	62	以免	1（略）	0.1%	81	与其	1（略）	0.1%
Total		990	100%	Total		1810	100%	Total		1149	100 %

続いて、各サブコーパスにおける上位3語の接続詞に着目して分析する。JPM、JPHとCHNの上位3語の接続詞は合計5種類（“**和**”、“**所以**”、“**但是**”、“**而**”、“**但**”）である。この5種類に“**与**”を加え、計6種類の接続詞の使用頻度について、サブコーパス間で有意差があるかどうかを検証する（表4）。なお，“**与**”はCHNの上位9語には含まれている一方、JPM・JPHの上位9語には含まれていない。さらに、『新漢語水平考試（HSK）語彙[10]』に基づき、語彙別の接続詞の難度を示す（LはLevelの略）。例えば、並列関係を表す“**和**”と“**与**”はそれぞれHSK 1級（最低級）、HSK 4級（中級）の語彙であり、数字が大きくなるほど難度は高くなる。

[10] 漢語水平考試（略称：HSK）は中国の教育部が認定する国際的な中国語の語学検定試験である。HSKの公式サイトには、1級（最低級）～6級（最高級）で使用される語彙（5000語）のリストは《新汉语水平考试（HSK）词汇（2012年修订版）》の名称で配布されている。URL：http://www.chinesetest.cn/godownload.do〔最終アクセス日：2021年9月15日〕

表 4　高頻度の接続詞についての多重比較 [11]

接続詞	和訳	難度	Mean			Post Hoc Test（Sig.）		
			JPM	JPH	CHN	JPM vs. CHN	JPH vs. CHN	JPM vs. JPH
和 /cc	と	L1	9.65	7.72	2.98	.000***	.000***	.105
所以 /c	だから	L2	5.22	6.95	0.39	.000***	.000***	.031*
但是 /c	しかし	L2	3.61	3.17	0.64	.000***	.000***	.661
而 /c	しかし	L4	0.93	0.38	2.09	.000***	.000***	.048*
与 /cc	と	L4	0.17	0.21	0.82	.005**	.007**	.944
但 /c	しかし	L2	1.20	1.25	1.57	.229	.302	.663

結果は、CHN と比べて、“**和**”、“**所以**”、“**但是**”は、JPM と JPH の方が過剰使用している（$p < 0.05$; Mean: JP > CHN）のに対し、“**而**”と“**与**”は CHN の方が多く使用している（$p < 0.05$; Mean: JP < CHN）。“**但**”の出現頻度に関しては、3 つのコーパス間に有意差がない（$p > 0.05$）ことが分かった。石川（2008: 221）は、日本人大学生が書いた英語エッセイには、*so, but* などの接続詞が英語母語話者より過剰使用していることを報告しているが、本研究においても、日本人中国語学習者の中国語作文における“**所以**”（*so*）、“**但是**”（*but*）の過剰使用について石川（2008: 221）の指摘と同じような傾向が見られた。また、学習者の方が過剰使用している上位 3 語の接続詞（L1･L2 級別：“**和**”、“**所以**”、“**但是**”）は難度の低い語彙であり、学習者の方が過少使用している“**而**”と“**与**”（L4 級別）は難度の高い語彙である。例えば、並列関係を表す接続詞“**和**”と“**与**”に関しては、学習者は難度の低い“**和**”を母語話者より多く使用しているのに対し、難度の高い“**与**”をほぼ使用していない。図 5 のこれらの 5 つの接続詞の分布がサブコーパス間の比較により、学習者の作文では、難度の低い接続詞が頻繁に使用されていることも明らかになった。

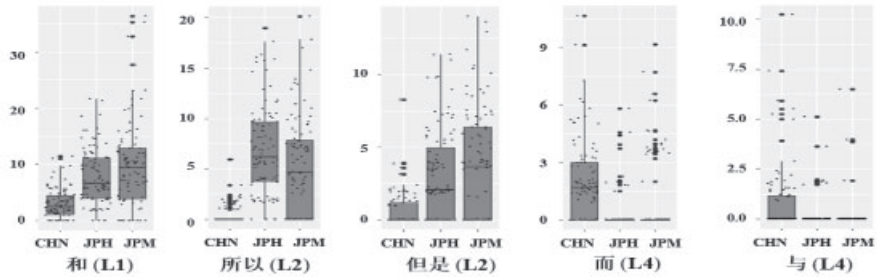

図 5　接続詞別の相対頻度の箱ひげ図

[11] NLPIR の「Chinese part-of-speech tag set」によって、「/cc」は並列関係を表す接続詞のタグであり、「/c」は並列以外の接続詞のタグである。URL：http://kgb.lingjoin.com/nlpir/html/readme.htm〔最終アクセス日：2021 年 11 月 15 日〕

表５　作文教科書の例文における５つの接続詞の相対頻度

	和	所以	但是	而	与
JPM の教科書	7.43	1.24	0.83	1.24	0.00
JPH の教科書	3.58	0.00	0.00	0.00	0.00

母語話者より、学習者の方が難度の低い接続詞を過剰使用し、難度の高い接続詞を過少使用している原因の一つは、学習者にとって難度の低い接続詞を使用しやすく、難度の高い接続詞を使用しにくいと考えられる。また、もう一つの原因は学習者が使用した教科書に関係があると考えられる。本研究では、学習者が使用した作文教科書における作文例文（それぞれ７篇）におけるこれらの５つ種類の接続詞の出現頻度を調べた（表5）。その結果、作文例文において難度の高い“**而**”、“**与**”の相対頻度が低く、使用例が極めて少ないことが確認された。教科書であまり使用されていない高難度の接続詞は、学習者が目にする機会が乏しく、学習者の産出語彙に含まれていないと推察される。

4.4 接続詞の出現位置

日本人英語学習者の英作文コーパスでは、英語母語話者の作文に比べ、文頭で接続表現が使われる割合が多いことが報告されている（今尾 , 2018）。中国語接続詞の場合においても、同じ現象が見られるかを確認するため、CasualConc（Imao, 2021）の「項目位置プロット」の機能を利用し、各コーパスの接続詞の文ごとの相対出現位置を可視化した（図 6）。「First」は文頭、「Last」は文末の意味で、その間は 10％ごと刻みになっており、「色が濃いほどその位置の割合が高くなっている」（今尾 , 2018）。図 6 を見ると、CHN、JPH、JPM コーパスにおいて、文頭で接続詞が使われている割合はそれぞれ 16.9%（194/1149）、24.8%、25.6% であるので、JPH と JPM では、文頭で接続詞が使用される割合が CHN より約 10％多いことが分かった。

図６　接続詞の文ごとの相対出現位置

表６　上位３語の接続詞の文頭での使用割合

Group	和	所以	但是	而	与
CHN	0.0%	31.0%	37.2%	14.4%	0.0%
JPH	0.0%	13.6%	40.1%	5.6%	0.0%
JPM	0.0%	15.9%	50.0%	18.2%	0.0%

最後に、各サブコーパスの上位３語の接続詞が文頭で使用される割合を調べる（表６）。まず、並列関係を表している“**和**”（主に単語や連語の並列に用いる）と“**与**”（主に名詞や代名詞の並列に用いる）は文頭で使用できない。反転関係を表すことができる“**而**”はJPMのほうが文頭で使われる割合が高いが、JPMには“**而**”の出現頻度（0.93）はCHN（2.09）ほど高くない。そして、結果を述べる際に用いる“**所以**”について、CHN（31.0%）の方が文頭で多く使用しているが、CHNにおける“**所以**”の出現頻度（0.39）はJPMとJPH（5.22、6.95）に比して非常に低い。それに対して、反転や逆接の意味を表すことに用いる“**但是**”は、JPM（50.0%）とJPH（40.1%）の方が文頭で多く使っている。JPMとJPHでは、“**但是**”の出現頻度（3.61、3.17）はCHN（0.64）より著しく高いため、これがJPMとJPHの方が文頭で接続詞の使用割合が多いことの要因だと考えられる。下記例文（6）及び例文（7）で示すように、学習者の作文には、句点の後ろに“**但是**”が置かれる例がよく見られる。日本語エッセイで用いられる接続詞で出現頻度が最も高いのは「しかし」であり、97%の「しかし」が文頭で使用されていることが報告されている（今尾 , 2018）ため、日本語「しかし」の文頭使用傾向が、JPMとJPHの中国語作文での“**但是**”を文頭で使用する傾向に影響している可能性があると考えられる。

（6）高中时，我属于篮球俱乐部。我不属于大学篮球俱乐部。**但是**在业余时间，我经常去体育馆打篮球。**但是**最近我忙于上大学，不打篮球。（JPM_74）

（訳：高校時代ではバスケットボール部に所属していました。私は大学のバスケットボール部に所属していません。**しかし**、暇さえあればよくジムに行ってバスケをします。**しかし**、最近は大学を通うことが忙しくてバスケをすることができません。）

（7）宫岛是日本三景之一。海中站的牌楼一定很漂亮，我们很期待。我们和我们车一起坐船去了。**但是**发生了一点儿也不想象的问题。那个牌楼正在施工，我们没看见了。真是太可惜了。**但是**宫岛还有世界遗产的严岛神社。这个神社的历史从6世纪开始的，所以值得一看。**但是**为了看那个牌楼，我们要去宫岛再一次。（JPH_72）

（訳：宮島は日本三景のひとつである。海の駅のパゴダはきっと美しいに違いないと、私たちはとても楽しみにしていました。私たちは私たちの車と一緒に船で乗って行

きました。しかし、想像に及ばない問題が発生しました。そのパゴダは工事中で見られませんでした。本当に残念でした。しかし、宮島には世界遺産である厳島神社もあります。この神社の歴史は6世紀に遡るので、一見の価値があります。しかし、パゴダを見るためには、私たちはもう一度宮島に行かなければなりません。）

5 結論と今後の課題

本研究では、出現頻度に着目し、日本人中国語学習者の作文における接続詞の使用実態を明らかにした。まず、接続詞の使用頻度と語彙多様性について、学習者が母語話者より全体的に接続詞を過剰使用しており、上級になると接続詞の語彙多様性が増える傾向が見られるものの、学習者が使用している接続詞の語彙多様性は母語話者より著しく低いことが確認された。次に、高頻度接続詞の使用状況について、学習者が作文で用いる接続詞に占める高頻度語の比率が母語話者に比べて高く、難度の低い接続詞が繰り返し使用されている一方で、難度の高い接続詞はほぼ使用していないことが明らかになった。また、接続詞の出現位置について、学習者は母語話者より文頭での接続詞の使用割合が約10%多いことが分かった。

なぜ接続詞の使用実態が以上のような傾向を見せるのか。すでに4.1と4.3で検討したように、その原因は、学習者が中国語接続詞の正しい用法や省略方略をまだ習得できていないことや、教科書に現れた接続詞の影響、母語である日本語の干渉などに求められるのではないかと思われる。そのため、日本の中国語教育では、学習者の母語を背景として考慮した上で、接続詞の教え方や教科書に掲載する接続詞の選定にもっと工夫する必要があるのではないかと考える。本研究の目的は、中国語教育の問題点を探り、教材の改善に資することであるが、課題も残っている。今後は、コーパスデータを拡大し、叙述文以外の作文（意見文など）や母語の異なる中国語学習者が書いた作文を加えることにより、接続詞の使用実態等をより詳細かつ母語横断的に分析、考察することができる。得られた研究結果や知見を今後の中国語教育現場に応用していきたい。

参考文献

崔建新, 张文贤（2002）. 不同语体下连词使用率的统计与分析,《第七届国际汉语教学讨论会论文选》, 389-404.

陈芳（2013）.《中高级水平日本留学生汉语连词使用特点及偏误分析》.（Student thesis）.（Retrieved from http://cdmd.cnki.com.cn/Article/CDMD-10165-1014138546.htm）.

苏鹰（2002）. 日语中连接句子和段落的接续词——与汉语连词的对比,《湖南大学学报（社会科学版》, 16（3）: 187-189.

盧濤（2008）「漢語連詞使用錯誤分析」『広島外国語教育研究』（11）, 185-196。

王蕊（2009）「日本語上級レベル学習者の接続表現の使用状況に関する調査―中国語母語話者のストーリーテリングテストを中心に―」『ポリグロシア』17, 117-128。

小林雄一郎（2017）『R によるやさしいテキストマイニング：機械学習編』オーム社。

髙野愛子・上村圭介（2017）「レジスター別出現頻度に基づく順接接続詞の文体差の評価：現代日本語書き言葉均衡コーパス（BCCWJ）の用例分析から」『語学教育研究論叢』（34）, 273-293。

石川慎一郎（2008）『英語コーパスと言語教育―データとしてのテクスト』大修館書店。

石川慎一郎（2012）『ベーシックコーパス言語学』ひつじ書房。

石黒圭（2008）『文章は接続詞で決まる』光文社新書。

石黒圭・阿保きみ枝・佐川祥予・中村紗弥子・劉洋（2009）「接続表現のジャンル別出現頻度について」『一橋大学留学生センター紀要』（12）, 73-85。

今尾康裕（2018）「日本の大学生英語学習者によるエッセイでの接続表現を探る：日本語エッセイ・英語母語話者によるエッセイと比較して」『言語文化共同研究プロジェクト』（2019）, 5-23。

Altenberg, B. & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S Granger (Ed.), *Learner English on computer.* (pp. 80–93). London: Longman.

Appel, R. (2020). A Contrastive Interlanguage Analysis of Linking Adverbials in EFL Writing: Identifying L*1* Related Differences. *Learner Corpus Studies in Asia and the World*, 4, 24-40.

Bolton, K., Nelson, G., & Hung, J. (2002). A corpus-based study of connectors in student writing: Research from the International Corpus of English in Hong Kong (ICE-HK). *International Journal of Corpus Linguistics, 7*(2), 165-182.

Chen, C. W. (2006). The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International journal of corpus linguistics*, *11*(1), 113-130.

Crewe, W. (1990). The illogic of logical connectives. *ELT Journal, 44*(4), 316–325.

Heino, P. (2010). *Adverbial connectors in advanced EFL learners' and native speakers' student writing*. Student Thesis, University of Stockholm, Stockholm.

Imao, Y. (2021). CasualConc (Version 2.1.6) [Computer software]. Available from https://sites.google.com/site/casualconcj/casualconc/.

Ishikawa, S. I. (2011). A corpus-based study on Asian learners' use of English linking adverbials. *Themes in Science and Technology Education*, *3*(1-2), 139-157.

Larsen-Walker, M. (2017). Can Data Driven Learning address L2 writers' habitual errors with English linking adverbials? *System*, *69*, 26-37.

Nakayama, S. (2021). A Corpus-Based Analysis of Japanese EFL Learners' Linking Adverbial Use. *The Educational Review, USA*, *5*(6), 164–171.

Narita, M., Sato, C., & Sugiura, M. (2004). Connector Usage in the English Essay Writing of Japanese EFL Learners. *Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC 2004*, *1*, 1171–1174.