

第二言語としての日本語ライティング評価ワークショップ

——評価基準の検討——

田 中 真 理*
長 阪 朱 美**
成 田 高 宏***
菅 井 英 明****

キーワード：第二言語としての日本語，ライティング評価，ワークショップ，評価基準，マルチプル・トレイト評価表

要 旨

本稿では、第二言語としての日本語におけるアカデミック・ライティングの評価基準作成の一環として行われたライティング評価ワークショップについて報告する。日本語教育においても近年ライティングの重要性は認識されてきたが、その評価方法に関しては、日本語教育学界で共有できるライティング評価基準がないこともあり、個々の日本語教師の判断に任されてしまうことが多い。しかし、大規模試験に日本語教師が直接関わることはなくとも、機関やクラスのライティング評価において信頼性、公平性を保つためには評価に関する共通の認識が必要である。そこで、筆者らは広く活用できる評価の基本的枠組み（ガイドライン）を提案することを目的に、独自の評価基準を作成した。本稿では、以上の観点から作成された評価基準を用いたワークショップの経緯と実際の状況を報告するとともに、提示した評価基準が有効に機能したか、また、採点者間ミーティングやアンケートから日本語教師がライティング評価をどのように捉えているかを考察、検討する。

ワークショップで使用したのは、「目的・内容」「構成・結束性」「読み手」「日本語A：正確さ」「日本語B：適切さ」の5トレイトを0～6のレベルで評価するマルチプル・トレイト評価表である。参加した20名の日本語教師による2種の小論文の評価結果は全体的に信頼性の高いものであること、一致しやすい小論文やトレイトがある反面そうでないものがあること、また単発的な個人レベルでの評価のずれは誰にでも起こりうることが示唆された。さらに、提示した評価基準が総体的には有効に機能した可能性が示された。

アンケートからは、採点者間ミーティングは高く評価されたものの、その手順や方法、評価基準のレベル記述に改善点のあることが分かった。今後、日本語教育学界でライティング評価の基本的枠組みを共有していくためにも、公のワークショップの機会の設けられることが望まれる。

*TANAKA Mari：名古屋外国語大学教授

**NAGASAKA Akemi：恵泉女学園大学教授

***NARITA Takahiro：元サハリン国立大学講師

****SUGAI Hideaki：香港理工大学副教授

1. はじめに

1-1. 目的

日本語教育におけるライティングの重要性は、記述問題を試験科目として課す大規模試験（日本留学試験）が近年行われていることもあり、教育の現場でも認識されてきている。しかし、実際に学習者が書いた作文をどのように客観的に評価するかについては、評価基準や評価手順についての共通認識がないため、教師一人一人の判断に任されてしまうことが多い。そこで、筆者らは平成17年度国立国語研究所日本語教育短期研修においてライティング評価ワークショップを実施した。ワークショップでは初級課程修了段階から大学の専門教育への移行段階までの学習者の作文(小論文)に対して広く活用できる「評価基準」とその使用方法を紹介し、さらに、グループでのディスカッション（採点者間ミーティング）を試み、参加者が実際に評価を体験することによって、各自の現場におけるライティング評価について考えるという機会を提供することを目的とした。

筆者らが上記「評価基準」を作成したのは、(1)現在、日本語教育学界で共有できるライティング評価基準がなく、(2)大規模試験の評価に日本語教師一般が直接関係することは少ないとしても、個々の機関のプレースメントテストや修了テストの信頼性、また、コースやクラスにおけるライティング評価の公平性を保つためには評価基準が必要であり、それらにも活用できるようなライティング評価の基本的枠組みを提案したいと考えたからである。

本稿は、以上のような観点より作成された評価基準を用いたワークショップについて報告するとともに、提示した評価基準がワークショップで有効に機能したか、また日本語教師がライティング評価をどのように捉えているかを考察、検討することを目的とする。

1-2. ライティング評価ワークショップ・評価トレーニング

前述したように、日本留学試験が記述問題を課している。その採点に関しては、日本学生支援機構が採点者を募集し、採点方法等についての講習会を行っているが（日本学生支援機構）、その具体的な内容に関しては試験の性格上非公開である。一方、同じパフォーマンス評価である口頭能力評価については、ワークショップや評価トレーニング等の報告が既になされている。機関内のワークショップとしては西川他（2003）が、短時間の評価トレーニングの試みとしては鈴木（2005）がある。さらに、庄司他（2005）では非母語話者の評価者を対象とする口頭能力の評定練習について報告している。しかし、ライティング評価ワークショップについての報告は、これまでのところ行われていない。

英語教育では、Test of English as a Foreign Language (TOEFL)、International English Language Testing System (IELTS) 等の大規模試験においてパフォーマンス評価が使われている。信頼性の

高い評価を行うためには評価者トレーニングが欠かせず (Hamp-Lyons, 1991, 2007; Weigle, 2002), TOEFL や IELTS や Cambridge ESOL Exams 等ではトレーニングを受けた評価者による評価が行われ, その報告等 (e.g., Show & Weir, 2007; Taylor & Falvery, 2007) も行われている。

Weigle (1994, 1998) は, トレーニングの重要性は評価者が評価表のスコアやクライテリアについて共通理解を得ることができ, それにより評価者の個人的な経験に基づくバイアスを防ぐことができる点や, 評価者間の一致が最優先ではなく他の評価者との一致の必要性に気付くことができる点にあるとしている。トレーニングにより評価者の自己一貫性 (self-consistency), つまり評価者内信頼性 (intra-rater reliability) を高めることはできる。しかし, 評価の厳しさの偏り等の評価者による個人差を取り除くことは難しいと言われている (Kondo-Brown, 2002; Weigle, 1994, 1998)。また, その一方で, 評価者トレーニングで評価の一致を強調することにより, 評価者の専門的知識 (expertise) や経験が軽んじられたり, 評価が評価表のクライテリアの表層的な面に集中してしまう恐れがある等の指摘もある (cf. Weigle, 1994)。

さらに, 近年ではインターネットによるオンライン評価者トレーニングも行われるようになった。時間的・空間的な点からもそのメリットは大きい。Elder, et al. (2007) は, オンライントレーニングは対面式 (face-to-face) のトレーニングと同様の結果が得られたが, 評価プロセスに慣れていない評価者にはオンラインよりも対面式を先に行うべきだと述べている。また, 評価者個人が持っているバイアスを減らすには, 対面式で各評価者にそれぞれフィードバックをするのが最も効果的である (Knoch, et al., 2007) と報告されている。

1-3. ライティング評価の基本概念

パフォーマンス評価としてのライティング評価は, 実際に大学や職場で必要であろう言語使用能力を予測し, 直接に学習者が産出したものを観察するものなので, 多肢選択等の間接テストに比べると妥当性, 真正性が高いと言える。しかし, その一方で, 学習者のアウトプット形式に自由が認められるので, (1)何をどう書かせるかをコントロールし, (2)どのように評価するかを決めることが重要となる (cf. 菅井, 2003)。 (1)がプロンプト (prompt) の役割で, (2)が評価のガイドライン, とりわけ評価基準 (評価表) の役割だと言えよう。White (1984) は, さらに, 高い信頼性を維持するのに必要なものの1つとして, 基準 (benchmark) となるサンプル・スクリプトが必要だと述べている。

今回のワークショップでは, 評価基準としてマルチプル・トレイト評価表を使用した。マルチプル・トレイト評価表では, あるいくつかのトレイト (特性) に注目し, そのトレイトを総合的に評価する。トレイトは機関やコースやクラスの目的によって決定すればよいとされている (cf. Hamp-Lyons, 1991)。ここでは日本語教育に必要な基本的なトレイトとして, 「目的・内容」「構成・結束性」「読み手」「日本語 A: 正確さ」「日本語 B: 適切さ」の5つを設定したが (評価基

準 A：トレイト別基準：資料 1)，これらのトレイトは目的に応じて追加したり減らしたり自由に調整できる。また、日本語教育でよく使われている分析的評価表では、「内容：30%」「構成：20%」「文法：25%」のように重み付けをし、各クライテリアの総和を算出するのが一般的であるが、マルチプル・トレイト評価表では各トレイトが独立しているので重み付けもなく、総和を算出することも前提としていない（詳しくは、田中，2005；田中・長阪，2004，2006 を参照）。本評価基準では、初級修了段階から大学専門課程へ移行するまでのアカデミック・ライティング練習段階を想定し、最高レベルを「6」として「6」～「1」のレベルを設定し、トレイト別に各レベルの具体的な記述を試みた。さらに、大規模試験への応用も考慮して「0」の定義も行った（評価基準 B：レベル別基準：資料 2）。

トレイトの中身を具体的に記述したものを基準説明（descriptor）と呼ぶが、これがないと、評価者は恣意的にトレイトの中身を解釈してしまうことになる。例えば、評価基準 A（資料 1）の「目的・内容」のサブトレイト「内容」を見ると、「トピック」「主張」「サポート」と下位項目に分かれ、その右側に実際の小論文で何を見るのかが具体的に記されている。さらに、どのような順序で見えていくのかが表の右端に「補足説明」として記されている。

2. ライティング評価ワークショップ開催までの経緯

ワークショップに必要なライティング評価ガイドラインは、予備調査（田中・長阪，2006）を経て作成された。ワークショップの方法は主に Weigle（2002）を参考にしている。

2-1. ワークショップ開催までの準備

- (1) 評価表：予備調査を経て、最終的には評価基準 A（トレイト別基準）と評価基準 B（レベル別基準）の 2 本立てとした。
- (2) プロンプト：授業、予備調査で使用したプロンプトの中から、アカデミック・ライティングの代表的なディスコース・モードである「説明」（exposition/explanation）と「論証」（argumentation）をそれぞれ主要モードとする「ファーストフードとスローフード」（以下：F & S）と「通学授業と遠隔授業」（以下：E-learning）の 2 つを選び、図表等の資料を使用しない「小論文形式」（600 字程度）とした（資料 3）。
- (3) 小論文：2 年かけて「F & S」83 編、「E-learning」73 編を収集した（資料 4）。書き手は日本の国立大学（理系）に所属する学生で、年齢は大半が 20 代であり、母語は中国語、韓国語、インドネシア語・マレー語、タイ語等である。日本語能力レベルは初級後期から上級（日本語能力試験 4 級程度～1 級合格）までの広範囲に及ぶ。参考のために日本人大学生の小論文（7 編）も含めた。評価に当たっては、書き手の氏名を消し番号で処理した。

- (4) **研究者による評価**：上記(3)で収集した小論文を筆者ら研究者4名が、上記(1)の評価基準Aと評価基準Bを使って個別に評価した。
- (5) **元データ**：上記(4)から日本人学生の小論文を除いたものを「元データ」(F&S：76編，E-learning：66編)とし、4名の評価の内的一貫性を示す信頼性係数(クロンバックの α 係数)を調べた。両方のプロンプトにおける全てのトレイトで0.8以上、特に「日本語A：正確さ」「日本語B：適切さ」では0.9以上の高い α 係数が得られた。
- (6) **サンプル小論文**：「元データ」から各プロンプト3レベルのサンプル小論文を選んだ。これらのサンプルには4名の評価の一致度の高いものを選んだが、さらに評価者間ミーティングを行い、点数を調整してワークショップ時に示す**参考評価点**を決定した。また、サンプル解説を作成した(資料5)。ワークショップ用のサンプル小論文はタイプした。
- (7) **練習用小論文**：ワークショップにおける練習では、参加者に評価基準に慣れてもらうために、評価の安定した小論文、各プロンプト3編をレベル順に評価してもらった(cf. Weigle, 2002)。また、手書き(文字の読みにくさ等)の影響を排除し、評価基準を使つての評価に集中できるように、小論文はタイプして使用した。
- (8) **本番用小論文**：講習後に行う本番用小論文には、サンプル小論文や練習用小論文よりも評価の難しいものを、各プロンプト5編選んだ。これらは、実際のライティング評価を鑑み、手書きのままとしたが、できるだけ文字の読みやすいものを選んだ¹。
- (9) **プロンプトのガイドライン**：想定される「読み手」や規定の600字を超えた場合の扱いやプロンプトの指示文の解釈を一定にするための簡単なガイドラインを作成した(資料3)。

2-2. ワークショップ

- (1) **参加者**：ライティング教育の経験がある日本語教師で、定員を20名とした。全員日本語母語話者で、内訳は、①性別：男性3名、女性17名、②年齢：20代～50代(30代、40代それぞれ7名)③日本語教育歴(回答のあった18名)：平均9.34年(標準偏差：4.25年)、最大：19.5年、最小：3年、④作文評価経験：全員あり、⑤勤務先(複数回答)：日本語学校：12名、大学：5名、その他4名で、4名が同時に大学院にも所属している。
- (2) **配布資料**：①研修の目的、②プログラム(時間配分)、③総合的評価表(TOEFL WRITING Scoring Guide)・分析的評価表(ESL Composition Profile)、④評価基準A及び評価基準B、⑤プロンプトとガイドライン、⑥サンプル小論文・3レベル 参考評価点と解説(F&S：1F～3F, E-learning：1E～3E)、⑦練習用小論文(タイプ書き)(4F～6F, 4E～6E)、⑧本番用小論文(手書き)(7F～11F, 7E～11E)、⑨フィードバック票、⑩アンケート調査票(一部を除い

¹ サンプル、練習用、本番用の小論文番号(例えば7E, 7F)は、必ずしも同一の書き手ではない。2つのプロンプトで対応するかのように同じ番号が用いられているが、これらは便宜的な理由からである。

て冊子体で配布：資料1～資料5)

- (3) プログラムと当日の流れ：短期研修の枠内という時間制約があり、かなりの過密スケジュールになっている。グループは5名から成る4グループとし、できるだけ背景や所属機関の異なる教師で構成した。

【午前の部】

【午後の部】

10:00-10:10	研修の趣旨とスケジュールの説明	13:30-14:00	グループによる報告と全体での質問と確認
10:10-10:40	評価基準 A とプロンプトの説明		
10:40-11:20	評価基準 B とサンプルの説明	14:00-15:00	参加者による本番用小論文評価
11:20-11:25	(休憩)	15:00-15:15	(休憩)
11:25-12:00	参加者による練習用小論文評価	15:15-15:45	全体ディスカッション
12:00-12:30	グループ・ディスカッション	15:45-15:55	アンケート記入
12:30-13:30	(昼食)	16:00	終了

プログラムに沿って、まず(a)評価に関する一般的説明(総合的評価表、分析的評価表と今回使用するマルチプル・トレイト評価表の特徴)、(b)評価基準のトレイトの説明(特に、これまで日本語教育にはなかった「読み手」のトレイトを設定した理由や具体的な内容)²、(c)評価表の見方(各トレイトの基準説明をどのような順序で見っていくか等)、(d)プロンプトの解説、(e)サンプル小論文のレベル別解説を行った。その後、各自で練習用小論文を評価した後、各グループでディスカッションを行った。午後、各グループでの評価点と参考評価点を開示し、グループと主催者(筆者ら)間で質疑応答を行った。その後、参加者は本番用小論文の評価を個別に行い、アンケートを記入して終了した。

3. ワークショップでの評価結果とその分析

3-1. ワークショップで使用した小論文

ワークショップで使用したサンプル・練習用・本番用小論文の詳細について述べる。表1のサンプル小論文は、大きく3レベルに分け、レベル順に提示し、トレイト別の参考評価点を示した。また、なぜその評価点になるのか解説を付けた(資料5)。サンプルには3つのレベルにおいて典型的だと思われるものを選んだので各小論文の5つのトレイトにはほぼ同じような点数が並んでいるが、実際の評価においては「日本語」が「5」でも「構成・結束性」が「2」というこ

²「読み手」のトレイトの設定理由の1つは、書き手(学習者)に(教師ではない)真の読み手を想定して目的に合った文章が書けるようにすることである。また、トレイトの内容は、「読み手への配慮」「読み手にとっての面白さ」であるが、後者は評価者によるばらつきが出ることを考え、選抜試験等のハイクラス・テストでは、このトレイトをはずすことを提案している(田中・長阪, 2004; 田中他, 2007)。

表1 サンプル小論文の詳細 (F & S, E-learning)

F & S	書き手の 国籍	目的・ 内容	構成・ 結束性	読み手	日本語 A: 正確さ	日本語 B: 適切さ	E-learning						
							書き手の 国籍	目的・ 内容	構成・ 結束性	読み手	日本語 A: 正確さ	日本語 B: 適切さ	
1F	ハイチ	2	2	3	2	2	1E	ラオス	2	2	3	2	2
2F	タンザニア	3	3	4	4	4	2E	中国	4	3	4	4	4
3F	インドネシア	5	5	5	6	6	3E	中国	5	4	5	5	5

とはあり得る。また、ここに示されている点数はあくまでも筆者ら4名で検討した参考評価点で絶対的なものではなく、前後1点(±1)の差は十分あり得ることを説明した。

表2の練習用小論文も、サンプル小論文と同様の観点から選び、参加者に個別に評価してもらった。各プロンプトの小論文は、レベル順になっている。

表2 練習用小論文の詳細 (F & S, E-learning)

F & S	書き手の 国籍	目的・ 内容	構成・ 結束性	読み手	日本語 A: 正確さ	日本語 B: 適切さ	E-learning						
							書き手の 国籍	目的・ 内容	構成・ 結束性	読み手	日本語 A: 正確さ	日本語 B: 適切さ	
4F	タイ	3	2	3	2	2	4E	ハイチ	1	2	1	1	2
5F	中国	4	4	5	5	4	5E	中国	4	4	5	4	4
6F	中国	5	5	5	6	6	6E	中国	6	6	6	5	6

表3は本番用小論文の詳細である。サンプル小論文や練習用小論文にはプロンプトの解釈が割れそうな小論文は含まれていなかったが、本番用小論文には、評価の割れそうなもの、例えば、**E-learning**の「あなたなら、どちらの方法を選びますか」というプロンプトの指示に対して「併用案」を書いている小論文の評価も含まれている。今回のプロンプトのガイドラインでは、通学授業と遠隔授業の併用案でも、理由が明確に述べられ説得力があると判断されれば大きく評価を下げることはしないとしたが、クラスで練習するような場合には「目的」(課題の達成)を重視し、中身がどんなに素晴らしくても課題が達成されていなければ、「1」や「2」を付けるという姿勢をとることは、むしろ、各トレイトが独立しているマルチプル・トレイト評価表の本来の意図に合っているだろう。

表3 本番用小論文の詳細 (F & S, E-learning)

F & S	書き手の 国籍	目的・ 内容	構成・ 結束性	読み手	日本語 A: 正確さ	日本語 B: 適切さ	E-learning						
							書き手の 国籍	目的・ 内容	構成・ 結束性	読み手	日本語 A: 正確さ	日本語 B: 適切さ	
7F	中国	3	3	4	3	3	7E	インドネシア	2	2	2	2	2
8F	中国	3~4	3~4	4	3~4	4	8E	中国	4	4	4	4	4
9F	中国	5	4	5	4~5	5	9E	韓国	5	6	6	6	6
10F	モンゴル	4~5	4	4	5~6	6	10E	モンゴル	3~4	3~4	4	5	5
11F	マレーシア (中国語)	6	5	6	6	6	11E	マレーシア (中国語)	6	6	6	6	6

3-2. 練習用・本番用小論文の評価結果と分析

表4はワークショップ参加日本語教師20名による練習用小論文の評価の平均値と標準偏差(*SD*)である。*SD*を見ると、4F~6F、4E~6Eの全てのトレイトにおいて1.0以内に収まっており、提示した評価基準が理解され、安定した評価が行われていたことが分かる。

表4 ワークショップ参加日本語教師20名による練習用小論文の評価の平均値と標準偏差(*SD*)

	目的・内容		構成・結束性		読み手		日本語A：正確さ		日本語B：適切さ	
	平均値	(<i>SD</i>)	平均値	(<i>SD</i>)	平均値	(<i>SD</i>)	平均値	(<i>SD</i>)	平均値	(<i>SD</i>)
4F	2.05	(0.51)	2.21	(0.41)	2.42	(0.49)	2.16	(0.49)	1.84	(0.36)
5F	3.79	(1.00)	3.87	(0.86)	4.11	(0.85)	3.89	(0.55)	4.11	(0.72)
6F	4.74	(0.91)	5.21	(0.61)	5.39	(0.80)	5.16	(0.59)	5.42	(0.67)
4E	2.00	(0.56)	2.05	(0.39)	2.32	(0.65)	1.95	(0.39)	1.95	(0.51)
5E	3.71	(0.63)	3.84	(0.59)	3.84	(0.67)	3.53	(0.75)	3.89	(0.64)
6E	5.18	(0.78)	5.45	(0.54)	5.26	(0.91)	5.11	(0.45)	5.11	(0.64)

表5は本番用小論文の平均値と標準偏差(*SD*)である。評価の割れそうなものを意図的に含めたために表4の練習用小論文より*SD*は大きい、全体的には大ききなばらつきはない。

表5 ワークショップ参加日本語教師20名による本番用小論文の評価の平均値と標準偏差(*SD*)

	目的・内容		構成・結束性		読み手		日本語A：正確さ		日本語B：適切さ	
	平均値	(<i>SD</i>)	平均値	(<i>SD</i>)	平均値	(<i>SD</i>)	平均値	(<i>SD</i>)	平均値	(<i>SD</i>)
7F	2.55	(0.59)	2.60	(0.66)	2.60	(0.80)	2.88	(0.77)	2.60	(0.74)
8F	2.83	(0.66)	3.00	(0.78)	3.40	(0.97)	3.20	(0.75)	3.13	(0.86)
9F	4.43	(1.04)	4.80	(0.87)	4.85	(0.96)	4.55	(0.67)	4.65	(0.85)
10F	4.20	(0.98)	4.23	(0.98)	4.55	(1.07)	4.80	(0.81)	4.83	(0.91)
11F	5.35	(0.79)	4.95	(0.87)	5.13	(0.84)	5.00	(0.45)	5.35	(0.65)
7E	2.10	(0.62)	2.40	(0.58)	2.15	(0.79)	2.08	(0.64)	1.85	(0.65)
8E	3.93	(1.00)	3.95	(0.81)	3.50	(1.03)	3.93	(0.68)	3.90	(0.77)
9E	4.65	(1.06)	4.63	(1.23)	4.63	(1.13)	4.65	(0.73)	4.90	(0.77)
10E	4.00	(1.05)	3.23	(1.10)	4.30	(0.95)	4.60	(0.66)	4.75	(0.83)
11E	5.68	(0.46)	5.33	(0.60)	5.58	(0.58)	5.45	(0.59)	5.70	(0.46)

トレイトの観点から見ると、言語能力(「日本語A：正確さ」「日本語B：適切さ」)の*SD*は全て1.0以内に収まっている。レベルの観点から見ると、低いレベル(7F, 7E)や高いレベル(11F, 11E)の評価が安定しており、中間レベルの評価の難しいことが分かる。平均値も、表3の参考評価点と大きくは異ならない。違いが認められるのは、7Fの「読み手」と、9Eの「構成・結束性」「読み手」「日本語A：正確さ」「日本語B：適切さ」である。7Fの「読み手」は、日本

語教師全体の評価としてはまとまっているが（平均値：2.60, SD ：0.80）、参考評価点「4」とは異なっている。その理由として、研究者（参考評価点）が他のトレイトと独立させて「読み手」を評価しているのに対し、日本語教師の多くが他のトレイトの影響を受けている可能性が考えられる。例えば、参加者の1人は「内容のみを評価しようとしても（トレイトごとに評価しようとしても）、つい語彙や文法力などにひっぱられ、せっかくトレイトが分かれているのに、結局何となく全体的感覚で評価してしまっている」とフィードバック票に記している。一方、9Eは SD が大きいことから日本語教師の中で評価が割れていることが分かる。そこで、次に20名の日本語教師の評価が参考評価点と実際にどの程度一致していたのか分析する。

表6は、本番用小論文における日本語教師20名と参考評価点のトレイト別相関係数である。これを見ると、総じて両者の相関は高く、20名の平均値（表5）が参考評価点（表3）と大きく違わないことから、信頼性のある評価の行われていることが分かる。特に、日本語能力に関しては、評価者全員において高い相関が認められ、表5の SD が小さいという結果と一致する。しかし、「構成・結束性」「読み手」では、相関の高い評価者もいる一方で、そうではない評価者もいる。この結果は、表5の「構成・結束性」「読み手」の SD にも表れている（e.g., 「構成・結束性」：9E, 10E, 「読み手」：10F, 8E, 9E）。

表6 本番用小論文におけるワークショップ参加日本語教師20名と参考評価点との相関係数

評価者	a	b	c	d	e	f	g	h	i	j
目的・内容	0.62	0.95 **	0.90 **	0.79 **	0.91 **	0.73 *	0.85 **	0.72 *	0.70 *	0.95 **
構成・結束性	0.62	0.87 **	0.60	0.64 *	0.88 **	0.45	0.80 **	0.81 **	0.78 **	0.81 **
読み手	0.60	0.79 **	0.92 **	0.64 *	0.86 **	0.65 *	0.82 **	0.65 *	0.78 **	0.78 **
日本語A：正確さ	0.77 **	0.90 **	0.79 **	0.68 *	0.87 **	0.86 **	0.90 **	0.95 **	0.86 **	0.86 **
日本語B：適切さ	0.76 *	0.88 **	0.92 **	0.73 *	0.81 **	0.87 **	0.95 **	0.94 **	0.95 **	0.81 **
評価者	k	l	m	n	o	p	q	r	s	t
目的・内容	0.75 *	0.69 *	0.84 **	0.94 **	0.76*	0.78 **	0.89 **	0.75 *	0.83 **	0.83 **
構成・結束性	0.46	0.53	0.66 *	0.86 **	0.30	0.91 **	0.69 *	0.77 **	0.71 *	0.95 **
読み手	0.66 *	0.83 **	0.55	0.80 **	0.65 *	0.86 **	0.80 **	0.56	0.60	0.76 *
日本語A：正確さ	0.69 *	0.86 **	0.90 **	0.93 **	0.90 **	0.84 **	0.83 **	0.75 *	0.86 **	0.76 *
日本語B：適切さ	0.70 *	0.87 **	0.88 **	0.88 **	0.91 **	0.96 **	0.93 **	0.81 **	0.87 **	0.86 **

** p <0.01, * p <0.05

次に、具体的に各評価者の評価点を示し、検討を加える。特に、評価の割れている9E（資料4）について評価の安定している11Eと比較しながら見ていく。表7は、9E, 11Eの日本語教師20名のトレイト別評価点である。下の最頻値、中央値、参考評価点の ± 1 の範囲内に収まらないものを表中に網かけで示した。このような範囲内に収まらない評価点は、10編中6編にあるが、20名の全評価点の4.7%にすぎない。その6編のうち、9Eでその割合が12%で最も高く、次いで10Eにおいて10%である。逆に、11Eのように範囲内に収まらない評価点が全くないものは、

表7 日本語教師 20 名による「9E」及び「11E」のトレイト別評価点
 9Eのトレイト別評価点

評価者	9Eのトレイト別評価点					評価者	11Eのトレイト別評価点				
	目的・内容	構成・結束性	読み手	日本語 A 正確さ	日本語 B 適切さ		目的・内容	構成・結束性	読み手	日本語 A 正確さ	日本語 B 適切さ
a	5	5	5	6	6	a	6	5	6	5	6
b	5	6	5	6	6	b	6	6	5	6	6
c	6	4	6	4	5	c	6	5	6	5	6
d	4	4	4	4	4	d	6	6	6	5	6
e	4	5	4	4	5	e	5	5	4	5	5
f	3	3	3	5	4	f	6	6	6	6	6
g	6	6	6	5	6	g	6	6	6	5	6
h	2	5	3	5	6	h	6	6	6	6	6
i	6	6	6	5	5	i	5	4	5	5	5
j	5	4	5	4	4	j	6	6	6	5	6
k	3	2	2	5	5	k	5	6	5	6	6
l	5	5	6	5	5	l	6	5	6	6	6
m	5	5	5	5	5	m	5	5	6	6	5
n	5	5	5	5	5	n	6	5	6	6	6
o	4	3	4	4	5	o	5.5	4.5	5.5	5	5
p	6	6	6	5	5	p	5	5	5	4	5
q	5	6	4	4	5	q	6	5	6	6	6
r	4	4	4	3	3	r	6	5	6	6	6
s	5	3	5	4	4	s	6	5	5	6	6
t	5	6	5	5	5	t	5	6	5	5	5
平均値 (SD)	4.65 (1.06)	4.63 (1.23)	4.63 (1.13)	4.65 (0.73)	4.90 (0.77)	平均値 (SD)	5.68 (0.46)	5.33 (0.60)	5.58 (0.58)	5.45 (0.59)	5.70 (0.46)
最大値	6	6	6	6	6	最大値	6	6	6	6	6
最小値	2	2	2	3	3	最小値	5	4	4	4	5
最頻値	5	6	5	5	5	最頻値	6	5	6	6	6
中央値	5	5	5	5	5	中央値	6	5	6	5.5	6
参考評価点	5	6	6	6	6	参考評価点	6	6	6	6	6

8F, 9F, 11F, 11E の4編である。また、このような ± 1 の範囲内に収まらない評価点は特定の評価者に集中しているのではなく、20名中14名に認められる。つまり、単発的な評価のずれは、どの評価者にも起こりうると考えられる。

9Eは通学授業と遠隔授業の併用案であり、もっともSDが大きい。例えば、評価者h(表7)は「目的・内容」で「2」を付けており、評価結果記入票に「説明不足。聞きかじった話を、自分のものとしてよく消化しないうちに使っている印象がある」と記している。一方、評価者p(表7)は「独創性もある。主張も一貫している」と記して「6」を付けている。評価点「2」は、評価者20名の平均値、最頻値、中央値、参考評価点からかけ離れているが、表6を見ると、評価者hは、全体的にも、また「目的・内容」においても信頼できる評価をしている。したがって、評価者hの9Eにおける「目的・内容」の評価は、この小論文における単発的な個人レベルのずれだと考えられる。

9Eにおいては、さらに「構成・結束性」で最頻値、中央値、参考評価点が「6」ないし「5」

であるが、4名が「2」ないし「3」を付けている。「構成・結束性」に関しては、個々の小論文の「目的・内容」との関連もあり（田中・初鹿野・坪根，1998），プロンプトのガイドラインであらかじめ方針を示すことは難しい。今後の検討課題とする。

9Eの参考評価点に関しては、参加者から、なぜ全体的に「6」なのかという疑問が出されている。一方、11Eの「6」の評価にはみなが納得している。9Eは前述した併用案の例として出したものであったが、プロンプトのガイドラインで方針を示してあったためか、参加者から併用案に関しての意見や疑問は殆どなかった。なぜ日本語教師が「6」であることに疑問を持ったかについては、2点考えられる。

第1点は、評価基準B（レベル別基準）の「6」のレベル説明が「ほぼ完璧にできている」や「母語話者に近いレベルで、正確に／適切に日本語を使用している」と記されており、基準説明に忠実に従えば「5」のレベルに近かったからであろう。また、筆者ら研究者は「元データ」全てを評価しており、1～6レベルの小論文を広く把握したうえで9Eが高いレベルに入ると判断したのに対し、ワークショップ参加者は、1つのプロンプトにつき5編の小論文、サンプルや練習小論文を含めても各プロンプト11編の小論文に当たっただけで、さらに「完璧な6」が存在するはずだと考えたとも推察される。もう1点は、小論文を見る際、日本語教師の場合には個々の細かい点に注意が向けられているのに対し、研究者は各トレイトの中身を総合的に見ていることが考えられる。例えば、日本語の「正確さ」に関して9Eには確かに漢字や活用、助詞等の間違いも認められるが、日本語教師の場合にはその個々の間違いに目が向き、全体的に使われている文型や表現のレベルの高さを評価することが難しくなっているのではないだろうか。「適切さ」に関しても、全体的には小論文にふさわしいレジスターを備えているので、研究者4名には「6」と判断されたのだろう。

3-3. アンケート調査に関して

ワークショップ終了後のアンケート調査（田中他，2007）から、マルチプル・トレイト評価表が参加者に理解され、教育現場での活用が可能な評価基準と受け止められたことが分かった。評価基準Aについてはほとんどの回答者が「役に立つ」と答え、とくに「読み手」の重要性に関心が集まると同時に、その評価の難しさについても指摘があった。評価基準Bについては、多くが「役に立つ」と回答したが、3-2の9Eの評価において述べたように、レベルと基準説明に改善すべき点のあることが明らかになった。

採点者間ミーティングについては、他の評価者との意見交換は有意義ではあったが、ディスカッションの時間が不足していたことや指示が曖昧だったことから納得できる評価の調整まで至らなかったと答えた参加者が多かった。今後、ミーティングの目的、手順を明確にしたうえで採点者間ミーティングの役割を考えていく必要があるだろう。

4. 今後の課題

今回のライティング評価ワークショップは、反省すべき点も多いが、初の試みとしては概ね成功したと言えよう。また、今回提示した評価基準、評価方法については、評価がそれほど難しくないと予想される小論文が対象であり、ライティング教育・評価の経験のある日本語教師が評価者であるという限定付きではあったが、有効に機能した可能性が示された。評価の割れそうなり難しい小論文に関しては別に検討する予定である。

今回は時間的制限があり無理であったが、ワークショップは経験者が対象の場合でも最低2日必要である。1日目は評価基準の把握と評価が難しくないと予想される基本的な小論文の評価、2日目は評価が割れそうなものの評価とそのフィードバックに当てる。さらに、日本語教育、特にライティング教育・評価の経験に応じて複数コースが必要であろう。

ワークショップのアンケートにもディスカッションで他の人の評価に対する考え方が分かって有意義であったという意見があったが、採点者間ミーティングは自己の評価を省み、評価の認識を互いに新たにする機会となる。今後は、ある1つの共通した評価基準を使って評価を一致させること、即ち、評価者間信頼性 (inter-rater reliability) を高めることを目指していかなければならない。

今回のワークショップで明らかになったことの1つは、評価が一致しやすいトレイトとそうでないトレイトがあるということである。今回の結果では、前者が言語能力、後者が「構成・結束性」「読み手」で、「読み手」に関しては、注2に記したように目的に応じてはずせばよい。「構成・結束性」に関しては、個々の小論文のさらなる分析が必要であろう。

今後の課題としては、個人レベルでのずれについての検討がある。ずれは、評価基準の問題や疲労等から起こる場合もあるが、評価に個人の言語観、教育観、ライティング教育経験、背景知識等が関係することは避けられない。どのような個人的要素が評価のどの部分 (トレイト、レベル、特定の小論文等) に影響するのかを特定することは難しいが、個人差を是正、減少することは可能だと思われる。今後、さらに研究を進め、その成果を評価基準やワークショップ、トレーニング方法に反映させたい。

評価基準は、各機関・コース・クラスの目的に合わせて作成すべきものであるが、まずは日本語教育学界で、基本的なライティング評価の枠組みについて、ある程度共通した認識を持つ必要がある。評価者となる日本語教師も、評価基準やプロンプトに合わせて柔軟に対応し、評価の安定度 (評価者内信頼性) を向上させるよう努めるべきであろう。そのためにも、ライティング評価の公的なワークショップの機会を設けることが望まれる。

謝 辞

小論文の研究使用を快諾してくださった学生の方々、評価基準の検討並びにアンケートにご協力いただいた短期研修参加者の先生方に記して感謝の意を表したい。なお、本研究は平成 16-18 年度科学研究費補助金基盤研究 (C) 「第二言語によるライティングについての基礎研究: Good writing とは何か」(課題番号 16520321) 及び平成 19-21 年度科学研究費補助金基盤研究 (C) 「第二言語としての日本語ライティング評価: Good writing のさらなる追求」(課題番号 19520448) の助成を受けた。

参 考 文 献

- 庄司恵雄 (代表) (2005) 「非母語話者教員による日本語口頭能力の判定に向けて」平成 16 年度学術振興会 科学技術研究補助金研究基盤研究 (C) (1) 企画調査研究成果報告書
- 菅井英明 (2003) 「記述テストの特徴と比較」菅井英明 (代表) 「日本語教育における評価法に関する基礎的資料整備とその分析」平成 13 年度～14 年度文部科学省科学研究費補助金基盤研究 (C) (2) 研究成果報告書, 80-95
- 鈴木秀明 (2005) 「短時間の評価トレーニングが教師の発話評価に及ぼす効果」『言語科学研究』11, 神田外国語大学大学院, 77-94
- 田中真理 (2005) 「日本語教育におけるライティング評価」国立国語研究所編『日本語教育年鑑 2005 年版』くろしお出版, 42-52
- 田中真理代表 (2007) 「第二言語によるライティングについての基礎研究: Good writing とは何か」平成 16 年度～18 年度科学研究費補助金基盤研究 (C) 研究成果報告書
- 田中真理・初鹿野阿れ・坪根由香里 (1998) 「第二言語としての日本語における作文評価—「いい」作文の決定要因—」『日本語教育』99, 日本語教育学会, 60-71
- 田中真理・長阪朱美 (2004) 「日本語と英語を目標言語とするライティング評価基準の展望: 第二言語としての日本語のライティング評価基準作成に向けて」『第二言語としての日本語の習得研究』7, 第二言語習得研究会, 214-253
- (2006) 「第 2 言語としての日本語ライティング評価基準とその作成過程」国立国語研究所編『世界の言語テスト』くろしお出版, 253-276
- 西川寛之・西部由佳・山中郁・山辺真理子 (2003) 「パフォーマンス・アセスメント中心の口頭表現能力テスト—比較可能性の保証を高める評価者の涵養—」『日本語教育学会春季大会予稿集』日本語教育学会, 183-188
- Elder, C., Barkhuizen, G., Knoch, U. & von Randow, J. 2007. Evaluating rater responses to an online training program for L2 writing assessment. *Language Testing* 24, 37-64.
- Hamp-Lyons, L. 1991. Scoring procedures for ESL contexts. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts*, Norwood, NJ: Ablex, 241-276.
- 2007. Worrying about rating. *Assessing Writing* 12, 1-9.
- Knoch, U., Read, J. & von Randow, J. 2007. Re-training writing raters online: How does it compare with face-to-face training? *Assessing Writing* 12, 26-43.
- Kondo-Brown, K. 2002. A FACETS analysis of rater bias in measuring Japanese second language writing performance. *Language Testing* 19, 3-31.
- Show, S.D. & Weir, C.J. 2007. *Examining writing: Research and practice in assessing second language writing*, Studies in Language Testing 26, Cambridge: Cambridge University Press.
- Taylor, L. & Falvey, P. (Eds.). 2007. *IELTS collected papers: Research in speaking and writing assessment*, Studies

- in *Language Testing* 19, Cambridge: Cambridge University Press.
- Weigle, S.C. 1994. Effects of training on raters of ESL compositions. *Language Testing* 11, 197-223.
- 1998. Using FACETS to model rater training effects. *Language Testing* 15, 263-287.
- 2002. *Assessing writing*. Cambridge: Cambridge University Press.
- White, E.M. 1984. Holisticism. *College Composition and Communication* 35, 400-409.
- 日本学生支援機構 〈http://www.jasso.go.jp/eju/kizyutu_boshu.html〉, 2007年10月25日参照

資料 1

評価基準A: トレイト別・基準説明

トレイト		基準説明	[表の見方の 補足説明]
1. 目的・内容	目的 課題の達成 内容 トピック 主張 サポート	<ul style="list-style-type: none"> このライティングの目的、すなわち、課題が達成されているか (例えば、「AとBについて比較・対照し、自分の意見を述べる」、「ある意見について賛成か反対か論証し、それを大学新聞に投稿する」など) トピックに合った内容か 主張(main idea: 一番言いたいこと)が明確か: <ul style="list-style-type: none"> 文章全体を通して一貫性があるか 妥当性(正確さも含む)、説得力があるか 主張がサポートされているか: <ul style="list-style-type: none"> 理由、例、説明などが適切で(正確さも含む)、客観性、説得力があるか 	
2. 構成・結束性	文章全体 段落 文	<ul style="list-style-type: none"> 文章全体の構成が明確か: <ul style="list-style-type: none"> 「はじめに」(Introduction)と「本文」(Body)と「結論・まとめ」(Conclusion)に相当する部分があるか 文章全体が論理的・意味的につながっているか 段落と段落のつながりがスムーズか: <ul style="list-style-type: none"> 前後の段落が意味的につながっているか 段落と段落をつなぐ接続表現(「まず」「次に」「しかし」「一方」など)が効果的に使われているか 指示語(「これらの～」など)が効果的に使われているか 段落内の文と文のつながりがスムーズか: <ul style="list-style-type: none"> 文と文をつなぐ接続表現(「たとえば」「しかし」「したがって」「特に」など)が効果的に使われているか 指示語(「この～」「それらは」「その中でも」など)が効果的に使われているか 	・ライティングの「内容」自体の妥当性、客観性、正確さなどは、1の[内容]において評価する ・3の[読み手]においては、1の[内容]において評価できない側面—読み手に対する配慮があり、その結果、読み手に分かりやすくなっているか、また、読み手にとって面白いか—という点を評価する ・[A「正確さ」とB「適切さ」]に入る項目を厳密に分けることは難しいが、[A「正確さ」]では誤り、意味不明の箇所などがないかどうかを見て、[B「適切さ」]では誤り、意味不明ではないが、このライティングにふさわしい形式やスタイルであるかどうかを見る ・評価表の基準説明に記載されているとは、全て「基本的に」にできているかどうかというところであり、特に[A「正確さ」]や[B「適切さ」]において、誤りや不適切な箇所の数を問題にするものではない
3. 読み手	配慮 面白さ	<ul style="list-style-type: none"> 読み手を意識して書かれているか (例えば、大学新聞に投稿するという課題では、その読者など) 読み手に対する配慮があるか (例えば、読み手にないと思われる文化的背景を説明するなど) 読み手にとって興味深いか: <ul style="list-style-type: none"> 創造性、オリジナリティ、新鮮さなどがあるか 引き込まれて読んでいくか 	
4. 日本語 (言語能力)	A 正確さ 文法 文型・構文 語彙 句読法 表記 B 適切さ 形式 レジスター a. 表現・語彙 b. 文末スタイル c. 漢字とかなの使用バランス	<ul style="list-style-type: none"> 文法が正確か: <ul style="list-style-type: none"> 助詞、動詞や形容詞などの活用形、時制、アスペクト、ヴォイスなどが正確か 文が正確に組み立てられているか: <ul style="list-style-type: none"> 複文の主節と従節の関係や呼応関係(「なぜなら、…からである」など)が正確か 単文の羅列や同じ文型の繰り返しではなく、文脈にふさわしい文型や日本語らしい構文が使いこなせているか 語彙が正しく使われているか 句読点(、)や符号(「」)などが適切に使えているか 正しい表記(漢字・ひらがな・カタカナなど)が使えているか このライティングに合った形式(「小論文・レポート」「投書文」などの形式)が使われているか (例えば、「小論文」が随筆調になっていたり、「レポート」の最後に読み手(教師)への個人的なメッセージがついていたりしないか、など) このライティングに合ったレジスター(以下のa～c)が適切に使われているか <ul style="list-style-type: none"> a. このライティングに合った表現・語彙が適切に使われているか <ul style="list-style-type: none"> 「話しことば」と「書きことば」の区別ができていないか 不適切な敬意表現がないか(例えば、「遠隔授業はほどほどに」としても「便利」など) b. 「普通体」と「丁寧体」の区別ができ、文末スタイルが適切に使われ、基本的に統一されているか c. 漢字とかながバランスよく使用されているか (例えば、「小論文」がほとんど「ひらがな」で書かれていないかなど) 	

資料 2

評価基準B:レベル別・基準説明

トレイト	レベル・スコア	基準説明
1. 目的・内容	<p>6 Very good とてもよい</p> <p>5 Fairly good かなりよい</p> <p>4 Good よい</p> <p>3 Fair ふつう</p> <p>2 Not good 十分ではない</p> <p>1 Poor まだまだ</p>	<p>明確な目的・課題の達成、トピックにあった内容、明確で一貫性のある主張、主張の十分なサポート</p> <ul style="list-style-type: none"> ほぼ完璧にできている 目的・課題の達成に問題点はないが、際立ってはいない。主張は一貫していない。読みの視点・立場はない。読み応えのある内容になっていない。など 目的・課題の達成はなされたが、一部が十分でない。主張の一貫性にやや欠ける。サポートはあるが具体性に欠ける。など 目的は一応示され、主張も分かるが、問題点がある。「サポート」が十分でないために、客観性、説得力に欠ける。主張にやや偏りが見られ妥当性に欠ける。など 書いたことはおおよそ分かるが、「内容」としてのまとまりがなく、課題の達成は十分である。量的・質的トピックについて言及が少ない、あるいは、反対にトピックとは関係のないことへの言及が多い。主張の一貫性がない。など 反対にトピックについては書かれているが、目的が明確でなく、言いたいことが理解できず、課題は達成されていない
2. 構成・結束性	<p>6 Very good とてもよい</p> <p>5 Fairly good かなりよい</p> <p>4 Good よい</p> <p>3 Fair ふつう</p> <p>2 Not good 十分ではない</p> <p>1 Poor まだまだ</p>	<p>文章全体の構成、適切な段落分け、段落と段落の関係、段落内の文のつながり</p> <ul style="list-style-type: none"> ほぼ完璧にできている 文章全体の構成や段落の問題はないが、段落内にやや不自然さがある。段落間の文のつながりにやや不自然さがある。など 文章全体の構成や段落の一部が十分でない。「はじめに」(Introduction)や「結論・まとめ」(Conclusion)が文だけから成る。あるいは「本論」(body)の段落分けがやや細かすぎたり、荒すぎたりする。など 文章全体の(意味的な)流れはあるが、形式的段落(改行による段落)が不自然である。「はじめに」または「結論・まとめ」のどちらかが明らかに欠落している。段落分けがあまりに細かすぎたり、荒すぎたりする。など 文章全体の構成がかなり不自然である。形式的段落は作られているが、段落と段落が意味的につながっていない。など 文章全体の構成が認められない。形式的にも段落分け、意味的にも段落に属するものがなく、「構成」についての知識が認められない
3. 読み手	<p>6 Very good とてもよい</p> <p>5 Fairly good かなりよい</p> <p>4 Good よい</p> <p>3 Fair ふつう</p> <p>2 Not good 十分ではない</p> <p>1 Poor まだまだ</p>	<p>読み手に対する配慮、読み手にとっての興味深さ</p> <ul style="list-style-type: none"> 読み手に背景知識がない場合にはそれを提供し、読み手を引き込み、読み手にも考えさせる工夫がある。など 読み手に背景知識がない場合にはそれを提供しているが、やや読み手を引き込んでいく要素に欠ける。など 読み手を惹きつけているが、説明不足の箇所があったり、読み手を引き込んでいく要素に欠ける。など 当然読み手にはないと思われる知識や情報(例えば、国の文化や習慣など)についての説明がないために、その国の事情を知らない読み手には分かりづらく、興味を持たせることは難しい。など 書き手中心(読者不在)のライティングだという印象を与える 書き手中心(読者不在)のライティングであり、読み手には理解できない
4. 日本語(言語能力)	A 正確さ	<p>正確な文法、多様で効果的な文型、正確な語彙・句読法・表記</p> <ul style="list-style-type: none"> 母語話者に近いレベルで、正確に日本語を使用している 大きな誤りはなく、特に「多様で効果的な文型」を使用している 多少の誤りがあったとしても、意味を十分理解できる程度の正確さがある。しかし、「多様で効果的な文型」はあまり使われていない 誤りがあるが、意味は何とか理解できる程度の正確さがある アカデミック・ライティングを書けるレベルにまだ達していない。意味のよく分らない箇所がかなりあり、文章全体の理解を妨げる アカデミック・ライティングを書けるレベルではない。全体に不正確で、意味が分らず、読者の心が困難である
	B 適切さ	<p>適切な形式、レジスター(表現・語彙、文末スタイル、漢字とかなのバランス)</p> <ul style="list-style-type: none"> 母語話者に近いレベルで、適切に日本語を使用している 重大な不適切な箇所はなく、よくに課題(のジャンルなど)にふさわしい表現・語彙が使われている 表現・語彙には多少の不自然さがあったとしても、文末スタイルがそろい、漢字も適度に使われている 表現・語彙が不自然であったり、「話しことば」がかなり混じっていたり、文末スタイルが統一されていなかったり、漢字の使い方がなかったりする アカデミック・ライティングを書けるレベルにまだ達していない。表現・語彙がかなり不自然で、「話しことば」と「書きことば」の区別や「普通体」と「尊体」の区別ができていない。あるいは、漢字がほとんど使われず、ひらがなで書かれている。など アカデミック・ライティングを書けるレベルではない。表現・語彙が極めて乏しく、文章全体が稚拙で、漢字もほとんど使われていない。など
	0 N/A. 該当せず	<p>白紙やそれに近い状態。あるいは、課題を無視して書かれている場合や大部分が目録言語(日本語)以外の言語で書かれている場合(注) 目録言語(日本語)以外の言語が単語レベルで使われていても、大きく評価を下げることはしない。4. Bレベルの「漢字がほとんど使われず、ひらがなで書かれている」と同じような観点から評価する</p>

資料3

プロンプト(課題文)とガイドライン

<プロンプト: 通学授業と遠隔授業>

大学が通学授業と遠隔授業(e-learning)について、学生から小論文を募集しています。大学で勉強する方法として、2つ考えられます。1つは、学生が大学に行って、教師や他の学生と授業で直接コミュニケーションをとりながら勉強する方法(通学授業)です。もう1つは、大学に行かずに、インターネットなどの情報ネットワークとコンピュータを利用して、教師とコミュニケーションをとりながら、自宅などで勉強する方法(遠隔授業:e-learning)です。

あなたなら、どちらの方法を選びますか。理由をあげて、あなたの意見を600字程度で書いてください。

<ガイドライン>

目的・内容

- ・ どちらを選ぶかが書いてあるか、また、その理由が述べられているか
- ・ プロンプトに従い、どちらを選ぶかが書いてあることが原則だが、通学授業と遠隔授業の併用案も、理由が明確に述べられ、説得力があると判断されれば、大きく評価を下げることはしない

読み手

- ・ 大学(小論文を募集している広報部や授業改革委員会など)を想定して書かれているか

日本語 B 適切さ

- ・ 「小論文」の形式、レジスターが使われているか
- ・ 普通体の方が望ましいが、文末スタイルが統一されていることの方が重要である

その他

- ・ 今回は、長さ(600字程度)をかなり上回っていても、評価を下げることはしない

資料4

ライティング No. 1 <遠隔授業と通学授業>

9E

[タイトル:

]

[名前:

][2005年 月 日]

大学の授業では通学授業と遠隔授業を適度に混合した方がいい
 と思う。そうしたら双方の長所だけを取る事ができるためである。では、
 各授業の長短には何かあるのだろうか。

まず、通学授業は真のコミュニケーションができる。実際に、人達の顔を見ながら、話し合えるのが、心の真のつながりをもたらすのである。ほかに、授業以外の時間にも一緒にご飯を食べたり、歩き回りながら、授業以外のいろいろな話しかける。

その反面、インターネットでの情報ネットワークとコンピュータを利用の全くないとしたら、急変している情報化社会に適応できなくなるはずだ。大学では必ずネットワーク時代に依りて、学生にコンピュータ^{操作性}コミュニケーションの方法を養うべきである。それは実際にそのよは授業をやってみるのが一番いいということに違いない。

遠隔授業の長所はまさにそういうことにある。つまりパソコンのコミュニケーションの経験ができる。短所は皮相的な人間関係が生じるということだ。みんなネットワークでしか会えないことになって、実際にはいない、つまり仮象の人間関係になってしまうということである。

従って、大学では双方を適度に混ぜてやった方がいい。例えば、一週間に一日は遠隔授業をやる。もしくは、毎日の最後の授業は遠隔授業に関する方がいいと思う。なぜなら21世紀には実際に接している人との関係とネットワークのコミュニケーションしている人との関係、両方が上手にできる人を要するためである。

サンプル作文 3 E 評価と解説

3 E

通学授業 (attending classes) と遠隔授業 (e-learning)

タイトル:

私は学生が大学に行って、教師や他の学生と授業で直接コミュニケーションをとりながら、勉強する方法、つまり通学授業を選びたいと思う。

なぜならば、大学は小さい社会と考えられる。学生たちは大学で単に知識的面の勉強ではなく、将来社会への進出前の準備でも言えるだろう。キャンパス内で授業をとりながら、先生と他の学生とのコミュニケーションもしている。なにか分からないことがあれば、すぐまわりの人に聞け、他の人が困難があった時、自分も手伝ってあげられる。こう言う人間と人間直接的なコミュニケーションは、将来社会への進出するときは異常に大事なのだ。この能力は、学校内人と人のコミュニケーションをしないながら、次第に身に付けるものだ。ただコンピュータのスクリーンの前に座って、相手の顔も見えずにキーボードを打って文字で交流することは、その能力を養成できなと思う。新聞も雑誌もたくさん例を挙げた、長い時間インターネットで人と交流する人のコミュニケーション能力は退化していく。人間のコミュニケーションは、もともと基本なのは相手との直接話し合うことである。学校でも職場でも、そのコミュニケーション能力が低下すれば、勉強も仕事もスムーズに進めないと思う。

だからこそ、私はやはり大学を通っての通学授業を選びたいと思う。

1. 目的・内容: 5 Fairly Good かなりよい

主張が一貫しており、コミュニケーションとはどういうことか、キャンパス内で授業を受けるとはどういうことかを具体的に説明し、通学授業の利点をサポートしている。ただし、遠隔に対する描写で「長い時間インターネットで人と交流する人のコミュニケーション能力は退化」という部分は、十分な根拠が示されていないので、サポートとして弱い。

2. 構成・結束性: 4 Good よい

文章全体の構成は一応できている。ただし、第2段落は、途中から遠隔授業に対する批判となっているが、新聞や雑誌のいうインターネットの欠点の部分は十分な説明を要するので、新たに段落を設けて論を進めたほうが分かりやすい。さらに、最後の段落は単なるくり返りで、1行で終わっており、段落の構成としては弱い。

3. 読み手: 5 Fairly Good かなりよい

大学は小さな社会であり、人間のコミュニケーションの基本は直接話し合うという考え方や、「まわりの人に聞ける」「自分も手伝ってあげられる」等、実際の大学生生活を生き生きと描写しているところには、読み手は引き込まれていく。しかし、インターネットについては、筆者のやや一面的な立場から書かれている。

4. 日本語 A 正確さ: 5 Fairly Good かなりよい

文法に大きな誤りはなく、多様な文型の使用、正確な語彙の使用がみられるが、表記(「小い」)や語彙(「ずく」、「異常に大事」)等で、母語話者とは異なるレベルにある。

4. 日本語 B 適切さ: 5 Fairly Good かなりよい

普通体を使って統一し、小論文としてまとまったスタイルになっている。しかし、「なにか分からないとこ」「こう言う人間」等、話しことばがあり、改善すべき点が残っている。