

Lexical Frequency Profile を用いた L2 ライティング における語彙的豊かさの評価

杉森直樹

Abstract

This study investigates the validation of the Lexical Frequency Profile (LFP) measure for the assessment of the lexical richness of essays written by Japanese learners of English. This study also examines how the LFPs of the essays may vary according to the students' proficiency level. Two essay writing assignments were given to 40 Japanese college students and their essays were subjected to LEP analysis using a software called *Range*. The result of the analysis shows that the high-proficiency students significantly used the academic words (AWL) more frequently than the low-proficiency students. Concerning the stability of the profile, the LFPs of the first 2000 words showed considerable stability between the two essays. The LFPs of the words beyond 2000, however, showed no stability between the essays. This study revealed the potential of the LFP as a statistical measure to assess the lexical richness of a composition written under the condition that dictionary use was allowed for writing.

Keywords : Lexical Frequency Profile, essay writing, productive vocabulary, learner corpus, lexical richness

はじめに

Brown (2001) によれば、英語学習者のライティングに対する評価指標の一つに、その使用語彙 (vocabulary) が挙げられている。ライティングにおける使用語彙の評価においては、発表語彙として学習者が適切な語彙を正しく使用できているかどうかや、語彙レベルの幅も含めてどの程度多様な語彙を使用できているかなどがその基準となる。しかしながら、エッセイライティングにおける使用語彙の評価基準に関しては、多くの場合、大まかな scoring rubric や descriptor が示されているだけであり、実際には評価者の判断に任されることも多い。この様な評価上の問題点を改善する方法の一つとして、コンピュータを用いた使用語彙の評価法が挙げられる。この方法では、学習者のライティングを専用のソフトウェアにより計量的に分析して使用語彙の評価を行うもので、学習者のライティングを客観的に評価することが可能となる。Read (2000) によれば、語彙的豊かさ (Lexical richness) を計測するための統計的評価基準には、異語数 (Type) と総語数 (Token) の比率を表す Type-token ratio (TTR)、内容語と機能語の比率を表す Lexical density (LD)、内容語のみの TTR を計測する Lexical variation (LV) などがあ

るとされているが、これらの指標に加えて、使用語彙のレベルまで測定する Lexical Frequency Profile (LFP) と呼ばれる指標がある。LFP は、学習者がどのレベルの語彙をどの程度使用しているかを計測する指標であり、基本レベルの平易な語彙から、高度なアカデミック語彙までの使用割合を計量的に分析するものである。

本研究は、日本人英語学習者の L2 エッセイライティングにおける使用語彙の豊かさを LFP を用いて統計的に分析し、学習者の使用語彙の特徴を調査すると共に、LFP の評価指標としての妥当性を考察するものである。

1. LFP による使用語彙の評価法

LFP は Laufer and Nation (1995) で提唱され、英語学習者のライティングにおける使用語彙の豊かさを示す指標の一つである。LFP は、学習者のライティングにおける使用語彙を 4 つの頻度レベルに分類し、それぞれのレベルに属する単語家族 (word family) の割合を求めて、それらの相対的使用割合を計測するもので、4 つのレベルとは、1) 最も使用頻度が高い 1000 語 (1st 1000)、2) 次に使用頻度の高い 1000 語 (2nd 1000)、3) Academic Word List (AWL) の 836 語、4) それ以外の語 (NiL: Not in the lists) である。

例えば、200 語で書かれた英文エッセイのうち、150 語が 1st 1000 に含まれ、20 語が 2nd 1000 に含まれ、20 語が AWL に、残りの 10 語が NiL になったとすると、このエッセイの LFP は 75%—10%—10%—5% となる。これらの語彙のレベル判定の基準となる語彙リストについては、1st 1000 と 2nd 1000 の上位 2000 語は、West (1953) の General Service List (GSL) を使用しており、3) の AWL は、元は University Word List とされていたものを Coxhead (1998) で改訂された Academic Word List (AWL) を使用している。LFP の算出にこれらの語彙リストを用いるものに加えて、最近では、British National Corpus (BNC) の頻度データをベースにした語彙リストを用いて、語彙のレベルを 16 段階に分析できるものもある。

Laufer and Nation の LFP に関する先行研究では、熟達度 (proficiency) が異なる 3 つのグループの学習者にそれぞれ 300 語程度のエッセイを 2 つ書かせてその LFP を分析している。それによれば、1) 1st 1000 語の語彙は proficiency の低いグループの方が使用割合が多い、2) 2nd 1000 語に関してはグループ間での使用頻度に有意差は無い、3) AWL の語彙の使用頻度は、proficiency の低いグループが有意に低い、4) Proficiency の高い学習者は、2nd 1000 と NiL 以外のレベルの LFP がトピックによって変動しやすい、5) Proficiency の低い学習者は、どのレベルでも LFP は変動しにくい、という結果が報告されている。また、学習者の英語力の差が特定の語彙レベルの語の使用割合の差として LFP に表れたことから、LFP は学習者の熟達度を測定する指標になるとされている。

2. 研究目的

LFP の指標としての妥当性に関しては、Meara (2005) のように、モンテカルロ法を用いたシミュレーション実験でその問題点を指摘する研究もあり、更なる検証が必要であると考えら

れる。また、Laufer and Nation の研究で被験者となった学習者は、ビクトリア大学での EFL 学習者とイスラエル人の英語学習者（大学生）であり、書かせる時間を制限して授業中に実施した *timed writing* のデータを分析したものである。一般に、ライティングの評価研究においては、書かせる時間を制限し、辞書の使用を許可しないという条件で書かせることが多いが、Weigle (2002) によれば、従来、辞書使用を許可しないで書かせた場合のライティング評価研究は多くなされてきたが、辞書使用を許可した場合についてのライティング評価研究はまだ不十分であると述べられている。そこで、本研究においては、これまであまり調査が行われてこなかった日本人英語学習者のエッセイの LFP を調査することとし、ライティングの条件も先行研究とは変更して、宿題として自宅でエッセイを書かせる形とした上で、辞書使用を許可した形で書かせることとした。これらのことをふまえて、本研究では、日本人英語学習者のエッセイライティングにおいて、1) エッセイのトピックが異なると LFP はどの様に変化するか、2) 上位学習者と下位学習者に同じトピックでエッセイを書かせた場合、両者の LFP にはどのような差が見られるのか、の 2 つのリサーチ・クエスチョンを設定し、これらを調査することとした。

3. 調査・分析方法

本研究は、理系学部の日本人大学生を対象とし、1 年生前期のライティングの授業に於いて書かせたエッセイを収集した。被験者となる学習者は、入学時にリスニング、リーディング、文法・語彙の問題で構成された独自のプレースメントテストを受験し、そのスコアに基づいてレベル別のクラスに分けられている。今回は、共に *intermediate* レベルに属する上位クラス 20 名、下位クラス 20 名のエッセイ課題 2 回分をデータとして収集した。なお、上位クラスと下位クラスとの間の英語力の差に関しては、プレースメントテストの得点で分散分析を行った結果、両クラスの間には有意差があることが確認された ($F(1, 39) = 315.90, p < .001$)。

ライティング課題は、250-300 語の *descriptive* なエッセイを書くもので、3 週間の間隔を空けて 2 つのエッセイを書かせた。与えられたエッセイライティングの課題は、Essay 1 のタイトルが、"*Things I Like to Do*" で、Essay 2 のタイトルが、"*What I Should Do to Be a Better Student*" である。エッセイ課題は電子データとして収集してファイル化したが、エッセイ中のスペリングミスについては、誤りが明白なものについては正しいものに修正した。また、固有名詞については分析から除外した。LFP の分析には、Paul Nation により開発された *Range* と呼ばれるソフトウェアを使用した¹⁾。*Range* に使用したワードリストは、前述した GSL の 2000 語と AWL である。学習者が書いたエッセイのファイルを *Range* に読み込んで 1st 1000, 2nd 1000, AWL, NiL のレベルで LFP の分析を行った。この LFP の分析を、2 種類のエッセイそれぞれにおいて行った。

この様にして算出された LFP に対して、最初に、同一の学習者に異なるトピックで複数のエッセイを書かせた場合の LFP の値の変動を見るため、上位群、下位群に分けて、2 種類のエッセイ間の LFP の値の差を検証する分析を行った。今回は、調査対象とする学習者の数が少ないこともあり、ノンパラメトリック検定を用いて、4 つの語彙レベル毎に Wilcoxon signed-rank test を行って 2 種類のエッセイの LFP の値の差を検証した。次に、学習者の英語力の差の LFP に対する影響を調査するため、エッセイごとに、上位群と下位群の LFP の差を Mann-Whitney U-test

を用いて検証した。

4. 結果

表1は、上位群の被験者が書いた2つのエッセイのLFPの平均値と検定の結果を示している。また、表2は、下位群の被験者が書いた2つのエッセイのLFPの平均値と検定の結果を示している。これらの表が示すように、上位群と下位群の両方に於いて、1st 1000と2nd 1000の語彙の使用割合に関しては、トピックが異なってもLFPの値には有意差が見られず、これらのレベルではトピックの影響を受けない傾向が認められた。その一方で、AWLの語彙とNiLの語彙に関しては、2つのエッセイのLFPの値には有意差が見られ、これらのレベルの語彙の使用割合が、トピックによって変動することを示している。

表1 上位群のLFP平均値

Upper Group	1st 1000	2nd 1000	AWL	NiL	
Essay 1 Mean	80.81	8.68	3.68	6.83	(%)
(SD)	(4.02)	(2.87)	(2.16)	(2.91)	
Essay 2 Mean	80.23	7.40	7.20	5.17	(%)
(SD)	(4.36)	(2.18)	(2.61)	(2.03)	
<i>T</i>	92	58	15***	47*	
<i>p</i>	.627	.079	<.000	.030	

p* < .05 *p* < .01 ****p* < .001

表2 下位群のLFP平均値

Lower Group	1st 1000	2nd 1000	AWL	NiL	
Essay 1 Mean	83.22	7.01	1.98	7.80	(%)
(SD)	(2.95)	(2.54)	(1.44)	(2.21)	
Essay 2 Mean	82.77	6.69	4.90	5.64	(%)
(SD)	(4.82)	(2.60)	(1.71)	(2.09)	
<i>T</i>	103	93	10***	46*	
<i>p</i>	.941	.654	<.000	.028	

p* < .05 *p* < .01 ****p* < .001

表3と表4は、それぞれEssay 1とEssay 2における上位群と下位群のLFP平均値と検定の結果を示している。これによれば、両方のエッセイにおいて、平易な単語が多い1st 1000のレベルは、下位群の学習者の方が多く使用しているが、2nd 1000からAWLのレベルの単語になると、

上位群の学習者の方が多く使用するようになるという傾向が見られる。検定の結果では、1st 1000, 2nd 1000, NiL の 3つのレベルでは LFP には上位群と下位群の間で有意差は認められなかった。一方、AWL レベルの LFP の値には、どちらのエッセイにおいても上位群と下位群との間で有意差が認められた。

表 3 Essay 1 の LFP 平均値

Essay 1	1st 1000	2nd 1000	AWL	NiL	
Lower Level Mean (n=20)	83.22	7.01	1.98	7.80	(%)
(SD)	(2.95)	(2.54)	(1.44)	(2.21)	
Upper Level Mean (n=20)	80.81	8.68	3.68	6.83	(%)
(SD)	(4.02)	(2.87)	(2.16)	(2.91)	
<i>U</i>	268.5	135.0	113.5*	247.5	
<i>p</i>	.063	.079	.019	.199	

p* < .05 *p* < .01

表 4 Essay 2 の LFP 平均値

Essay 2	1st 1000	2nd 1000	AWL	NiL	
Lower Level Mean (n=20)	82.77	6.69	4.90	5.64	(%)
(SD)	(4.82)	(2.60)	(1.71)	(2.09)	
Upper Level Mean (n=20)	80.23	7.40	7.20	5.17	(%)
(SD)	(4.36)	(2.18)	(2.61)	(2.03)	
<i>U</i>	267.0	178.0	96.0**	229.0	
<i>p</i>	.070	.552	.005	.433	

p* < .05 *p* < .01

5. 考察

ここでは、本研究で設定した2つのリサーチ・クエスチョンに分けて考察を行うこととする。

5.1 トピックの LFP に対する影響

LFP の変動の有無に関しては、上位学習者と下位学習者で同じ結果を示した。両グループ共に、1st 1000 語と 2nd 1000 語に関しては、トピックが異なっても、それぞれのレベルの語彙の使用割合には有意差が認められず、LFP の値は定常的であるといえる。一方、AWL レベルの語と NiL の語に関しては、Essay 1 と Essay 2 の間で、使用割合に有意差が認められた。このことから、2000 語レベルまでの発表語彙は比較的安定しているが、「Beyond 2000」のレベルになると、

LFPはトピックの影響を受けて変動しやすい傾向があることを示唆している。Laufer and Nation (1995)の先行研究では、proficiencyの高い学習者は、2nd 1000とNiLのレベルのLFPは安定しており、それ以外のレベルは変動しやすいとされているが、本研究においても2nd 1000語レベルのLFPが安定してAWLレベルで変動が認められたことは、このことを一部裏付けるものとなっている。また、proficiencyの低い学習者のLFPはトピックによって変動しにくいという結果も報告されているが、今回、下位群においてもAWLレベルとNiLレベルの語彙のLFPの変動があったことは、限定的ながら辞書使用の効果があったと考えられることと、後述するように、エッセイで使用された日常語彙がNiLレベルに判定されたためと考えられる。Essay 1は比較的日常生活に関するテーマのエッセイであったため、自らの考えを述べる形のEssay 2より日常語彙の使用割合が多く、そのためNiLレベルの語彙の使用割合が高くなったといえる。AWLレベルのLFPの値からもEssay 2の方が難しいトピックであったことが分かる。

5.2 英語力の差とLFPの関係

学習者の英語力の差とLFPの関係に関しては、Essay 1とEssay 2の両方で、下位群の方が1st 1000語とNiLの語をより多く使用し、上位群が2nd 1000語をより多く使用する傾向が見られたが、どれもLFPの値には統計的な有意差は見られなかった。しかしながら、AWLレベルの語に関しては、どちらのエッセイにおいても、上位群の学習者の方がこのレベルの語を有意により多く使用していることが確認された。このような傾向については、条件は異なるものの、Smith (2004)でも同様の結果が報告されている。

Laufer & Nationの先行研究では、1st 1000語に関してはproficiencyの低いグループの学習者の方が使用割合が多いとされている。本研究においても、有意差は無かったものの、その傾向があることが示されており、下位群の学習者の方が平易な単語を使用する割合が高くなる傾向が見られる。また、2nd 1000語に関してはグループ間での使用頻度に有意差は無いとされているが、これに関しても同じ結果が得られた。これについては、今回のエッセイライティング課題では辞書使用が許可されていたため、下位群の学習者が和英辞書を使用して、このレベルの発表語彙を増やしたことにより、両者の差が無くなった可能性があると考えられる。そのため、今後、辞書使用を許可しない形で同様のエッセイライティングの課題を与えて分析を行う必要がある。AWLに関しては、proficiencyの低いグループの方が使用割合が低いとされているが、本研究でも、このレベルでは上位群と下位群の間のLFPには有意差が認められた。今回のエッセイで使用された語彙のうち、AWLレベルと判定された語彙をリストアップして調査したところ、上位群でも下位群でも高頻度で使用される語彙には共通性が見られたが、上位群の方が、使用した語彙の種類がより豊富であった。このことは、たとえ同じように辞書の使用を許可されたとしても、下位学習者は、このレベルの語彙を、上位学習者ほどには発表語彙として十分に使いこなすことができていることを示している。言い換えると、2nd 1000までの語彙に関しては、辞書使用によって上位学習者に近い割合の語彙を使用することができたが、その辞書使用の限界がAWLの語彙レベルにあるといえ、辞書が自由に使用できる環境に於いても学習者の英語力の差が「Beyond 2000」のレベルの語彙の使用割合に現れる可能性を示唆するものとなっている。NiLレベルの語彙の使用割合に関しては、2つのエッセイの両方に於いて群間にはLFP

の有意差が見られなかった。これは、前述したように、NiL の語の使用割合はトピックの影響があることと、LFP の分析に用いた語彙リスト自体の問題があると考えられる。今回の分析で使用した語彙リストは、1st 1000 語及び 2nd 1000 語は GSL がベースとなっており、AWL は学術用語のリストであるため、これらのリストには、*hobby, movie, baseball, guitar* 等の、今回のエッセイで学習者が比較的良好に使用した日常語が含まれておらず、結果的にこれらの語は NiL レベルと判定されてしまっている。NiL と判定された語のリストを分析すると、上位群、下位群共に AWL を超えるような低頻度語よりも、このような日常語が多く含まれていることが判明した。そのため、このレベルでは上位群と下位群との間に差が見られなかったものと判断される。従って、今後、BNC のワードリストをベースにした語彙リストに基づいて LFP を測定すれば、AWL レベルを超える語の使用割合に於いても、上位群と下位群との間に差が見られる可能性があると考えられる。

今回行った分析結果をまとめると、GSL と AWL のワードリストを使用する場合には、トピックの影響は AWL と NiL のレベルの LFP に現れ、それ以外のレベルは比較的安定することが判明した。また、学習者の英語力の差は AWL レベルの LFP に現れることが示された。これらのことから、今回の調査結果を見る限り、LFP の値はライティングにおける使用語彙から学習者の英語力を判定する指標としては妥当なものであると言える可能性が示された。

おわりに

本研究では、収集したエッセイの種類や被験者の数が十分であったとはいえないが、今回の調査対象となった intermediate レベルの上位層と下位層の学習者の英語力の差は、エッセイライティングにおいては、AWL レベルの語彙の使用割合の差として表れることが示され、たとえば同じように辞書使用を許可されたとしても、上位群の学習者と下位群の学習者との間には、発表語彙としてエッセイに使用できる語彙に依然として差が存在することが判明した。言い換えると、上位群の学習者は、アカデミックレベルの語彙をより多く使ってエッセイが書けるが、下位群の学習者は、たとえ辞書使用を許可されたとしても、それを効果的に利用して自分が本来持っている発表語彙より高いレベルの語彙をエッセイライティングに使用することに限界があり、和英辞書の使用能力自体に差があると考えられる。そのため、ライティングの授業においては、学習者が和英辞書を適切に使用することのできる能力を養成するような辞書指導が必要であると言える。現状では、英和辞書の辞書指導ですら十分に行われているとは言いが、和英辞書の辞書指導については、更に不十分であるという実態を考えると、日本人英語学習者のライティング力を高めるには、和英辞書の辞書指導を十分に行い、アカデミックな語彙や低頻度語を発表語彙として適切に使用できるような指導も必要であるといえる。

今回の研究結果から、LFP には、エッセイにおける使用語彙を用いた評価基準としての一定程度の妥当性があると言えるが、ライティングにおける使用語彙の評価に関しては、このような計量的、統計的な評価方法だけでなく、教師による holistic な評価と併用する必要があることは言うまでもないことであり、今後、より多くの学習者のエッセイデータを集めて、両者の関連についての研究を行うことが必要であり、多角的な面から LFP の妥当性を検証する必要がある。

るといえる。また、今回のような single word の語彙使用だけでなく、使用されたフレーズや構文にも上位群と下位群の差があると考えられるので、今後これらの調査を行って、使用語彙との関係を分析する予定である。

注

- 1) *Range* は 2009 年 8 月現在、Victoria University of Wellington の Paul Nation 教授のホームページからダウンロード可能となっている (<http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx>)。

参考文献

- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy* (2nd ed.). White Plains, NY: Longman.
- Coxhead, A. (1998). *An academic word list*. English Language Institute Occasional Publication No. 18. Wellington, NZ: School of Linguistics and Applied Language Studies, Victoria University of Wellington.
- Goodfellow, R., Lamy, M., & Jones, G. (2002) Assessing learners' writing using lexical frequency. *ReCALL*, 14 (1), 133-145.
- Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal* 25 (2), 21-33.
- Laufer, B. (2005). Lexical frequency profiles: From Monte Carlo to the real world: A response to Meara. *Applied Linguistics* 26 (4), 581-587.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics* 16 (3), 307-322.
- Meara, P. (2005). Lexical frequency profiles: A Monte Carlo analysis. *Applied Linguistics*, 26 (1), 32-47.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Smith, R. (2004). The lexical frequency profile: Problems and uses. *JALT2004 at Nara Conference Proceedings*. The Japan Association for Language Teaching.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- West, M. (1953). *A general service list of English words*. London: Longman, Green.