

台湾における日本語学習者の作文の計量的分析 - *jWriter* を使用して -

A Quantitative analysis of compositions written by learners of Japanese as a second language in Taiwan -Use of *jWriter*-

太田 栄次

Eiji OHTA

要 旨

本研究では、筆者が担当する台湾人日本語学習者の文章作成の言語面での習熟度や達成度を把握するため、自動評価システム *jWriter* を用いて、学習者の作文について計量的分析をおこなった。分析課題は以下の2点である。課題①筆者が担当する台湾人日本語学習者2年次生の上学期の作文と下学期の作文を分析し、文章の変化について計量的分析を行うこと。課題②台湾人日本語学習者下学期の作文と日本人大学生の作文を比較し、筆者が担当する台湾人学生の日本語文章力の現状を把握すること。結果、課題①では、1文の形態素数の増加、使用語彙の難易度の上昇により、評点の有意な上昇を認めた。また、使用される品詞や助詞の多様化や動詞の使用や接続詞、連体詞の使用の増加などから、単純な文章から、複文などのより複雑な構造を持つ文章が多く使用されるようになったことが示唆された。課題②では、下学期の台湾人学生の作文における文章の長さや使用語彙の難易度は日本人大学生には及ばないこと、台湾人日本語学習者の文章は、助詞の多様な使用、動詞の使用頻度の向上という点において、さらなる改善が必要であることが示唆された。

キーワード：台湾の日本語学習者、日本語作文、計量的分析自動評価システム *jWriter*

1. はじめに

言語教育及び学習において評価は学習者の言語能力の習熟度や達成度を把握する手段として重要な役割を果たしている。例えば、学習者にとっては、学習のモチベーションにもなりうるし、教育する側は、そのような評価を見ながら自身の教育内容を見直したりすることもできることなどがあげられる。また近年では、説明責任が重視されている背景などもあり、学生を評価する際に、十全な説明が求められているが、客観的な評価は、そのよ

うな要求にも応えることができる。

日本語学習者の作文評価をより公平に行うためには、多角的に視点で作文を評価する必要があり、その意味で、一般的には日本語学習者の作文は「言語面」「内容面」「構成面」の側面から行われる場合が多い。ただし、いずれの面においてもそれぞれの評価基準が不明確であるか、統一されていない場合、評価の公平性及び一貫性が損なわれる可能性がある。評価者間の主観性を最小限に抑え、客観的な評価を行うために、共通の評価基準の策定が必要である。一方で作文の評価に限らず、評価というものはその結果に差が生じやすいものであり、評価に主観的なものが紛れ込むことを排除することは極めて難しく、例えば、採点者内の影響では、ハロー（光背）効果、評定の系列効果、課題選択、採点の手間、採点者間で想定される影響では、採点項目、採点者の属性などが採点に影響することが一般的に知られている。

そのような中で、客観的な評価システムとしての自動評価システムが注目されており、日本語教育においても、いくつかの自動評価システムが見られるようになってきた。それぞれのシステムは、多様な言語統計量を用いて言語の特徴を捉え、その難易度や習熟度を評価し、日本語教育や評価の支援のために利用されている。ただし、日本語教育におけるライティングの自動評価システムに限定すると管見の限りでは「*jWriter*」「*Good Writing Rater*」に限られる。ただし、「*Good Writing Rater*」はアカデミックライティングを評価するものであり、日本語学習者の一般的な作文評価ということになると、現在のところ *jWriter* しかない。

2. 日本語学習者作文評価システム *jWriter* (<https://jreadability.net/jwriter/>)

jWriter は日本語学習者が書いた作文の文章の習熟度及び論理性を、計量テキスト分析の手法を用いて自動評価するシステムである（李・長谷部・村田2019）。*jWriter* によって得られる言語統計量は、総文字数、文数、段落数、延べ語数、異なり語数、平均文長（一文の平均延べ語数）、語種別語彙数、難易度別語彙数、タイプトークン比（語の多様性を示す指標、以下、TTR）、接続表現数等の45種である。そのうち、平均文長、総文字数、動詞数、難易度別語彙数、TTR、和語数、漢語数によって作文の習熟度が算出される。推定の公式は、「 $1.637 + \text{平均文長} \times 0.045 + \text{中級後半語} \times 0.021 + \text{TTR} \times -0.430 + \text{動詞} \times 0.015 + \text{中級前半語} \times 0.011 + \text{総文字数} \times -0.004 + \text{和語} \times 0.007 + \text{漢語} \times 0.007$ 」（ $R^2=0.760$ ）である。

つまり文に関する項目は1文の長さや動詞の使用数で測り、語彙の使用についてはレベル別語彙の使用率や和語や漢語の使用率などでとらえたうえで、これらの変数を考慮して

習熟度を算出するといったものである。比較的単純な指標を組み合わせて計算されているため、どの言語母語話者も使用することができ、大きな負荷なく使用することができる。

尚、習熟度に関する公式は、「多言語母語の日本語学習者横断コーパス (International Corpus of Japanese as a Second Language: I-JAS)」にある12言語を母語とする日本語学習者(初級, 中級, 上級の3レベル)373名のエッセイを基準データとして用い、重回帰分析によって得られたモデルに基づいている(Lee & Hasebe 2020)。

また、*jWriter* では、作文能力の評定値のほか、評定値を基に付与される作文評価・ガイドラインという5段階評価(入門, 初級, 中級, 上級, 超級)に加え、「語の使い方に関する診断的評価」として「語の多様性」「漢語力」「長文作成力」「難解語」に関するフィードバックが得られる。画面上には各項目の値域ごとに用意されたアドバイスのテキストと共に、グループ平均と実際の値とを示したグラフが表示される。この機能により、学習者は自らの作文能力における具体的な課題を見出すことができる。また、教師は学生のその時点での作文能力に関する詳しい情報を知ること、今後の指導につなげることができる。以上のような文章の習熟度の評価に加え、2022年から文章が持つ論理性の度合を計算モデルで推定し、可視化する論理性の自動評価も行うことができるようになった。この論理性の評価の算出公式やその作成過程については、本研究の分析対象外であるため省略する。

3. 先行研究

本章では *jWriter* の自動評価についての有用性に関する先行研究及び台湾人日本語学習者に対する *jWriter* の実証研究について主なものを紹介する。

小森ら(2022)では *jWriter* の評価と教師の評価(「包括的評価」と「内容」「構成」「言語」「誤用」の内容別評価)の関係性について比較した結果、*jWriter* の評価と最も強い相関関係があったのは、教師の「包括的評価」と「言語」であることが明らかになった。

また、*jWriter* では、習熟度の算出には使用されないものの、各品詞別の使用語数や接続詞の分類別の使用語数、自動詞、他動詞の数なども測定することができる。小森ら(2022)では、*jWriter* の言語統計量を使って、中級学習者と上級学習者の日本語作文における言語的側面の差を詳細に分析している。結果として、作文の習熟度推定の公式に含まれないものであっても、名詞、形容詞、副詞、助詞、代名詞などにおいて中級学習者よりも上級者学習者の方が多く出現していることが明らかになった。

一方、*jWriter* を使用して台湾人日本語学習者の作文の縦断的分析を行ったものに陳相州

(2022)がある。陳相州(2022)は台湾の大学1年生から4年生までの日本語学習者における日本語作文を対象に、*jWriter*の5つの診断的評価項目(長文作成力、語の多様性、漢語力、難解語、接続表現)の各評価点と、学習者の日本語レベル及び学習年について調査したものであり、調査項目である5つの診断的評価項目のうち「長文作成力」「語の多様性」「漢語力」「難解語」の4項目が、1年生から4年生の学習経過に沿って伸びが見られ、「接続表現」では、1年生から4年生の学習を経ても伸びが見られなかったとしている。この研究結果は、今後の台湾人日本語学習者に特化した作文の言語的側面での変化を見るうえで非常に有益なものである。

4. 研究課題

以上の先行研究より、*jWriter*による作文の自動評価が台湾人日本語学習者の作文における計量的な分析が言語的側面の評価や判定に対して一定程度有用であることは明らかであるが、小森ら(2022)や陳相州(2022)が示したように、*jWriter*習熟度の算出には使用されないものの、日本語学習者の作文を評価するうえで、有用な指標があるようだ。*jWriter*では作文の習熟度推定の公式に含まれないものでも、文の中で使用されている助詞(が/を/で/に/と/から/の/は/も)、品詞(名詞/動詞/形容詞/助詞/助動詞/副詞/接続詞/連体詞)、機能別接続詞(順接/逆説/並列、列挙/対比/例示/理由/譲歩/まとめ)、動詞の自他についての計量的分析をすることができるが、特に台湾人日本語学習者の作文における言語面をより包括的に評価するためには、台湾人に特化した縦断的データに基づいて、これらの項目の計量的変化を分析することが、学習者が産出する作文の変化をより客観的にとらえ、特に台湾人日本語学習者に対してより効果的且つ詳細なフィードバックを与えるうえでも、より多様な指標に基づく自動評価を可能にするためにも有益であると考えられる。そこで、本稿では、以下の2点を調査の課題として設定することとする。

課題①*jWriter*を用いて、筆者が担当する台湾における日本語学習者(以下JFL)2年次生の上学期の作文と下学期の作文を比較する。具体的にはJFL上学期と下学期の作文について以下の2点を分析する。

- 1)各時期の習熟度を比較し、変化が見られたのであれば、その原因を特定すべく、各項目について分析する。
- 2)習熟度の産出に寄与しない項目についても、上学期から下学期における計量的変化を分析する。

課題②*jWriter* を用いて、JFL 下学期の作文と日本人大学生の作文を比較し、筆者が担当する JFL2 年次の学生の日本語文章力の現状を把握すること。具体的には JFL 下学期と日本人大学生の作文について以下の 2 点を分析する。

- 1) 習熟度を比較し、違いが見られたのであれば、その原因を特定すべく、各項目について分析する。
- 2) 習熟度の産出に寄与しない項目についても、計量的差異を分析する。

5. 課題 ①

5.1 対象

筆者が担当する「大ニ日文習作」を受講している JFL2 年次生（初級～中級）29 名である。

分析対象としたものは、JFL が書いた作文（6 テーマ）であるが、今回の研究目的は 2 年次上学期¹と下学期の作文をそれぞれ分析し、文章の上達について考察することであるため、上学期の作文、3 テーマと下学期の作文 3 テーマに分けてそれぞれ比較することとする（Table 1 参照）。したがって、上学期、下学期共に分析対象となる作文数は 29 名×3 テーマの合計 87 である。また、いずれの作文も、学習者にテーマと字数を指定し、後日提出させたものをデータ化したものである。

Table 1 各作文テーマと提出期限

比較群	提出期限	作文テーマ	平均語数
上学期	2021/9/27	自己紹介	474.4
	2021/10/6	日記を書く	435.2
	2021/10/28	趣味の紹介	438.5
下学期	2022/5/13	日本語を学習する理由	455.2
	2022/5/30	インターネットの問題点	443.3
	2022/6/29	台湾の国民性	482.7

5.2 分析方法

それぞれの作文を *jWriter* で数値化し、それぞれの項目について、対応のある t 検定にて

¹ 台湾の大学は 1 年 2 学期制であり、9 月に新学期を迎え 6 月に終業を迎える。9 月～1 月を上学期とし、2 月～6 月を下学期とする。

2群（上学期、下学期）間の差を検定した。比較する項目は *jWriter* において、文章の習熟度に関する変数である以下の8項目、①評点（習熟度の計算モデルにより算出された点数）②文の長さ（1文中の平均形態素数）、③TTR（延べ語数の比率を表し、数値が大きいほど語彙数が豊富であることになる）、使用語彙の難易度を測る指標として④中級前半語の使用率⑤中級後半語の使用率の二つ（語彙の難易度は「日本語教育語彙表」に準拠）、⑥動詞の使用率⑦和語の使用率⑧漢語の使用率である。また習熟度の判定には寄与しないものの、*jWriter* で数値化できる以下の4項目についても比較した。①助詞別使用比率（が／を／で／に／と／から／の／は／も）、②品詞別使用比率（名詞／動詞／形容詞／助詞／助動詞／副詞／接続詞／連体詞）、③接続詞の使用比率（順接／逆説／並列、列挙／対比／例示／理由／譲歩／まとめ）、④動詞の自他の使用比率を比較する。

5.3 結果

Table 2 は上学期作文と下学期作文における各項目の平均の比較である。評点は *jWriter* の習熟度で算出されたもの、文の長さは1文における平均形態素数、TTR は異なり語数÷総語数で算出される数値である。%で示すものは、接続詞の下位分類と自他動詞以外は、総形態素数を分母として算出したものである。比較のために差（下学期の平均から上学期の平均を引いたもの）を示す。またその差に有意差が認められる場合には**($p<0.01$)で、有意傾向が認められる場合には *($p<0.05$)で示す。

Table 2 上学期作文と下学期作文の比較（各項目別の平均）

	上学期 (N=87) 平均(SD)	下学期 (N=87) 平均(SD)	差 (下学期－上学期)	
・評点	1.67(0.40)	2.31(0.42)	+0.64	($t(86)=6.58, p<0.001$)**
・文の長さ	16.2(3.53)	23.4(3.90)	+7.2	($t(86)=10.4, p<0.001$)**
・TTR (/総形態素数 %)	0.41(0.05)	0.38(0.05)	-0.03	($t(86)=3.04, p=0.004$)**
・中級前(%)	19.3(0.05)	24.5(0.04)	+5.2	($t(86)=5.38, p<0.001$)**
・中級後(%)	7.9(0.04)	14.4(0.04)	+6.5	($t(86)=7.77, p<0.001$)**
・動詞 (%)	5.1(1.53)	6.3(1.84)	+1.2	($t(86)=3.25, p=0.002$)**

・和語 (%)	66.3(3.24)	66.7(6.20)	+0.4	($t(86)=0.56, p=0.579$)
・漢語 (%)	14.1(2.93)	13.8(2.38)	-0.3	($t(86)=0.92, p=0.365$)

【助詞】 (/ 総形態素数 %)

・が (%)	2.3(0.88)	2.6(0.89)	+0.2	($t(86)=1.37, p=0.174$)
・を (%)	4.0(1.28)	3.5(1.19)	-0.5	($t(86)=2.04, p=0.043$)*
・で (%)	1.4(0.66)	1.5(0.66)	+0.1	($t(86)=1.94, p=0.058$)
・に (%)	2.8(0.91)	3.1(0.99)	+0.3	($t(86)=1.32, p=0.194$)
・と (%)	1.7(0.75)	2.1(0.85)	+0.4	($t(86)=2.29, p=0.023$)*
・から (%)	0.7(0.38)	0.2(0.33)	-0.5	($t(86)=5.31, p<0.001$)**
・の (%)	3.3(1.37)	2.9(0.99)	-0.4	($t(86)=1.73, p=0.087$)
・は (%)	3.6(0.99)	3.1(1.14)	-0.5	($t(86)=2.28, p=0.025$)*
・も (%)	1.0(0.53)	1.2(0.69)	+0.2	($t(86)=1.57, p=0.120$)

【品詞】 (/ 総形態素数 %)

・名詞 (%)	22.6(2.53)	19.5(2.44)	-3.1	($t(86)=6.19, p<0.001$)**
・動詞 (%)	5.1(1.53)	6.3(1.84)	+1.2	($t(86)=3.34, p<0.001$)**
・形容詞 (%)	3.9(1.11)	3.7(1.59)	-0.2	($t(86)=0.85, p=0.398$)
・助詞 (%)	26.6(2.27)	27.5(2.48)	+0.9	($t(86)=1.98, p=0.049$)*
・助動詞 (%)	12.6(2.11)	11.5(1.53)	-1.1	($t(86)=3.03, p=0.003$)**
・副詞 (%)	1.7(0.91)	2.1(0.85)	+0.4	($t(86)=2.30, p=0.023$)*
・接続詞 (%)	0.3(0.34)	0.5(0.31)	+0.2	($t(86)=3.70, p<0.001$)**
・連体詞 (%)	0.6(0.52)	0.9(0.58)	+0.3	($t(86)=3.02, p=0.003$)**

【接続詞】 (/ 接続詞の総数 %)

・順接 (%)	6.5(17.6)	24.7(21.0)	+18.2	($t(86)=4.77, p<0.001$)**
・逆接 (%)	22.4(32.4)	17.1(21.6)	-5.3	($t(86)=0.95, p=0.342$)
・並列 (%)	49.0(34.8)	30.5(24.0)	-18.5	($t(86)=3.07, p=0.002$)**
・対比 (%)	0.0(0.00)	1.3(5.60)	+1.3	($t(86)=1.76, p=0.082$)
・例示 (%)	12.2(20.9)	20.4(18.5)	+8.2	($t(86)=2.05, p=0.042$)*
・理由 (%)	0.5(3.34)	3.2(7.56)	+2.7	($t(86)=2.37, p=0.019$)*

・譲歩 (%)	0.0(0.00)	0.0(0.00)	0.0	
・まとめ(%)	0.4(2.67)	0.8(3.88)	+0.4	($t(86)=0.59, p=0.554$)

【自他動詞】 (/ 動詞の総数 %)

・自動詞(%)	34.7(18.2)	32.1(14.2)	-2.6	($t(86)=0.79, p=0.430$)
・他動詞(%)	65.3(18.2)	67.9(14.2)	+2.6	($t(86)=0.79, p=0.430$)

5.4 考察

上学期作文と下学期作文との比較では、習熟度の評点が有意に上がった。その要因として、主に1文中の形態素数が伸び、文章が長くなったこと、使用する語彙もより難易度の高い語が多く使用されるようになったことが挙げられる。一方で和語と漢語の使用率には有意な変化が認められなかった。

次に *jWriter* の評定値推定の公式には含まれていないものの、有意な差を示した項目についてみていく。

助詞の使用では、より単純な文型に使用される「は」「を」の使用が有意に減少し、その他の助詞の使用が増加した。これは、「は」がより早い段階で習得されるからである。Yagi (1992) では、英語、中国語、ベトナム語を母語とする中級初期の学習者の作文に現れる助詞の正用順序を調査し、「は」「を」の正用率が「が」を上回ることを報告している。自由作文では習得が遅れている用法は使用が回避されることが知られており、台湾の JFL においても、学習の早い段階では「は」を多用しているのが、その他の格助詞の習得が進むにつれ、置き換った可能性がある。学生 A、B の作文で比較してみても、前期の文章では、ほとんどの文章が、「AはB」の型で書かれている（【学生 A】全5文中4文、【学生 B】全12文中5文）のに対し、後期はその使用が減少している（【学生 A】全5文中1文、【学生 B】全6文中2文）。

【学生 A】前期（趣味の紹介）

私の趣味は色々あります。映画や読書や音楽などです。その中には一番好きなのは小説を読むことです。初めて読んで好きになった小説は東野圭吾の《ナミヤ雑貨店の奇跡》です。推理小説の世界は色々な人のストーリーが素敵に繋がっています。

【学生 A】後期（私が日本語を学習する理由）

高校生の時に初めて日本に旅行に行ったとき、私^は日本の文化や風景が大好きになりました。その時に日本語が全然わからなかったので、少し不便な経験がありました。例えば、道に迷ったとき、どのように助けを求めたらわかりません。複雑な新幹線にも同じ問題があるんです。したがって、日本のことをもっと理解するために、台湾へ帰ってから日本語をしっかりと勉強することにしました。

【学生 B】前期（日記を書く）

先月に父の日をお祝いしたために、家族とくらすしを食べて行った。私たち^はマグロすしやサケすしやエビすしなどを食べた。食事の後、皿を入れて、ガシャポンをすることができた。兄^はキーホルダーをもらった。私もマスキングテープをもらった。とても楽しかった。それから、家に帰って、簡単なパーティーをした。私たち^は一緒に私が作ったケーキを食べた。そのケーキの種類^はチーズだった。父がいちばん好きな種類だから、父^は嬉しそうだった。その後、一緒に写真を撮って、フェイスブックにあげた。この素晴らしい一日を残すことができた。

【学生 B】後期（私が日本語を学習する理由）

私が小学生の時いつも兄と一緒に日本のアニメを見ていたのは、どんどん日本語に興味があったからです。アニメ^は中国語のタイトルがありますが、タイトルが見なくてもキャラクターが何を言っているのか理解できるようになりたいです。高校に入った後、先輩の勧めで日本語についてのサークルに参加しましたが、この時から日本語を一生懸命に勉強することにしました。子供の頃の夢を叶えることができますが、他の言語を学ぶことで未来の仕事に役に立つこともできると思っています。さらに、将来^は夢と現実を合わせて、マスコミ関係の仕事をしたいと思います。例えば、アニメや映画やドラマを翻訳する仕事です。

「から」の使用に関して、その使用率が有意に減少した。接続助詞としての「から」の使用状況として、畠山（2012）は、接続助詞「から」の習得が類似の表現である「ので」の習得に先行する可能性について示唆している。「から」は体言または連体形に接続しているのか、終止形に接続しているかによって格助詞と接続助詞に分けられるが、特に「から」の使用実態を確かめるため、格助詞の「から」、接続助詞の「から」、接続助詞の「ので」について再度カウントし直してみた。対象は前期の最も早い時期に書かれた「自己紹介」の文章と後期最も遅く書かれた「台湾の国民性」の文章とし、それぞれのテーマに沿って

書かれた学生 29 名分の文章から格助詞の「から」、接続助詞の「から」、接続助詞の「ので」を抽出した (Table 3 参照)。

Table 3 各作文テーマと「から」「ので」の使用数【使用率 (使用数/文数)】

テーマ	格助詞「から」	接続助詞「から」	接続助詞「ので」	総文数
前期： 自己紹介	15 【0.04】	34 【0.10】	6 【0.02】	343
後期： 台湾の国民性	9 【0.05】	14 【0.08】	13 【0.07】	186

それぞれのテーマ別による使用数を比較してみると、明らかに接続助詞の「から」の使用が減少し、接続助詞「ので」の使用が増加していることがわかる。また、総文数が異なるため 1 文中での使用率を比較してみても、同様の傾向を見出すことができる (接続助詞「から」の使用率：0.10→0.08、接続助詞「ので」の使用率：0.02→0.07)。この結果より上学期から下学期にかけて「から」の使用が減少したのは、主に接続助詞の「から」が「ので」や他の表現に置き換っていたのだと言える。具体的に学生 C、D の作文で比較してみると、前期は「から」を使っていたところに、後期の文章では「ので」や「ため」の使用が見られるようになっている。

【学生 C】前期 (趣味の紹介)

私は〇〇出身の台湾人です。4 人家族です。今は、台北に住んでいます。私は内向的な性格だと思います。韓国の俳優が好きですから、韓国ドラマを見るのが好きです。推理も好きですから、名探偵コナンを見ることも好きです。

【学生 C】後期 (私が日本語を学習する理由)

私は毎年日本へ遊びに行くので、日本語が上手になればもっと便利だと思います。例えば、買い物に行くとき商品の紹介も理解できます。また、私は日本のアニメが好きです。早くアニメを見るために、私にとって日本語を勉強することがしなければなりません。そして、日本の観光客が多いので、もし日本人が問題がある場合、私は助けることができます。

【学生 D】前期（自己紹介）

私は〇〇といいます。日本のアニメやゲームが好きです**から**、日本語学科に入ったのです。いつか、自分のゲームを作りたいです。休みの時、ゲームをしたり、ゲームの実況を見たりします。ゲーム実況は、ゲームの面白いところを楽しみながら、日本語の勉強もできます**から**、よく見ます。

【学生 D】後期（私が日本語を学習する理由）

私が日本語を勉強する理由は、趣味の**ため**です。小さなときから、日本のアニメが好きです。しかし、全部のアニメが中国語に翻訳されるのは無理な**ので**、せめて見たい部分が自分で理解できるようになりたくて、日本語を勉強し始めました。

接続詞の使用に関しては、上学期から下学期にかけて有意に増加した。これは、文章を下学期では、文章をつなごうとする傾向が強くなったからであろう。詳細を見ると「並列、列挙」の使用が減り、「順接」、「例示」、「理由」の使用が増加している。董芸（2019）では、学習歴により「並列・継起」の接続詞の出現率が下がり続ける傾向が見られること、その要因として、①並列・継起の接続関係を、接続詞ではない別の表現を使用して表すようになること、②並列・継起の接続詞が省略されることの2点を指摘している。一方、学習歴が長くなるにつれて、接続詞のバリエーションが増え、並列・継起の接続詞の使い分けができるようになり、作文はより緻密な論理性を持つことになると述べている。文章を適切につなぐためには、文章の前後の論理関係を適切に把握する必要があるが、石黒（2000）が指摘するように、接続詞の「そして」は用法が幅広く、機能がわかりにくい。上学期では、そのような比較的論理関係があいまいな「並列・列挙」（例：「そして」、「また」など）の偏った使用があり、下学期にかけて、文章の前後関係に着目できるようになった結果、「並列」以外の接続詞の使用が増加したと考えられる。また、接続詞の使用の中で「例示」や「理由」の使用が増加したことから、より具体的に文章を展開しようとしていることがわかる。

品詞の使用傾向の変化では、名詞、形容詞、助動詞の使用が有意に減少したのに対して、動詞、助詞、副詞、接続詞、連体詞の使用が有意に増加した。これらの品詞の使用量データは *jWriter* の評定値推定の公式には含まれていないため、これらの変化が直接評定値の向上に寄与したということではない。しかしながら、動詞の増加に伴い、助詞や副詞の使用が増加し、文章が長くなった結果、評点が向上したと考えられる。動詞の使用の変化について、石川（2021）は、I-JAS「多言語母語の日本語学習者横断コーパス（International Corpus

of Japanese as a Second Language) における産出データおよび日本語母語話者の産出データから、習熟度と動詞に使用の関係を量的に調査した。その結果、延べ語数、異なり語数とも、初級から中級にかけて増加し、母語話者を上回る水準に達した後、再び減少する逆 U 字型の分布を見せると指摘している。今回見られた動詞の使用量の増加は、この初級から中級にかけての動詞使用量の増加と重なる。ただし、石川 (2021) では使用された動詞の質的な違いにも着目し、自動詞他動詞の誤用やコロケーションの誤りなど学習者にのみ見られる動詞の使用があることなどが指摘されているため、動詞使用量の増加に伴い、動詞の誤用も増加している可能性が高いことも併せて指摘しておく。

以上より、JFL2 年次前期の作文と後期の作文を計量的に比較分析した結果、評点が上昇したことから、文章の上達が明確になった。また、詳細に分析すると、使用される品詞や助詞の多様化や動詞の使用や接続詞、連体詞の使用の増加などが認められ、単純な文章から、複文などのより複雑な構造を持つ文章が多く使用されるようになったことが垣間見える。

6. 課題 ②

6.1 対象及び分析方法

対象及び分析：筆者が担当する「大二日文習作」を受講している JFL2 年次生 29 名が後期に作成した作文 (29 名×3 テーマ=87) と日本人大学生 63 名 (2018~2021) が 1 年次に作成した作文 (テーマ「今までで最も印象に残っている言葉」) を分析対象とする。分析項目は 4.2 で示したものと同様であるが、ここでは、対応のない t 検定にて 2 群 (JFL、日本人大学生) 間の差を検定した。

6.2 結果

Table 3 は JFL2 年次生下学期作文と日本人大学生の作文における各項目の平均の比較である。

Table 4 JFL の作文 (下学期) と日本人大学生の作文の比較
(各項目別の平均)

台湾人	日本人	差
(N=87)	(N=63)	
平均(SD)	平均(SD)	(日本人-台湾人)

・評点	2.31(0.43)	3.02(0.51)	+0.71	($t(148)=7.73, p<0.001$)**
・文の長さ	23.4(3.90)	26.6(7.77)	+3.2	($t(148)=2.53, p=0.013$)*
・TTR	0.39(0.05)	0.33(0.03)	-0.06	($t(148)=8.10, p<0.001$)**
	(/ 総形態素数 %)			
・中級前(%)	24.5(4.07)	21.4(4.45)	-3.1	($t(148)=3.75, p<0.001$)**
・中級後(%)	14.4(3.89)	15.0(4.01)	+0.6	($t(148)=0.76, p=0.448$)
・動詞 (%)	6.3(1.84)	7.4(1.57)	+1.1	($t(148)=3.48, p<0.001$)**
・和語 (%)	66.7(0.05)	74.1(0.05)	+7.4	($t(148)=7.29, p<0.001$)**
・漢語 (%)	13.8(6.20)	13.4(4.36)	-0.4	($t(148)=0.64, p=0.520$)

【助詞】	(/ 総形態素数 %)			
・が (%)	2.6(0.89)	2.5(0.79)	-0.1	($t(86)=0.50, p=0.616$)
・を (%)	3.5(1.19)	2.9(0.81)	-0.6	($t(148)=3.04, p=0.003$)**
・で (%)	1.5(0.66)	1.4(0.55)	-0.1	($t(148)=0.62, p=0.538$)
・に (%)	3.1(0.99)	3.7(0.87)	+0.6	($t(148)=3.42, p<0.001$)**
・と (%)	2.1(0.85)	2.3(0.79)	+0.2	($t(148)=1.66, p<0.001$)**
・から (%)	0.2(0.33)	0.6(0.31)	+0.4	($t(148)=6.63, p<0.001$)**
・の (%)	2.9(0.99)	3.3(0.98)	+0.4	($t(148)=2.40, p=0.018$)*
・は (%)	3.1(1.14)	3.1(0.72)	0.0	($t(148)=0.25, p=0.804$)
・も (%)	1.2(0.69)	1.6(0.73)	+0.4	($t(148)=4.16, p<0.001$)**

【品詞】	(/ 総形態素数 %)			
・名詞 (%)	19.5(2.44)	20.8(2.56)	+1.3	($t(148)=2.71, p=0.008$)**
・動詞 (%)	6.3(1.83)	7.4(1.57)	+1.1	($t(148)=3.48, p<0.001$)**
・形容詞(%)	3.7(1.59)	3.0(1.00)	-0.7	($t(148)=2.72, p=0.008$)**
・助詞 (%)	27.5(2.48)	29.8(2.39)	+2.3	($t(148)=4.91, p<0.001$)**
・助動詞(%)	11.5(1.53)	11.1(2.01)	-0.4	($t(148)=0.47, p=0.637$)
・副詞 (%)	2.1(0.85)	1.5(0.58)	-0.6	($t(148)=4.32, p<0.001$)**
・接続詞(%)	0.5(0.31)	0.5(0.34)	0.0	($t(148)=1.32, p=0.189$)
・連体詞(%)	0.9(0.58)	1.7(0.64)	+0.8	($t(148)=6.75, p<0.001$)**

【接続詞】 (/ 接続詞の総数 %)

・順接 (%)	24.7(21.0)	11.6(20.2)	-13.1	($t(148)=3.31, p=0.001$)**
・逆接 (%)	17.1(21.6)	39.4(29.1)	+22.3	($t(148)=4.43, p<0.001$)**
・並列 (%)	30.5(24.0)	33.4(25.9)	+2.9	($t(148)=0.59, p=0.554$)
・対比 (%)	1.3(5.60)	0.0(0.00)	-1.3	($t(148)=1.88, p=0.063$)
・例示 (%)	20.4(18.5)	5.5(15.9)	-14.9	($t(148)=4.49, p<0.001$)**
・理由 (%)	3.2(7.56)	1.8(6.43)	-1.4	($t(148)=0.99, p=0.325$)
・譲歩 (%)	0.0(0.00)	0.0(0.00)	0.0	
・まとめ(%)	0.8(3.88)	0.3(2.50)	-0.5	($t(148)=0.71, p=0.479$)

【自他動詞】 (/ 動詞の総数 %)

・自動詞(%)	32.1(14.2)	46.0(9.24)	+13.9	($t(148)=6.16, p<0.001$)**
・他動詞(%)	67.9(14.2)	54.0(9.24)	-13.9	($t(148)=6.16, p<0.001$)**

6.3 考察

JFL 下学期作文と日本人大学生の作文との比較では、日本人大学生の作文の評点が有意に高いという結果であった。その要因として、主に1文中の形態素数が多く、文章が長いこと、使用する語彙も日本人大学生の方がより難易度の高い語を多く使用していることが挙げられる。一方で和語と漢語の使用率では、むしろJFLの方がより漢語を好んで使用していた。

これは、台湾で用いられる文字が漢字であることと関係していると思われる。台湾では公用語として標準中国語が用いられている。ただし、中国大陸の共通語（いわゆる「普通話」）と異なる側面があるため、「國語」や「華語」をよばれることもあるが、使用される文字は日本語で使用される漢字と若干の違いはあるものの、基本的に日本語でも使用される漢字である。したがって、2年次上学期のような学習歴が1年程度であったとしても、多くの漢語を使用したものと思われる。石黒（2004）において、中国語を母語とする日本語学習者が書く文章の特徴の一つに豊富な漢語使用が挙げられると述べているように、すでに豊富な漢字の知識を習得しているため、非漢字圏の学習者に比べて、更に日本語母語

話者と比べても、多くの漢字の使用がみられることが分かっている（李 2001、胡 2012）。ただし、台湾國語と日本語における漢語は異なる意味や使用のものも多く²、母語である台湾國語の知識が却ってマイナスに作用することがある。小森ら（2012）の実証実験では、中国語を母語とする日本語学習者を対象に漢語の中でも、日本語と中国語とで同義とされている日中同型語 20 語に焦点を当てて調査を行い、中国語と同じ連語形式（保持传统／伝統を保持する）と、中国語と異なる連語形式（建设家庭／*家庭を建設する）、漢語でなく和語を用いた方が適切な連語形式（整理头发／*髪の毛を整理する／髪の毛を整える）について、どの程度正確に習得しているかを分析した。その結果、中国語と同じ共起語が使用できない語については、日本語能力試験 1 級以上の学生でも習得が進んでいないと報告している。今回 JFL に見られた漢語の過剰使用においても、上期、下期ともに JFL が母語転用している例が少なからずあると思われる。jWriter では誤用の分析は自動でなされないため、漢語の使用の質的变化については見ることができないが、学習の早い段階から高い頻度で使用される漢語の中には多くの誤用が含まれていると推測される。

助詞の使用では、日本人大学生の方が JFL（下学期）より「を」「は」の使用率が有意に低く、その分より多種多様な助詞（「に」「の」「も」）を使用していた。品詞の使用では、日本人大学生の方が有意に高い使用率を示したものに、名詞、動詞、助詞、連体詞がある。ここから推測できることは、日本人大学生の方が JFL よりもさらに複雑な文章構造で文章を生成しているのではないかということである。加えて、日本人学生の文章において、接続表現よりも、連体詞の使用割合が高くなっていたことから、日本人大学生の方が複文などで文章をつなぐ際に、接続詞に頼るのではなく、「この」「その」「あの」のような指示連体詞を含む連体詞を多く使用して文を繋いでいる可能性が考えられるかもしれない。

自他動詞の使用割合から見ると、JFL では上学期、下学期にかかわらず他動詞を多く使用する傾向がある一方で、日本人学生は、自他動詞ともバランスよく使用しているようである。先行研究においても、守屋（1994）、小林（1996）、Morita（2004）が自動詞の習得の方が他動詞の習得よりも困難であるという見解を示しており、その点から見ると JFL において、学習初期段階では他動詞の使用が多くみられ、その後自動詞が増えていくということが予想される。今回の分析結果を見ると、上学期、下学期ともに他動詞の使用が自動詞の使用よりも同程度多く、偏りがみられた。しかしながら、この偏りが、単に習熟度によるものなのかどうかについては容易に断定することはできない。李（2020）は、YNU 書

² 中国語と日本語の漢字の異同については、小室（2019）に詳しい。

き言葉コーパスの分析を通して、自動詞他動詞の使用傾向を、母語別、習熟度、タスク別の3要因分散分析で分析したところ、習熟度とタスク別において自他動詞の使用に有意な差が見られたとしている。特にタスクの違い（発信が自発的であったか否か、読み手の属性や文章の長さの違い）によって、自他の使用頻度に差が見られたため、今回の分析においても、テーマの違いが使用頻度の差として現れた可能性は否定できない。テーマによって自他の使用頻度がどのように変化するか、また中級から上級にかけて自動詞の使用率がどのように推移するかについては今後の研究課題である。

7. まとめ

今回研究課題として以下の二点について、*jWriter* を用いて、筆者が担当する JFL の日本語作文を対象に計量的分析をおこなった。

課題①*jWriter* を用いて、筆者が担当する台湾における日本語学習者（以下 JFL）2 年次生の上学期の作文と下学期の作文を比較し、いくつかの項目についてその計量的変化を分析すること。

課題②*jWriter* を用いて、JFL 下学期の作文と日本人大学生の作文を比較し、いくつかの項目についてその計量的変化を分析し、筆者が担当する JFL2 年次の学生の日本語文章力の現状を把握すること。

その結果、課題①では、1 文の形態素数の増加、使用語彙の難易度の上昇により、評点が有意に上昇した。また単に文章が長くなっただけではなく、使用される品詞や助詞の多様化や動詞の使用や接続詞、連体詞の使用の増加などから、単純な文章から、複文などのより複雑な構造を持つ文章が多く使用されるようになったことが示唆された。ただし、評点にかかわる項目の中で、漢語と和語の使用率に関しては、JFL は上学期から下学期にかけてほとんど変化が見られず、早い段階から漢語の高い使用率がみられたのが特徴であった。

課題②では、JFL の文章の長さや使用語彙の難易度は日本人大学生にまだ及ばなかったこと、助詞の多様な使用、動詞の使用頻度の向上という点において、JFL の文章にさらなる改善が必要であることが示唆された。また、文章の一貫性、結束性を保証するためのストラテジーに JFL と日本人学生とで違いがある可能性があり、日本人学生は、文章をつなぐ際に接続詞に頼るよりも、連体詞なども多く使用していることが示唆された。今後は中

級程度の JFL に対して文のつなぎ方を指導する際、連体詞の積極な使用も促す必要があるかもしれない。

ここまで、JFL の計量的分析を通して、文章の上達を数値化することで、学生の文章の変化をより具体的に見ることができた。このような作文の数値化は、より包括的で確実性の高い作文評価の基礎データとして活用できるばかりでなく、学生へフィードバックをする際にも非常に有用であると思われる。具体的には、ルーブリック評価のように学習基準を“見える化”し、学習者自身に評価をさせるなどすれば、学習の達成状況を把握させることもできるかもしれない。また、指導の面においても、例えば、今回の分析を通じて観察することができたものでいうと、自動詞の使用量が不十分な学生に対して、自動詞の使用を促すような指導を実施したり、文の長さが十分でない学生に対しては、接続詞の使い方や連体修飾節の作り方の演習をさせるなどして、学生それぞれの特徴に合わせたよりきめ細やかな指導プログラムを立案することなども可能になると思われる。

ただし、台湾人 JFL に対して以上のようなより効果的な作文評価や指導を行うためには、今回の分析では量的にも範囲的にも十分であるとはいえない。今後より多くの作文分析をより網羅的に行い、その上達過程をより明らかにしていく必要があるだろう。

また、今回の分析において「内容面」「構成面」「誤用」には触れなかった。*jWriter* の評価は作文の量的側面というよりも質的、内容的側面ということができるこれらの項目についてその評価に反映されにくいという側面がある³。*jWriter* の評価と教師の評価の相関関係を検討した小森ほか (2018) では、日本語学 習者 30 名が書いた意見文に対して、教師 3 名による包括的評価 (7 段階評価) および項目 別評価 (内容、構成、言語、誤用、各 5 段階評価) と、*jWriter* による評価 (5 段階評価) を比較した結果、*jWriter* の評価は、教師による包括的評価と項目別評価の「言語」との間で、正の強い相関が示される一方で、「内容」、「構成」、「誤用」などの評価項目とは関係がなかったことが示された。これらの項目を自動評価システムによる作文評価に反映させることは、今後の課題であろう。特に誤用を如何に学習者の評価に落とし込むかということには十分に慎重でなければならない。誤用の判定は、評価者のさまざまなバックグラウンドや傾向によって左右されるものであるし、誤用自体も文章が長くなったり、或いは学習者が新しく学習した知識を挑戦的にアウトプ

³ *jWriter* には文章が持つ論理性の度合を計算モデルで推定し、可視化する機能がある。具体的には、文章の論理性を反映する接続詞などを抽出し、計算モデルに基づいて論理性を 4 段階 (優・高・中・低) で自動評価するものであるが、意見文以外の様々な文章のテーマについて、その内容を評価するものではない。

ットしようとする際に多く見られるようになるため、誤用が増えることが必ずしも学習者のレベルの低下を示すものではないと思われるからだ。ただし「誤用」「内容面」「構成面」などの作文の質的、内容的側面についてもそれを評価し、指導に生かす必要があることは言うまでもないことである。したがって、それらの側面に関しては、現時点では、自動評価ではなく、実際の教師による評価を行うなど効率的な役割分担についても考えていく必要があるだろう。

参考文献

- Morita, M. (2004) The Acquisition of Japanese Intransitive and Transitive Paired Verbs by English-Speaking Learners : Case Study at the Australian National University 『日本語教育論集 世界の日本語教育』第 14 号. 国際交流基金日本語国際センター、pp167-192
- Yagi, K. (1992) : The accuracy order of Japanese particles. *Japanese Language Education around the Globe*, 2, pp15-25.
- 石川慎一郎 (2021) 「絵描写作文課題における L2 日本語学習者の動詞使用と習熟度の関係」 『統計数理研究所共同研究レポート』 No.444、pp1-22
- 石黒圭 (2000) 「「そして」を初級で導入すべきか」 『言語文化』 37、一橋大学語学研究室、pp27-38
- 石黒圭 (2004) 「中国語母語話者の作文に見られる漢語副詞の使い方の特徴」 『一橋大学留学生センター紀要』 第 7 号、pp3-13
- 胡曉睿 (2012) 「日本語学習者の作文における漢語の使用について」 『第九回国際日本語教育・日本研究シンポジウム予稿集』 香港日本語教育研究会
- 小林典子 (1996) 「相対自動詞による結果・状態の表現-日本語学習者の習得状況-」 『文芸言語研究言語編』 29 号、筑波大学文藝・言語学系、pp41-56
- 小室リー郁子 (2019) 『中国語母語話者のための漢字語彙研究-母語知識を生かした教育をめざして-』 くろしお出版
- 小森和子・李在鎬・長谷部陽一郎・鈴木泰山・伊集院郁子・柳澤絵美 (2018) 「教師による評価とコンピュータによる自動評価はどの程度一致するのか—中上級日本語学習者の意見文の評価を対象に—」 『2018 年度日本語教育学会秋季大会予稿集』、pp 278-283.
- 小森和子, 三國純子, 徐一平, 近藤安月子 (2012) 「中国語を第一言語とする日本語学習者の漢語連語と和語連語の習得-中国語と同じ共起語を用いる場合と用いない場合の比較-」 『小出記念日本語教育研究会論文集』 20、小出記念日本語教育研究会、pp49-61

小森和子, 伊集院郁子, 李在鎬 (2022) 「日本語学習者の作文における自動評価と教師評価の比較」『明治大学国際日本研究』14-1、 pp 41-68

陳相州 (2022) 「作文自動評価システムによる台湾人日本語学習者の作文縦断分析」『東呉外語學報』53、pp71-93

董芸 (2020) 「日本語学習者の作文における並列・継起の接続詞の使用実態-中国語母語話者の縦断コーパスの分析を通じて-」『国立国語研究所論集』19号、pp127-138

畠山衛 (2012) 「日本語学習者による原因・理由を表す接続助詞「から」「ので」の語用論的使い分け能力の習得を探究する横断的研究」ICU 日本語教育研究 8、国際基督教大学日本語教育研究センター、pp 3 - 17

守屋美千代 (1994) 「日本語の自動詞・他動詞の選択条件-習得状況の分析を参考に-」『講座日本語教育』第29分冊、早稲田大学日本語研究教育センター、pp151-165

李在鎬 (2021) 「書くことを支援する自動評価システム『jWriter』(特集AI やICT が変える言語教育)」『日本語学 (2021 冬号)』40(4), 42-51.

Lee, Jae-Ho, & Hasebe, Yoichiro (2020) Quantitative Analysis of JFL Learners' Writing Abilities and the Development of a Computational System to Estimate Writing Proficiency. *Learner Corpus Studies in Asia and the World*, 5, 105-120.

(<http://www.lib.kobe-u.ac.jp/repository/81012493.pdf>)

李在鎬・長谷部陽一郎・村田裕美子 (2019) 「学習者作文の習熟度に関する自動判定とWebシステムの開発について」, 李在鎬 (編) 『ICT×日本語教育』38-53. 東京:くろしお出版.

李在鎬 (2020) 「学習者作文と自他の使用頻度」『自動詞と他動詞の教え方を考える』くろしお出版、pp43-56

李漢燮 (2001) 「韓国人学生の日本語作文における語種選択について: 日本語教育のためのアジアの対訳作文データの収集とコーパスの構築」平成 11-12 年度科学研究費補助金基盤研究 (B) (2) 課題番号 1691041 研究成果報告書、研究代表者: 前田 (宇佐美) 洋、pp45-50

本研究では、学習者作文評価システム jWriter (<https://jreadability.net/jwriter/>) を利用した。

(おおたえいじ、輔仁大學日本語文學系)

A Quantitative analysis of compositions written by learners of Japanese as a second language in Taiwan

-Use of *jWriter*-

Eiji Ohta

Department of Japanese Language and Culture of Fu Jen Catholic University/Assistant professor

Abstract

The purpose of this study is to conduct a quantitative analysis of Japanese compositions written by Japanese learners in Taiwan using a automatic evaluation system "jWriter".The following two issues were analyzed.1) To analyze the compositions of the second-year Taiwanese students in the first semester and the second semester and to conduct a quantitative analysis and "visualization" of the changes in their writing. 2) To compare the compositions of Taiwanese second semester Japanese language students with those of Japanese university students, and to grasp the current status of the Japanese writing ability of the Taiwanese students. The results of Analysis 1) The grade on compositions of student significantly increased, due to the increased number of morphemes in a sentence and the increased difficulty of the vocabulary used. The results of analysis 2) suggested that Japanese learners in Taiwan need to further improve the variety of word types and particle words used and to use more verbs in their writing.

Keywords: Learning Japanese as a Second Language in Taiwan,

Japanese compositions, Quantitative analysis,

Automated essay scoring "jWriter"