The University of Osaka
Institutional Knowledge Archive

Title	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
Author(s)	今尾, 康裕
Citation	言語文化共同研究プロジェクト. 2019, 2018, p. 5-23
Version Type	VoR
URL	https://doi.org/10.18910/72740
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

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

The University of Osaka

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―

今尾康裕

大阪大学言語文化研究科
〒560-0043 豊中市待兼山町 1-8
Email: imao@lang.osaka-u.ac.jp

概要　本研究は，日本の大学生が英語と日本語で同じトピックについて書いたエッセイを集めたバイリンガルコーパスを利用して，同じ書き手が英語と日本語で書いたエッセイでの文レベルの接続表現を比較するとともに，同じトピックで書かれた別の日本語母語話者と英語母語話者による英語エッセイとも比較した。その結果，同じ書き手では，日本語よりも英語の方が一定文数あたりの接続表現の使用頻度が高いことがわかった。また，英語母語話者のエッセイでは，一定文数あたりの接続表現の使用頻度は学習者のエッセイよりも高かったが，高頻度で使われる接続表現の種類には大きな違いが見られ，日本の大学生のエッセイでも異なるサンプルを比較しても，接続表現の使用傾向にそれとは異なる違いが見られた。それと同時に，英語・日本語エッセイでの比較や統制条件が異なるエッセイの比較には接続表現以外の要素も考慮に入れる必要があることも示唆された。

キーワード　学習者コーパス，接続表現，第１・第２言語エッセイ比較

Exploring connector usage in essays written by university students in Japan
―Through comparisons with essays in Japanese and English essays written by English native speakers―

Yasuhiro Imao
Graduate School of Language and Culture, Osaka University
1-8 Machikaneyama-cho, Toyonaka, Osaka, 560-0043 Japan

Abstract This study examined the use of sentence-level (full-clause-level) connectors in essays written by Japanese college students using a English-Japanese bilingual corpus. Comparisons were made between English and Japanese essays written by the same students as well as between English essays of the bilingual corpus and English essays written by a different group of Japanese college students as well as native speakers of English. The results show that writers use more connectors per fixed number of sentences when writing in English than in Japanese. English native speakers use higher number of connectors per fixed number of sentences, but their choices of connectors are found to be quite different from learners'. The tendencies of the connector usage in English essay are also different between two groups of college students and from essays written by English native speakers. At the same time, the findings suggest that considering elements other than connectors needs to be examined when comparing essays written in English and in Japanese or essays collected under

different instructions

Keywords learner corpus, connector usage, L1-L2 essay comparison

1. はじめに

第二言語でのアカデミックライティングには様々な要素が影響しており，特に英語においては，単語レベルから，句，文，文と文の結びつき（結束性），さらにはムーブなどの談話レベルや文章の構成まで，幅広い観点から研究が行われている。学習者コーパスを利用した研究が多く行われるようになった 1990 年代からは，一般コーパスや論文コーパスとの比較，さらに，英語母語話者エッセイコーパスとの比較などが行われ，特に語彙要素の使用頻度の傾向などが対象となってきた (Granger, 1998; Hinkel, 2002 など)。

語彙要素の中でも，文と文をつなぐ表現である，logical connectors や linking adverbials と呼ばれる接続表現に関するコーパスを使った研究が数多く行われている。初期のものでは，Milton & Tsang (1993) が，中国人学習者のエッセイと一般コーパスである BROWN，LOB，およびコンピュータサイエンスの教科書コーパスとの logical connectors の使用頻度の比較を行っており，その後も多くの研究が行われてきた。しかしながら，そのほかの多くの学習者コーパスの研究と同じく，文レベルの接続表現の研究は，学習者エッセイと一般英語や学術論文・教科書などの熟達者の英語文章との比較が中心で，英語母語話者のエッセイを収集した ICNALE や NICE コーパスなどを利用した一部の研究を除いて (Ishikawa, 2011; Narita, Sato, & Sugiura, 2004)，同じトピックで書かれた英語母語話者のエッセイや学習者が母語で書いたエッセイとの比較がされることはあまり多くない。そこで，本研究では，日本の大学生が同じトピックで書いた日本語エッセイおよび英語エッセイでの接続表現の比較をするとともに，同じトピックで書かれた別の日本人大学生エッセイおよび英語母語話者のエッセイとも比較を行い，今後の詳細な分析のための下準備として，接続表現の使用傾向を探ることを目的とする。

2. 文レベルの接続表現の研究

Milton & Tsang (1993) は，上記のように，学習者の書いた文章と一般コーパス・教科書コーパスでの，logical connectors の指標頻度の比較を行った。頻度集計した logical connectors は，文法書 (Celce-Murcia, Larsen-Freeman, & Williams, 1983) の分類に基づいて選ばれた単語レベルの等位接続詞と接続副詞であり，使用頻度の比率に基づいて学習者が過剰使用する logical connectors を示すとともに，学習者が使用する EFL 教科書の影響にその要因を求めて，教科書での logical connectors の扱いを分析している。学習者による接続表現の過剰使用仮説を検証するために，Granger & Tyson (1996) は，フランス語話者の英語エッセイ (ICLE) とそれに相当する英語母語話者のエッセイ (LOCNESS) および一般英語コーパス (LOB) での接続表現の過剰・過少使用について検証した。接続表現は，Quirk ら (1985) の接続表現に基づいて，一部を削除した上で，それらに独自のものを追加して分析している。過剰・過少使用については，母語の影響は見られたものの，ドイツ語母語話者のエッセイで確認したところ，母語に関係のない逸脱も見られた。そのほかにも，学習者においては文頭での接続表現の使用が過剰であることや，一般コーパスと母語話者エッセイコーパスとでは使用傾向が異なり，対応するコーパスとの比較が重要であると指摘して

いる。

香港の英語学習者のエッセイと，イギリスの英語母語話者の学生が書いたエッセイでの接続表現の使用頻度を，英語母語話者の書いた学術文章のと比較した Bolton ら (2002) の研究では，英語学習者と英語母語話者のエッセイでは，一般の学術文章と比較して過剰使用される接続表現が異なることが示された。Hinkel (2002) は，東・東南アジアの言語およびアラビア語を母語とする英語学習者（6 言語）と英語母語話者のエッセイを様々な言語要素で比較した。文レベルの接続表現では，英語学習者は英語母語話者と比べて，文章の進行や推移を示す接続表現では，どの言語の母語話者もかなり過剰使用をしているが，論理や意味関係を示す接続詞では，言語によって多少の過剰使用はあるが，それほど大きな差はないことが報告されている。

Altenberg & Tapper (1998) はスウェーデン語母語話者の書いた英語エッセイと英語母語話者のエッセイおよび，スウェーデン語母語話者がスウェーデン語で書いた文章での接続表現のカテゴリごとの使用頻度を比較した。スウェーデン語で書かれた文章と英語母語話者のエッセイでは，接続表現の使用頻度に大きな違いが見られなかったことから，スウェーデン語と英語での接続表現の使用傾向は大きく違わないとして，スウェーデン語母語話者の英語エッセイでの使用頻度が異なることから，母語の影響がそれほどないであろうと結論づけている。

2000 年代に入る直前には Biber ら (1999) が，接続詞ではなく，副詞を含めて副詞的に文を接続する表現を linking adverbial とし，enumeration and addition, summation, apposition, result/inference, contrast/concession, transition の 6 つのカテゴリに分類した上で，レジスターごとに使用頻度が異なり，会話やアカデミックな書き言葉での使用頻度が多いことなどを示した。

これらを含めて，linking adverbial や conjunctive adverbials など，ラベルの違いはあるが，アカデミックレジスターや学習者エッセイでの接続表現の使用やその習得のためのコーパスを利用した教育に関する研究が多く行われてきた (Chen, 2006; Charles, 2011; Conrad, 1999; Dorgeloh, 2004; Garner, 2013; Ishikawa, 2010; Larsen-Walker, 2017; Lei, 2012; Liu, 2008; Shaw, 2009; Zareva, 2011)。これまで見たように，これらの研究の中にも，学習者が書く英語での接続表現の過大・過少使用などを分析しているものもあるが，先に述べたように Ishikawa (Ishikawa, 2010) や Narita ら (2004) などのトピックを統一してデータが収取されたコーパスを利用した一部の研究を除き，学習者のエッセイと母語話者もしくは熟達者が書いた論文などの異なるジャンルの文章を比較しており，書く目的や文章の長さ，内容などがあまりにも違うものを比較をすること自体にどれだけ意味があるのか疑問が残る。また，Altenberg & Tapper (1998) などのように L1 と L2 の比較をした研究も見られるが，同じ書き手の文章ではなく，L1 での一般的な接続表現の使用頻度から学習言語である英語の文章への影響を推測するものである。

そこで，本研究では，同じ書き手による英語と日本語のエッセイを比較する目的で構築された関西大学バイリンガルエッセイコーパス (KUBEC) を利用して，接続表現の使用について英日で比較するとともに，同じトピックでの英語母語話者の書いたエッセイとも比較する。これまでの研究では，L1 と L2 で同じトピックで書かれたエッセイの語彙使用に関して比較できるものがほとんどないため，探索的にデータを概観する。

3. 研究手法

3. 1 使用したコーパス

本研究での中心となるコーパスは，関西大学バイリンガルエッセイコーパス（KUBEC）(山西・水本・染谷, 2013) で，関西大学を中心として 6 つの国立・私立で集められた英語と日本語で書かれたエッセイコーパスである。対象者はそれぞれの大学の学部生で，40 分程度で英語でエッセイを書き，その後，同じく 40 分程度を使い，同じトピックで日本語のエッセイを書いた。日本語でエッセイを書く際は，英語のエッセイを日本語に訳すのではなく，同じ内容を日本語で書き直すように指示された。データ収集の中心となった関西大学では，NICE[1] (Nagoya Interlanguage Corpus of English) と ICNALE [2](The International Corpus Network of Asian Learners of English) で使われたエッセイトピックのすべてで書かれたエッセイが収集され，それ以外の大学では，ICNALE のトピックのみで書かれたものが収容されている。今回は，すべての大学で収集された ICNALE の「アルバイト」のトピックで書かれたエッセイのみを利用した。また，同じトピックで書かれた ICNALE の日本語母語話者と英語母語話者が書いたエッセイとも比較した。

3. 2 分析対象

今回の分析では，文レベルの接続表現を対象とした。具体的には，linking adverbials と文レベルの等位接続詞で，英語のものは，Biber ら (1999) で扱われた linking adverbials を中心として，過去の接続表現を対象とした研究で扱われたものを対象とした (Bolton et al., 2002; Liu, 2008; Shaw, 2009)。日本語エッセイの接続表現は，石黒ら (2009) を参考に，後述する MeCab によるタグ付けで接続詞とされた表現のうち文レベルのもの，および英語の接続表現に対応する表現を対象とした。なお，本論では，これ以降 linking adverbials などの接続表現を「接続副詞」として表記していく。

3. 3 テキストの下処理

分析に使用したテキストのうち，KUBEC の英語エッセイは TreeTagger[3] で，日本語エッセイは MeCab + UniDic でタグづけされていたが，スペリングエラーがそのままであったこと，ICNALE のエッセイと共通のタグがついていないこと，また，タグ付の精度やタグの扱いに問題が見られたため，スペリングエラーを修正した上で，英語エッセイには Stanford CoreNLP (Manning et al., 2014) で，日本語エッセイには MeCab + IPADic でタグ付けした。そのタグ付け処理の際に，文ごとの頻度などを集計するために 1 行 1 文の形式に変換し，文の区切れを目視で確認して修正した。その後，英語エッセイの接続副詞に関しては，収集したリストのものを正規表現による検索置換で複合語のものは 1 つの単語として _ でつなぐ処理をした上で共通のタグを付けた。等位接続詞に関しては，目視で確認し，文レベルのものとそれ以外で識別できるようにタグを修正した。また，特に前置詞などの誤用などで，接続表現として使われているが，リストに対応するものがない場合は，接続表現のタグを付けたが，修正は行わなかった。その他にも，今回の分析では対象としないが，複合前置詞を 1 つの前置詞として扱う処理をしている。このような処理をしたため，ICNALE

1. http://sgr.gsid.nagoya-u.ac.jp/wordpress/?page_id=965
2. http://language.sakura.ne.jp/icnale/
3. http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/

のエッセイにおいては，他の研究などとは総語数が少なくなっている。

日本語のエッセイでは，「接続詞」のタグがついているもののうち，文レベルの接続表現でないものに別のタグに付与し，文レベルの接続表現のみを「接続詞」とした。また，石黒ら (2009) を参考にして，それぞれの接続表現を検索し，複数の単語に分かれているものは，1 つの単語としてまとめて「接続詞」とした。その後，文頭の表現を中心に英語の接続表現に対応する副詞的な表現を目視で確認して，「接続詞」と判断できるものに「接続詞」のタグを付けた。接続詞のタグ付け終了後には，日本語特有の漢字，仮名，数値の表記のゆれなどを，品詞タグとは別の形式のタグを付けて元の表記を記録した上で修正し統一した。これ以外にも，石黒ら (2009) のリストにもあるように，一部の接続表現において，使われている助詞によって意味や機能が変わらないと判断できるものに関しては，タグを付けして元の単語を記録した上で統一した表記にした。

これらの一連の処理は，CasualTagger の開発版 (Imao, 2018b) を主として利用した。

3. 4 分析

接続表現の頻度を集計した後に，それぞれのコーパスに特徴的な接続表現を抽出して比較した。また，学習者の英語エッセイでは，L1 に関係なく接続表現を文頭に使う傾向が見られる (Granger & Tyson, 1996) ことから，使用頻度の多い接続表現においては文の中での出現位置を確認した。頻度集計およびその後の特徴語抽出のための統計処理には，CasualConc (Imao, 2018a) のファイル情報と Word Count の機能を利用し，CasualConc で扱えない統計処理などは，簡単な統計処理は Excel で，それ以外は R も使いながら行った。接続表現の文中での使用位置については，位置項目プロット機能を利用した。

4. 結果および考察

4. 1 コーパスの概要

表 1 に今回分析対象とした 4 つのコーパスの概要を示す。KUBEC のエッセイには，英日共にタイトルがついているが，タイトルは頻度集計やその後の分析に含めていない。

表 1 コーパスの概要

	エッセイ数	総語数	平均 異なり語数	平均 総語数	平均文数	1 文あたりの 平均語数
KUBEC E	654	194599	137.89	297.55	19.96	14.91
KUBEC J	654	249947	152.90	385.21	20.85	19.26
ICNALE NNS	400	87447	111.69	218.62	16.78	13.03
ICNALE NS	200	44155	130.25	220.78	8.81	25.06

KUBEC の英語エッセイ (KUBEC E) （297.55 語）と日本語エッセイ (KUBEC J) （382.18 語）ではエッセイの長さがかなり異なることがわかる。考えられる要因は様々なものがあるが，母語である日本語の方が自由に運用できるという可能性は当然考えられる。日本語で文と文（ここでは完全に独立した文という意味ではない）をつなぐ表現には，動詞の活用や助詞なども含め，接続表現以外の方法も多くあり，それらを使うことで接続表現の頻

度が低くなっているかもしれない。その一方で，英語においては，そこまでの運用能力を持たないため，言語としては日本語と同様に多様な表現が可能であるものの，学習者は定型の接続表現に過度に依存していることが考えられる。ただ，それ以外にも，英語と日本語の文法が異なることにより単語の切り分けが異なることや，英語エッセイでの複合語の処理などの影響，日本語の品詞タグ付けに利用した IPADic の切り分けの影響なども考えられるため，ここでの分析だけで結論づけられるわけではない。

英語エッセイの比較では，KUBEC E （297.55 語）に比べて，ICNALE の日本語母語話者のエッセイ（ICNALE NNS: 218.62 語）と英語母語話者のエッセイ（ICNALE NS: 220.78 語） の総語数が短くなっている。これは，ICNALE では，データ収集での制限時間が相対的にやや短めであることと，相互数を 200 から 300 語程度に統制していること (Ishikawa, 2011) が影響していると考えられる。また，KUBEC E （19.96 文）と KUBEC J （20.85 文）の平均文数は KUBEC J がわずかに多いがほぼ同じであり，総語数の差は一文あたりの語数の差（14.91 語，19.26 語）に起因していることがわかる。これは，KUBEC のエッセイは，英語でエッセイを書いたのちに日本語でエッセイを書いているため，ほぼ同じ文数となっている可能性も考えられるが，1 文中の単語数は，運用能力の違いの他にも英日の文法の違いなどが現れているだけの可能性も残る。KUBEC E （19.96 文）と ICNALE NNS （16.78 文）での平均文数を見ると平均総語数ほどは差がなく，1 文あたりの語数では KUBEC E（14.91 語）と ICNALE NNS（13.03 語）で，ICNALE NNS のエッセイの方が，KUBEC E のエッセイよりも，平均するとやや短い文で書かれている。ICNALE NS （8.81 文）の平均文数は ICNALE NNS （16.78 文）の半分強くらいであり，1 文あたり 25.06 語と，文数は少ないがかなり長い文を書いていることがわかる。

4. 2 接続表現の使用頻度

次に，英語エッセイで接続表現としてタグ付けした接続副詞（句）と文レベルの等位接続詞，日本語エッセイでの文レベルの接続詞を含む接続表現の使用頻度を集計した（表 2）。

表 2 接続表現の使用頻度

	エッセイ数	接続表現数	1 エッセイあたり	100 語あたり	10 文あたり
KUBEC E	654	6105	9.33	3.14	4.68
KUBEC J	654	4989	7.63	2.00	3.85
ICNALE NNS	400	3173	7.93	3.63	4.73
ICNALE NS	200	893	4.47	2.02	5.07

文をつなぐ接続表現を比較する際に，表 1 にあるように平均総語数が異なるため，単純に比較することが難しく，英語と日本語では単語の切り分けや文法的な語の使用が異なるため単純に比較できない。そこで，3 つの単位あたりの頻度で比較を試みた。まずは，1 エッセイあたりの頻度の比較である。KUBEC E （9.33 回）と KUBEC J （7.63 回）を比べると接続表現の使用総数からもわかるように，KUBEC E の方が多くなっている。ただ，英語と日本語の文法的な切り分けや語の持つ機能の違いなどもあるため，単純に使用頻度に差が出ているのかを判断することは難しい。英語でのエッセイを比較すると，1 エッセイ

あたりでは，ICNALE NS （4.47 回）に比べて，日本語母語話者のエッセイではほぼ倍の頻度（KUBEC E: 9.33 回，ICNALE NNS: 7.93 回）で接続表現が使われていることがわかる。また，1 エッセイあたりの総語数が多い KUBEC E の方が，ICNALE NNS よりも頻度が多いが，エッセイ自体の長さ，分の多さが影響していることも十分考えられる。

次に 100 語あたりの頻度で比べたが同様の結果となった。KUBEC E （3.14 回）と KUBEC J （2.00 回）では，総語数の多い KUBEC J の方が接続表現の数が少ないことからもわかるように，100 語あたりの頻度でも少なくなっている。英語エッセイを比べると，学習者エッセイ同士では，KUBEC E は 3.14 回で ICNALE NNS は 3.63 回であり，ICNALE NNS の方が多い結果となった。これには，ICNALE NNS の方が 1 文あたりの長さが短いことが影響してるのであろう。ICNALE NS では 100 語あたり 2.02 回で，単位語数あたりで見ても，接続表現の使用頻度が少ないことが示された。

ここで対象としている接続表現は文をつなぐことがその機能であり，文の数や文の長さにその使用頻度は影響されるため，10 文あたりの使用頻度を確認した。KUBEC E （4.68 回）と KUBEC J （3.85 回）では，KUBEC J の方が使用頻度が低く，同じ書き手が英語で書く場合は，日本語で書くときよりも接続表現をより頻繁に使うことがわかる。ただし，当然ながら言語の差があるため，厳密にはそう言えない可能性も残る。英語エッセイで比べると，KUBEC E （4.68 回）と ICNALE NNS （4.73 回）では，10 文あたりの使用頻度がほぼ同じという結果であった。つまり，文の長さは ICNALE NNS の方が短いが，文をつなぐ接続表現の使用頻度はほぼ変わらないということになる。ICNALE NS （5.07 回）では，学習者のエッセイと比べて，文あたりで比べると頻度が高くなっており，文と文をつなぐという点において，英語母語話者が接続表現をあまり使わないということではなく，学習者と同じ，あるいはそれ以上に接続表現を使うが，学習者が文として分けてしまうところを従属節や関係節，名詞節，あるいは前置詞句などを用いて長い文を書き，それらを接続表現を用いてつなげているのであろう。

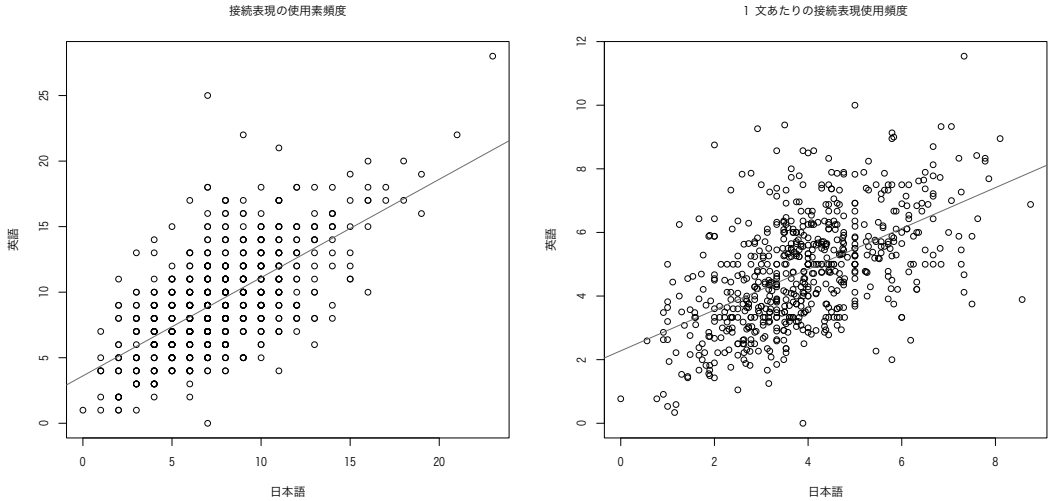

図 1 KUBEC E と KUBEC J での接続表現の使用頻度散布図

同一学習者が英日のエッセイで接続表現の使用傾向が同じかどうかを確認するために相関係数を確認したところ，素頻度では .64，10 文あたりの頻度では .54 であった。散布図

を見ても，ある程度ばらつきがあることが確認できる（図 1）これは，参加者それぞれの日本語の文章作成能力，英語運用能力などが異なるため，英語では日本語と同レベルの文章が書けていないことが考えられる。しかしながら，KUBEC には，参加者の文章構成能力や英語運用能力などの指標となる情報がないため，それらがどのように影響しているかは確認ができない。また，この中程度の相関が示すことは，参加者のすべてとは言わないが，多くが最初に書いた英語エッセイを単に日本語に訳すのではなく，指示通りに日本語で書き直しているのであろうことである。

実際に英語学習者（KUBEC E）の英語エッセイと同じ書き手の日本語エッセイ，英語母語話者（ICNALE NS）のエッセイからの抜粋を比較してみる。KUBEC は段落ごとの区切れがあるため，1 つの段落を抜き出してあるが，ICNALE は 1 パラグラフの文章が基本であるため，途中を抜粋している。

(1) **Finally**, we can say what we did during the university. We might have the image which college students are always hanging out. **In fact**, we tend to hang out a lot of time. **But** we have huge time during the university. **And** we have to decide everything on ourselves. **So** we should choose the study as a first choice. (KUBEC E #1)

(2) 最後は，大学で何をしてきたか言えるためである。私たちは，大学生はいつも遊んでいるというイメージを抱いているかもしれない。**実際**，私たちは多くの時間を遊びに費やす傾向がある。**しかし**，大学生は多くの時間を所有しており，すべてのことを自分自身で決断しなければならないのである。**だから**，私たちは勉強を一番目の選択肢にすべきである。(KUBEC J #1)

(3) **Second**, part time job might extend people connection. As for me, I had not much friends and the friends were only around the school. I didn't like using SNS, **so** my people network was only around the school and my family. I started to have a part time job just on this February. Until then my network was limited, **but** working in society, I'm getting know about not a student or teacher or family but working people in society. (KUBEC E #2)

(4) **２つ目に**，アルバイトを通して人脈が広がる可能性がある。学校での友人もあまり多くなかった私は，SNSもあまり好きでなく，人との関わりは学校と家族や親類のみに限られていた。アルバイトを始めた今年２月から，社会で働くということがどのようなものなのかを実際に知ることができた。(KUBEC J #2)

(5) A part-time job is also a great way to gain independence. With the money from a part-time job, students no longer have to rely on 'handouts' from their parents or others. **Instead** they can become independent and more prepared for the challenges of life beyond university. **However**, **above all**, a part-time job looks great on a resume, which is very important in an increasingly competitive job market. (ICNALE NS)

同じ書き手の英日のエッセイを比較すると，使用頻度が同じ程度の例（ (1) と (2)）では，ほぼ同じ構成で意味的に対応する接続表現が使われている（英語：6 文で 5 回，日本語：5 文で 3 回）。段落の最初の finally が「最後は」となって，副詞的な接続表現とは言えないためにそのように分類されていないが，実質的には同じ役割と言えなくはない。そう考えると，英語の接続表現のうち and に対応するものだけがないが，動詞の活用（おり，連用形）で文をつなげている形になっている。これは，母語である日本語で書いているため，接続表現を用いない表現方法が使えたということであろう。ただし，そうであったとしても，日本語の文章として読んでも接続表現が過剰に使用されている印象は受ける。使用頻度が大きく異なる例（ (3) と (4)）では，英語のエッセイでは 5 文で 3 回の使用だが，日本語は 3 文で 1 回である。厳密に同じ内容ではないが，日本語では接続表現を使わずに文がつなげられていることがわかる。

英語の文章を比較すると，英語学習者の書いた例 (1) は，短い単文の連続（6 文）で，その多くの文頭（5 回すべて）で接続表現が使われており，かなり接続表現が目につく。もう一つの例 (2) では，短い節を等位接続詞でつなげているのが確認できる。英語母語話者の書いた例 (5) は 4 文でそのうち 2 文には文頭で接続表現が使われているが（3 回），単文は少なく，前置詞句や句レベルの等位接続，関係節などを使って長い文が書かれており，接続表現で始まる文が 2 文が連続してはいるものの，これらの文例から読み手が受ける印象は大きく異なるであろう。

4. 3 使われている接続表現の比較

ここまでは接続表現全体の使用頻度の比較をしたが，高頻度で使われる接続表現を見ていく。表 3 は，KUBEC E と KUBEC J で使われている接続表現の上位 10 までの 100 語ごとと 10 文ごとの相対頻度である。

表 3 KUBEC の英日エッセイの頻度上位 10 接続表現

KUBEC E			KUBEC J		
	100 語ごと	**10 文ごと**		**100 語ごと**	**10 文ごと**
and	0.49	0.74	しかし	0.26	0.50
however	0.34	0.51	また	0.17	0.33
but	0.32	0.47	例えば	0.12	0.23
so	0.30	0.45	まず	0.11	0.21
therefore	0.15	0.23	そして	0.11	0.21
for example	0.14	0.21	なぜなら	0.09	0.17
in conclusion	0.10	0.15	だから	0.08	0.15
first	0.10	0.15	さらに	0.08	0.15
second	0.09	0.14	次に	0.06	0.12
moreover	0.08	0.12	しかしながら	0.06	0.12

KUBEC E では，2 番目の however を除き，and，but，so などの等位接続詞が頻度上位になっている。これらは，10 文あたりおおよそ 0.5 回程度かそれ以上の頻度で，1 エッセイあたりの平均文数がほぼ 20 であったことを考えると，平均して 1 エッセイあたり 1 回程度

は使用されているということになる。また，それ以下の接続表現は，情報のまとめや追加をする therefore や moreover，例を挙げる際の for example，文章の構成に関わる first，second，in conclusion などとなっている。論点を挙げる際の first，second などは，firstly，secondly などの表現は別の表現として集計しているため，このような列挙する表現の頻度はもう少し多いことになる。

KUBEC J では，最上位の「しかし」が 10 文あたりで 0.5 回程度で，1 エッセイあたりではおよそ 1 回程度となっている以外は，英語の上位 4 接続表現の頻度よりも低くなっている。接続表現の総頻度自体が英語よりも少ないことを考えると，英語での頻度よりも低いことは当然とも考えられるが，日本語でのエッセイでは接続表現が英語でのエッセイよりも一部の表現に頼ることが少ないとも言える。実際，英日では言語が異なるためまったく同じ基準で接続表現を抽出しているわけではないため純粋に頻度の比較ができるわけではないが，接続表現の種類の数はほぼ同じであることから（英: 137，日: 139），日本語の方が英語よりも特定の表現への偏りが少ないことがわかる。これは母語であるため日本語の運用能力が高く，決まり切った特定の表現以外も使えるためである可能性が考えられる。使われている接続表現を見ると，英語の上位にある列挙のための first，second に当たるものよりも，前後の文をつなぐような表現が多い。

KUBEC と ICNALE の学習者と母語話者のエッセイを比較するために，ICNALE NNS と ICNALE NS の使用した頻度上位の接続表現を確認する（表 4）。

表 4 ICNALE NNS と NS の頻度上位 10 接続表現

ICNALE NNS			**ICNALE NS**		
	100 語ごと	**10 文ごと**		**100 語ごと**	**10 文ごと**
and	0.67	0.88	and	0.80	2.00
so	0.65	0.84	but	0.27	0.67
but	0.59	0.77	so	0.19	0.47
for example	0.23	0.30	however	0.10	0.26
second	0.16	0.21	therefore	0.08	0.20
first	0.16	0.21	for example	0.06	0.15
however	0.11	0.14	also	0.06	0.15
of course	0.10	0.13	finally	0.03	0.09
therefore	0.09	0.11	next	0.02	0.06
third	0.08	0.11	furthermore	0.02	0.06

10 文ごとの頻度で見ると，ICNALE NNS の傾向としては，等位接続詞の and，so，but がそれ以外と比べて倍以上の頻度で使われていることがわかる。KUBEC E と比べると，in conclusion と moreover の代わりに of course と third が入っている他は頻度は異なるが同じ接続表現が上位を占めている。KUBEC E と ICNALE NNS を 10 文ごとの頻度で比較すると，KUBEC E と比べて ICNALE NNS では，however の頻度が低くなっている（KUBEC E: 0.51，ICNALE NNS: 0.14）。また，ICNALE NNS での and，so，but の使用頻度 (0.88，0.84，0.77) は，KUBEC E (0.74，0.45，0.47) よりもそれぞれ高くなっている。

ICNALE NS は，and の 10 文ごとの頻度が 1.99 であり，KUBEC E (0.73) や ICNALE NNS

(0.87) の 2 倍以上となっており突出して高い。ただし，平均文数は半分程度であるため，1 エッセイあたりにするとそこまで多いというわけではない。等位接続詞の but と so は，頻度の違いはあるが，KUBEC E や ICNALE NNS と同様に副詞的な接続表現と比べると高頻度で使われているのがわかる。これら以外の接続表現では，also，next，furthermore などの文をつなぐ表現が多く，ポイントを列挙する表現などが finally のみであった。ただし，あくまでも上位 10 までに入っていないだけで，そのような表現がほとんど見られないということではない。

ここまでは，頻度に基づくとはいえほぼ印象で多い少ないの比較をしてきたが，統計的に確認するために，対数尤度とマン・ホイットニー検定（MWU）によるキーワード抽出を試みた。前者は CasualConc の Word Count の機能で素頻度を基にして計算し，後者はファイル情報に付いている機能を利用して 100 文あたりの頻度を基にして計算し，最低頻度が 10 以上で，対数尤度は，$p = .01$ に当たる 6.63 以上，マン・ホイットニー検定も同様に $p = .01$ を基準として抽出した。表 5 は，KUBEC E と ICNALE NNS でのキーワード統計で，それぞれに特徴的と判断された接続語のリストである。対数尤度でのキーワード抽出は，通常大規模な参照コーパスに対して行うが，ここでは，小規模なコーパス同士でのキーワード抽出となるため，結果はあくまでも参考程度に扱う。

表 5 KUBEC E と ICNALE NNS の特徴語

KUBEC E		**ICNALE NNS**	
対数尤度	**MWU**	**対数尤度**	**MWU**
however	however	so	so
in conclusion	in conclusion	but	but
to sum up	to sum up	because	because
therefore	therefore	and	for example
in addition	in addition	for example	and
moreover	also	second	second
on the other hand	moreover	of course	third
also	on the other hand	third	of course
thus	first of all	first	
hence	thus	that is to say	
first of all	secondly		
for one thing	furthermore		

KUBEC E と ICNALE NNS を比べると，KUBEC E に特徴的と判断された接続表現は，アカデミックな文章に見られる接続副詞などが多い一方，ICNALE NNS に特徴的と判断されたものは，等位接続詞や first, second, third の要点の列挙に用いられる表現などとともに，文レベルの接続表現として誤用された従属接続詞 because も見られる。このような結果となった要因としては，KUBEC E のエッセイを書いた学生の方が ICNALE NNS のエッセイを書いた学生よりもアカデミックな書き言葉を身につけている，つまり，アカデミックな書き言葉の運用能力が高いことが考えられる。ただ，ICNALE NNS のエッセイの方が語数の統制が取られて平均で 100 語ほど短く，要点の列挙をしてあまり議論を掘り下げるような分量がないため，情報を足していくための in addition，moreover，also などの接続表

現や therefore や thus などの議論をまとめるような表現が比較的使われていないという可能性も否定できない。また，基本的には 1 段落で書かれているため，複数の段落で書かれている KUBEC E に比べて，要点を列挙する表現は使っても結びの段落などの最初によく使われる in conclusion や to sum up もあまり使われないのであろう。しかしながら，対数尤度の値も比較的高いものや，MWU の効果量の大きいものもは上位 2，3 に限られ，2 つのコーパスの違いはそれほど大きいとは言い切れない。

KUBEC E と ICNALE NNS の ICNALE NS に対しての有意な特徴語を示したのが表 6 である。これら 2 つの学習者コーパスに対する ICNALE NS の特徴語は and のみであった。熟達者の書く文章で and がより高い頻度で用いられることは，これまでの研究でも報告されている (Shaw, 2009)。対数尤度を使う際に，規模の小さい ICNALE NS を参照コーパスにすることに問題がないとも言えないが，MWU の結果とも大きく違わないことから，大きな問題はないであろう。

表 6 ICNALE NS に対する KUBEC E と ICNALE NNS の特徴語

KUBEC E		**ICNALE NNS**	
対数尤度	**MWU**	**対数尤度**	**MWU**
however	however	so	second
moreover	in conclusion	second	so
in conclusion	moreover	but	first
second	second	first	for example
first	first	for example	because
first of all	first of all	because	of course
then	for example	of course	third
to sum up	in addition	third	then
for example	then	then	but
in addition	secondly	moreover	moreover
of course	of course	in fact	in fact
so	to sum up	first of all	first of all
thus	third	actually	actually
therefore	thus		
in fact	therefore		
secondly	in fact		
third	firstly		
because	on the other hand		
in other words	because		
actually	so		
firstly	actually		
hence	in other words		
on the other hand			

これらのリストは，KUBEC E と ICNALE NNS のお互いの特徴語をある程度反映していると考えられる。KUBEC E には，多くの高頻度接続副詞が含まれ，従属接続詞の so や，ICNALE NNS の特徴語であった first，second，third なども入っている。つまり，KUBEC E のエッセイでは，ICNALE NS に対して，従属接続詞に関しては，so の使用が多いだけだ

が，接続副詞はその多くが過剰使用されていることがわかる。その一方で，ICNALE NNS では，KUBEC E との比較でも特徴語として抽出された接続表現に加えて，moreover，in fact，actually など，情報を加えていくために使われる副詞が多く使用されていることがわかる。

これらの要因としては，単純に接続表現が不必要なところで使用されている可能性が大いに考えられるが，これまでも述べたように，KUBEC の場合は ICNALE のエッセイに比べて平均文数が 2 倍以上になっており，文と文をつなぐ役割である接続表現の単純な使用頻度が高くなっていることも貢献していると考えられる。KUBEC のエッセイ収集の中心である関西大学ではライティングの授業でのデータ収集であったこともあるが，それ以外の大学でもライティングの指導を受けている可能性があり，接続表現を使用して文章を構成していくことだけでなく，同じ表現を繰り返さないことも学んでいれば，基本的な等位接続詞以外の使用頻度が高くなることも必然的に起きる可能性が十分ある。単純にエッセイ自体が短く，文の数が少なければ，さほど多くの異なる接続表現を使わないうちにエッセイが書き終わってしまう可能性が高いためである。

4. 4 接続表現の出現位置

これまでの研究で，学習者の書く英語の文章では，文頭以外で使うことが可能な接続表現でも，その多くが文頭で使われるということが報告されているため (Narita, Sato, & Sugiura, 2004)，接続表現の文の中での出現位置を確認する。図 2 は各コーパスの接続表現の 1 文中の相対出現位置を表したもので，最初のマスは文頭，最後のマスは文末で，その間は 10% 刻みになっている。色が濃いほどその位置の割合が高くなっている。

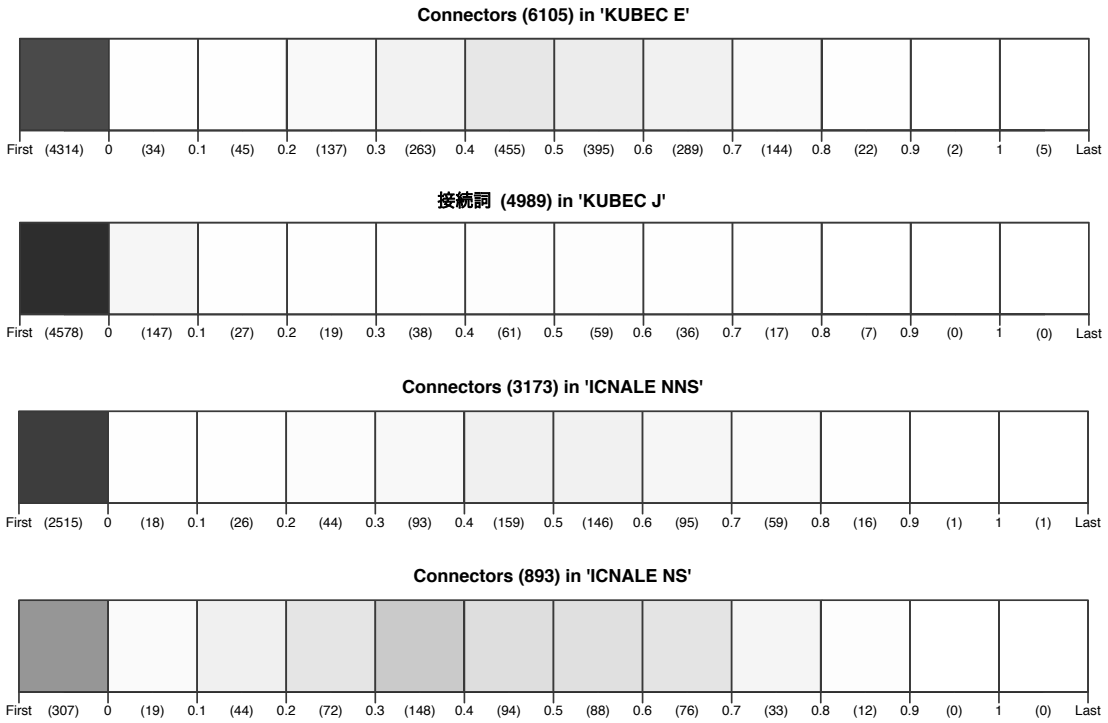

図 2 接続表現の相対出現位置と頻度

英日の学習者エッセイでは，明らかに文頭で接続表現が使われる割合が多くなっているが，英語母語話者のエッセイでは，文の中程にも多く現れている。具体的にどの程度文頭で使われているかを表 7 に示す。英語のエッセイに関しては，等位接続詞と linking adverbials などの接続副詞的表現（接続副詞）を分けて集計している。

表 7 接続表現の文頭での使用割合

	接続表現	等位接続詞	接続副詞
KUBEC E	71%	31%	93%
KUBEC J	92%		
ICNALE NNS	79%	65%	96%
ICNALE NS	34%	6%	81%

英語のエッセイに関しては，KUBEC E では，全体で 71% の接続表現が文頭で使われているが，等位接続詞では 31% で，接続副詞などの表現では 93% が文頭に位置している。ICNALE NNS では，KUBEC E よりも文頭で使われる等位接続詞の割合が 65% と大きくなっており，接続副詞では，KUBEC と同様の 96% であった。英語母語話者のエッセイでは，等位接続詞が文頭で使われるのはそのうちのわずか 6 % であり，その他の接続表現では 81% であった。日本語のエッセイである KUBEC J では英語と同様の区別をしていないが，接続表現全体の 92% が文頭での使用であり，学習者の英語エッセイでの接続表現と比べても文頭のみでの使用割合が高くなっている。英日の接続表現の使用位置については，言語の違いがあるため純粋に比較することは難しく，他の言語を母語とする学習者でも文頭での接続表現の使用が多い傾向が報告されていることから，それだけが要因とは言えないが，日本語での接続表現の使用傾向が英語エッセイでの文頭接続表現の過剰使用に何らかの影響を与えている可能性は十分考えられる。

最後に，文ごとの相対頻度の合計で上位 8 までの接続表現の文頭での使用割合を確認する（表 8）。

表 8 頻度上位 8 接続表現の文頭での使用割合

	KUBEC E	ICNALE NNS	ICNALE NS
and	17%	52%	1%
but	34%	64%	12%
first	99%	100%	100%
for example	93%	93%	89%
however	84%	90%	64%
second	100%	100%	67%
so	49%	78%	20%
therefore	91%	97%	51%

まず，どのコーパスでも最も高頻度で使われている接続表現の and を見る。長く規範文法やスタイルガイドなどで文頭での使用を避けるべきものとされてきた。その傾向は変わってきているとされているが，Biber ら (1999) によると，アカデミックライティングでは

文頭での and の使用頻度は非常に低い。ICNALE NS では 1% となっており，英語母語話者の書いたエッセイでも文頭での使用が避けられる傾向が見られる。10 文ごとの使用頻度ではそれほど大きな違いが見られない KUBEC E (0.74) と ICNALE NNS (0.88) ではあるが，文頭での使用割合を見ると，KUBEC E では 17% で多少見られる程度だが，ICNALE NNS では 52% と文レベルの接続詞として使われる and の半数を超えている。

これまで見た KUBEC E の例 (1) のように，おそらく運用能力のあまり高くないが学習者のエッセイでよく見られるように，ICNALE NNS のエッセイでも，短い文を and でつないで，それぞれを独立した文として続けていく例が多く見られた（例 (6)）。

(6) They are my precious friends except job time. **And** I learn etiquette or correct tongue, too. **And** Thanks to this experience, I grow up my sense of responsibility. (ICNALE NNS)

英語母語話者のエッセイでは，例 (7) のように長い文を and でつないでさらに長い文にする例が数多く見られた。これが，文ごとの and の使用頻度が高い要因になっている。

(7) It is well known that students to graduate from college with no work experience have an extremely tough time finding employment, **and** along with this, the employment which they can find is not as lucrative or as engaging as a work that can be found if they only had some work experience. (ICNALE NS)

次の but も and と同じように，ICNALE NNS では 64% が文頭での使用であり半数を超えている。KUBEC E と ICNALE NS でも，and よりは文頭での使用の割合が多いが，ICNALE NNS のような過剰とも言える使用割合ではない。10 文ごとの使用頻度も ICNALE NNS では KUBEC E と ICNALE NS の倍ほどであり，実際の使用状況を確認しても，and の場合と同様に，短い文を but を使ってつないでいくような例が多く見られた（例 (8)）。同様の傾向は，同じ等位接続詞の so でも見られた。また，ここでの例であるように，but の後にコンマが置かれる例が見られるが，これは日本語で逆接を表す接続詞の「しかし」の後に読点が置かれる影響があるのではないかと考えられる。

(8) **But**, I think there are more good things than bad things. I started part time job the other day. Now I am anxious **but** I expect. I will experience a serious ploblem. **But** then I will develop. So I agree with what taking part time job. (ICNALE NNS)

接続副詞を見ると，5 つの内，first はほぼ文頭で for example も同様の傾向が見られる。これらとは異なる傾向の second については，ICNALE NS での頻度が 6 回であり，ここでの割合が全体の傾向とするのは難しい。残りの 2 つ，however と therefore は，ともに学習者のエッセイでは高い割合で文頭で使用されているが，英語母語話者の場合は，文頭以外での使用割合が高いことがわかる。これら 2 つの接続副詞は，前の文とつなぐ役割であるが，主語の後ろなどに配置される形（例 (9) および例 (11)）と，2 つの文をコロンや等位接続詞でつないで 2 文（節）目の文頭に用いる形が多く見られ，相対位置をみると，however

では前者が therefore では後者の使い方が多いことが確認できる（図 3）。

(9) The one bad thing, **however**, is that as expected, I do sometimes have some problems balancing my school and work life. (ICNALE NS)

(10) This experience may translate to job opportunities after graduation or not; **however**, the experience of working itself and the added insight into the reality of the working world will always be personally beneficial. (ICNALE NS)

(11) The onus is **therefore** upon this student to determine its importance. (ICNALE NS)

(12) They would soon come to understand that they have to do this consistently and **therefore**, have some discipline and pride in the work that they do. (ICNALE NS)

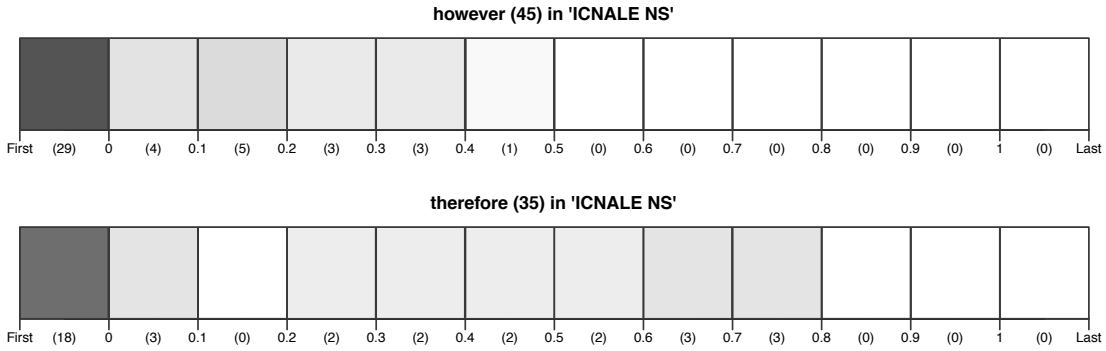

図 3 ICNALE NS のエッセイにおける however と therefore の相対位置と頻度

表 9 日本語エッセイでの頻度上位 8 接続表現の文頭での使用割合

	KUBEC J
しかし	97%
また	85%
例えば	95%
まず	99%
そして	93%
なぜなら	88%
だから	95%
さらに	92%

英語エッセイとの単純比較はできないが，KUBEC J の日本語エッセイでの頻度上位 8 までの接続表現の文頭での使用割合（表 9）を確認すると，文頭での使用がすべての表現で 80% を上回り，多くは 90% 以上であった。学習者の接続表現を文頭で使用する傾向は日本語母語話者に限ったことではないため，日本語での接続表現の文頭使用が日本語母語話者の英語エッセイでの接続表現の大きな要因になっているすることは難しいが，少なからず

影響がある可能性は考えられる。

5. おわりに

本稿では，日本の大学での英語学習者によって英語と日本語で同じトピックについて書かれたエッセイを集めたバイリンガルコーパスと，同じトピックで書かれた日本語を母語話者とする学習者と英語母語話者のエッセイのコーパスを比較し，言語間の比較，コーパス間の比較，英語学習者と英語母語話者の比較を行った。英語と日本語では文の構造も異なることから単純な比較はできないが，学習者によっては英日で同じような使用傾向である例も使用傾向が異なる例も確認でき，使用頻度に関しては全体として中程度の相関があった。英語と日本語の運用能力の外部指標がないために，使用傾向の差異が何に起因するかの特定は難しいが，ここで扱った接続表現以外の文（節）をつなぐ表現，英語でいえば，従属節や副詞句，日本語では助詞なども含めて分析する必要があるであろう。英語のエッセイを比較する場合にも，英語の運用能力などの指標があることが望ましい。また，同じトピックで同じような制限時間などの条件で集めたエッセイであっても，そのわずかな違いや統制の加減によって，文章レベルの構成などには大きな差が出ることも確認でき，語彙だけでなく文法などの言語使用の比較を含めて，条件の統制やデータ収集の種子などの重要さが再確認できたと同時に，同じトピックであるからと言って，言語要素などの頻度比較は慎重に行う必要があることもわかった。

今後は，英日の比較を中心に，どの要素が比較可能か，また，比較するために揃えるべき条件はどのようなものがあるかを精査した上で，書かれたエッセイを質的に考察することも含めて研究を行っていきたい。

参考文献

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.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of spoken and written English*. Harlow, Essex: Pearson Education.

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.

Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1983). *The grammar book: An ESL/EFL teacher's course*. Boston: Heinle & Heinle.

Charles, M. (2011). Adverbials of result: Phraseology and functions in the problem–solution pattern. *Journal of English for Academic Purposes*, *10*(1), 47-60. doi:10.1016/j.jeap.2011.01.002

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.

Conrad, S. M. (1999). The importance of corpus-based research for language teachers. *System*, *27*(1), 1-18.

Dorgeloh, H. (2004). Conjunction in sentence and discourse: Sentence-initial and and discourse structure. *Journal of Pragmatics*, *36*(10), 1761-1779. doi:10.1016/j.pragma.2004.04.004

Garner, J. R. (2013). The use of linking adverbials in academic essays by non-native writers: How data-driven learning can help. *CALICO Journal*, *30*(3), 410-422. doi:10.11139/cj.30.3.410-422

Granger, S. (Ed.). (1998). *Learner English on computer*. New York: Addison Wesley Longman.

Granger, S. & Tyson, S. (1996). Connector usage in the english essay writing of native and non-native EFL speakers of english. *World Englishes*, *15*(1), 17-27.

Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*. Mahwah, NJ: Routledge.

Imao, Y. (2018a). CasualConc (Version 2.0.8) [Computer software]. Available from https://sites.google.com/site/casualconc/

Imao, Y. (2018b). CasualTagger (Version 1.0.1) [Computer software]. Available from https://sites.google.com/site/casualconc/

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.

Ishikawa, S. (2010). A corpus-based study on asian learners' use of english linking adverbials. *Themes in Science and Technology Education*, *3*(1-2), 139-157.

Ishikawa, S. (2011). A new horizon in learner corpus studies: The aim of the icnale project. In G. Weir, S. Ishikawa, & K. Poonpon (Eds.), Corpora and language technologies in teaching, learning and research (pp. 3-11). Glasgow, UK: University of Strathclyde Publishing.

Larsen-Walker, M. (2017). Can data driven learning address L2 writers' habitual errors with english linking adverbials? *System*, *69*, 26-37. doi:10.1016/j.system.2017.08.005

Lei, L. (2012). Linking adverbials in academic writing on applied linguistics by chinese doctoral students. *Journal of English for Academic Purposes*, *11*(3), 267-275. doi:10.1016/j.jeap.2012.05.003

Liu, D. (2008). Linking adverbials: An across-register corpus study and its implications. *International Journal of Corpus Linguistics*, *13*(4), 491-518. doi:10.1075/ijcl.13.4.05liu

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & Mcclosky, D. (2014). The stanford corenlp natural language processing toolkit. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 55-60.

Milton, J. C. P. & Tsang, E. S.-C. (1993). A corpus-based study of logical connectors in EFL students' writing: Directions for future research. In R. Pemberton & E. S. Tsang (Eds.), *Studies in lexis* (pp. 215-246). Hong Kong: HKUST Language Center.

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

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive Grammar of the English language*. Essex: Pearson.

Shaw, P. (2009). Linking adverbials in student and professional writing in literary studies: What makes writing mature. In M. Charles, D. Pecorari, & S. Hunston (Eds.), *Academic writing:*

At the interface of corpus and discourse (pp. 215-235). London: Continuum.

山西博之・水本篤・染谷泰正 (2013).「関西大学バイリンガルエッセイコーパスプロジェクト―その概要と教育研究への応用に関する展望―」『関西大学外国語学部紀要』, *9*, 117-139.

Zareva, A. (2011). 'and so that was it': Linking adverbials in student academic presentations. *RELC Journal*, *42*(1), 5-15. doi:10.1177/0033688210390664