

「男はつらいよ」の英語字幕と英語検定教科書の語彙比較

古樋直己*

The Comparison of Vocabulary Between English Subtitles of a Famous Japanese Movie *Tora-san* And Authorized English Textbooks

Naoki FURUHI

This paper mainly examines the distribution, levels, and uniqueness of the vocabulary in the English subtitles of a famous Japanese movie called *Tora-san --- A Lovable Tramp ---*, and authorized English textbooks published in Japan in order to find a effective way of understanding English subtitles of Japanese movies and getting accustomed to the exposure to English. The findings are as follows. First, the words in the subtitles of *Tora-san* cover a narrower range than the textbooks; therefore, learners do not have to memorize wider range of vocabulary. Second, the words in the movie mainly consist of GSL1 and 2 levels and contain much fewer difficult words, thus, students should chiefly study rather easier words. Finally, the lexicons in the subtitles have many atypical ones, therefore, learners need to pay more attention to such words as appearing only in the movie.

Key words: vocabulary, word, subtitles, Japanese movie, *Tora-san*, corpus

1. はじめに

本研究では、日本映画の英語字幕に用いられている語彙と、高等学校用の文部科学省英語検定教科書に出現する語彙を様々な面から比較する。

この研究を実施する背景は下記の通りである。英語力は語彙力で決まる。これは、英語力は読解力と密接な関係があり、その読解力と深い関係を持つのが語彙力であるからである。読解力は、単語認識力、単語認識速度、語彙力、統語処理力、文意味処理力、ワーキングメモリ能力からなり、さらに、語彙力が読解力を構成する諸要因の中でも重要な位置を占める(門田, 2003, p.134)。また、Nation(2001, p.144), Nassaji(2006)においても、語彙力の重要性が指摘されている。英語力は語彙知識により大きな影響を受けているのである。

語彙力を向上させるにはどのような方法があるのだろうか。未知語に遭遇した際には辞書で意味を確かめたりする意図的学習、英文を読み、背景知識や前後のコンテキストから推測する偶発的学習などの

方法がある。どちらの方法に重点を置くかはともかく、両者を適切な組み合わせで実践していくのが語彙力向上へとつながる。膨大な語彙を習得するためには意図的学習だけでは不十分であり、偶発的語彙習得も併用する必要がある。

また、一般的な紙媒体のテキストを用いる学習と、音声や映像を用いた学習もある。古樋(2009)は、日本映画を日本語音声、英語字幕で使用することによって偶発的に起こる語彙習得について検証した。結果として、どのような英語力の実験協力者についても、偶発的な語彙習得が起こっていることを、統計的手法を採用し、立証している。他にも、植松(2004)でも、映画の英語字幕の英語学習への応用に効果があるとされている。

このように、映画の英語字幕を活用した方法は、語彙習得の手段のひとつとして十分な可能性を秘めたものである。

さて、映画の英語字幕が語彙学習に活用できることは明らかとなったが、はたして、英語字幕に出現する語彙と、英語の検定教科書に登場する語彙とは、どの程度一致するのか確認することを考えた。

本来は、教科書の英語を学習していけば、英語字幕も不自由しない程度に理解可能かどうかを確認するべきではある。しかし、これでは、理解している

原稿受付 平成 22 年 8 月 31 日

* 一般科目

とはどのようなことであるか明確に定義することが困難である。よって、本研究では、英語字幕と検定教科書に使用されている語彙が、それぞれどのようなものであるか検証するにとどめる。

2. 先行研究

ここでは、語彙習得、映画を用いた語彙習得に関する先行研究について概観する。

語彙はどの程度の数の習得を目標とすればよいのだろうか。英語を第1言語とする教養ある大人の場合 2 万ワードファミリー程度の語彙を知っている(門田, 2003, p.266)。見出し語換算では、この 1.6 から 2 倍となる(門田, 2003, p.36)ので 3 万から 4 万語に相当する。この語彙数を習得することは現実的ではないかもしれない。

実際、中学校学習指導要領(文部科学省, 1999a)、高等学校学習指導要領(文部科学省, 1999b)によると、中学校で 900 語程度、高等学校では、英語 I、英語 II、リーディングでそれぞれ、400 語、500 語、900 語程度習得することになっており、高校卒業までに合計で 2700 語程度学ぶ。教養あるネイティブスピーカーの語彙数とは一桁違う。

さて、語彙習得には様々な学習法がある。語彙習得の研究は多様である。たとえば、絵を用いた研究(Barcroft, 2002; Barcroft, 2003)や、テストなどの何らかの課題を与えた場合の語彙習得について研究(Barcroft, 2007; De Goot, 2006; Fukkink & de Gloppe, 2001)、テキストに飾り付けなどを施し、語彙を強調すると呼ばれる方法について検証したもの(Lee, 2007; Min, 2008; Rott, 2007)、キーワード法を用いたもの(Sagara & Alba, 2006; Rodríguez & Sadoki, 2000)、会話による言葉のやりとりによる語彙習得を研究したもの(Brouwer, 2003)、思考過程を声に出すよう実験協力者に依頼した研究(Daalen-Kapteijns et al., 2001)、第1言語と第2言語が1対1対応している場合と、第1言語ひとつに対して訳語が2つ存在する場合での反応速度を見ている研究(Jiang, 2004)、綴りと音に注目した研究(Chikamatsu(2006); Schiff & Calif, 2007)、テキストを読むときの実験協力者のスキルに重点を置いたもの(Fukkink et al. 2005; Nassaji, 2003)、語彙知識の深さをテーマとしたもの(Gredanus & Nienhui, 2001; Nassaji, 2006)などが挙げられる。

また、植松(2004)、古樋(2009)などのように、映画を用いた語彙習得に関する研究もある。植松では洋画の英語字幕、英語音声を活用しているが、古樋では日本映画の英語字幕、日本語音声を活用している。日本語音声の採用は、物語の理解しやすさを優先し

た結果である。また、日本映画を採用した理由は、親しみのある文化、風景、俳優が登場するからである。

3. 目的

日本映画の英語字幕は、検定教科書に用いられている語彙とどのように異なるのかを、調査するため、下記のリサーチクエスチョンを設定し、具体的に調べることにする。

- (1) 語彙の多様性に差は存在するのか。また、その差はどのようなものか。
- (2) 語彙の難度に差は存在するのか。また、その差はどのようなものか。
- (3) 語彙の特殊性に差は存在するのか。また、その差はどのようなものか。

4. 素材と方法

リサーチクエスチョンに挙げた分析を実行するため、はじめに、必要なデータを電子ファイルで準備する。

日本映画には「男はつらいよ」の第1作(山田洋次, 2008)を使用した。この作品を採用した理由は、第1に英語字幕が準備されていることである。また、今後、素材を増やし、研究の制度を向上させることを考えると、様々なジャンルの日本映画を採用するよりも、同じ作品の続編を用いたほうがよい。この点から言えば、本作品は 48 作が市場に出ており、英語字幕の分量を多く得るには最適である。ただし、本研究では、後述するように英語字幕を電子化するのに手間取り、第1作のみとしている。

なお、日本映画の英語字幕について分析するための台詞の電子化は下記のように行った。フリーソフトの DVDDecypter と SubRip で文字化けが多いながらも台詞をテキストファイル化した。その後、映画を英語字幕を表示させながら再生して、手作業で修正している。

英語検定教科書は、最近3年以内発行の英語 I 用、英語 II 用をそれぞれ 5 種類準備した。英語 I は *Vivid I*, *Exceed I*, *Tomorrow I*, *Powwow I*, *Prominence I*、英語 II は *Vivid II*, *Exceed II*, *Tomorrow II*, *Powwow II*, *Prominence II* である。これらの教科書に決定したのは、手元に英文のテキストファイルが収録されている CD-ROM があるからである。しかし、すべて、文部科学省の検定基準に沿ったものであり、使用語彙レベルなどは、一定の水準に収まっていると考えられる。

他に使用した素材として、石川(2008, pp. 166-167)の日本の英語教科書(英語 I, オーラルコミュニケーション I のもの), 韓国の英語教科書, 代表的なコーパスである FLOB と FRONW, 英米映画の各データも参考として用いている。

なお, リサーチクエスチョンに挙げている語彙の多様性, 難度, 特殊性の分析方法は, 石川(2008, pp. 165-167)を採用している。概要を述べる。延べ語数(tokens), 異なり語数(types), レマ語数(lemmas)をコンコーダンスソフトウェア Antconc で求める。さらに, 表計算ソフトウェアにて Guiraud 値も求めている。Guiraud 値は下記の式(石川, p. 77)で算出される。

$$\text{Guiraud} = \frac{\text{types}}{\sqrt{\text{tokens}}}$$

この値は, 語彙の多様性を表す。異なり語数を延べ語数で除した TTR(Type Token Ratio)では, 延べ語数の増加に伴い, 自然に低下するので, コーパスのサイズの差を考慮した Guiraud 値が存在する。この値は, 語彙多様性の指標中, 最も安定した結果を返すとされている。

5. 結 果

5-1 語彙の多様性の比較

表 1 には語彙の多様性の比較を行うために, 教科書と映画英語字幕について, 延べ語数, 異なり語数, レマ, Guiraud 値を示した。

表 1 語彙の多様性の比較 1

	英語 I5 種類す べて	英語 II5 種類す べて	英語 I お よび II す べて(各 5 種類, 合計 10 冊)	男は つら いよ
延べ語 数	30778	41503	72281	5648
異なり 語数	3413	4596	5868	1132
レマ	2599	3458	4348	993
Guiraud	19.5	22.6	21.8	15.1

Guiraud 値は, 映画のものが他よりもかなり低いことがわかる。たとえば, 教科書 10 冊の値の 7 割に満たない。この映画に用いられている英語字幕の語彙は, 同じ語彙の繰り返しが多く, 多様性に乏しいことがわかる。また, 英語 I の教科書と英語 II の教科書では, やはり英語 I のほうが英語 II よりも低い

値となっており, 学年が進むにつれて, 語彙の多様性が大きくなっていることがわかる。

つぎに, 表 2 には, 石川(2008, p.171)の情報を元に作成したものである。なお, 表 2 の「映画」とは, 石川が収集した英米の映画の台詞データコーパスである。

表 2 語彙の多様性の比較 2

	韓国教 科書(高 1 用 4 種 類)	FLOB	FRONW	映画	男は つら いよ
延べ語 数	185547	1199781	1206555	1396429	5648
異なり 語数	11111	43134	43222	35454	1132
レマ	9247	33181	33267	27361	993
Guiraud	25.8	39.4	39.3	30.0	15.1

表 2 においても, 本研究で取り上げている日本映画の英語字幕に用いられる語彙の多様性は低いことがわかる。また, 同時に, 日本の教科書の語彙の多様性が低いことも理解できる。

このように, 「男はつらいよ」の英語字幕の使用語彙は, ここで比べた他の素材よりも, 多様性が低く, 語彙の繰り返し使用が多いことがわかった。つまり, 少ない語彙を習得できれば, 理解可能である。

5-2 語彙の難度の比較

表 3 には語彙の難度を比較するために, 各素材の語彙が, どの程度の難易度の語彙で構成されているのかをパーセントで示した。

語彙の難易度は, GSL1 から GSL3 に分類している。GSL1 に分類されるものが最もやさしく, 数字が大きくなるにつれて難度が増す。この分析には, Range(<http://www.victoria.ac.nz/lals/resources/range.aspx>)というフリーソフトウェアを用いた。

表 3 からわかるように, GSL1, GSL2, GSL1 から 3 以外のその他には, 各素材に大きな違いはない。しかし, GSL3 には日本映画と他の素材で 4 倍から 6 倍の差がある。

これは, 日本映画には難度の高い語彙が非常に少ないことを意味する。少なくとも今回取り上げた日本映画の英語字幕を理解するには, 難度の高い語彙の使用頻度が低いことを念頭に置いておけばよいことがわかる。

表3 語彙の難度の比較 1

	英語 I5 種類す べて	英語 II5 種類す べて	英語 Iおよ び IIすべ て(各 5 種 類, 合計 10 冊)	男は つら いよ
GSL1[%]	84.9	82.7	83.7	83.5
GSL2[%]	5.2	5.8	5.6	7.1
GSL3[%]	1.2	1.8	1.5	0.3
その他[%]	8.7	9.7	9.3	9.2

表4 語彙の難度の比較 2

	日本 教科 書(英 語 Iと OCI 各 4 種類)	韓国教 科書(高 1 用 4 種 類)	FLOB	FROWN	映 画	男は つら いよ
GSL1	82.3	78.5	68.0	66.5	82.2	83.5
GSL2	6.3	5.8	4.7	4.6	5.0	7.1
GSL3	1.3	1.6	5.7	6.4	1.3	0.3
その他	10.1	14.1	21.6	22.5	10.5	9.2

つぎに、表4に表2と同じ素材の場合を挙げた。

やはり、今回取り上げた「男はつらいよ」の英語字幕のGSL3の低さは際立っている。同じ映画でも、英米の映画の23[%]に留まっている。

このように、今回の日本映画の語彙の難度は、他の素材と比較して、極めて低い。つまり、GSL1と2という基本的な語彙を習得していれば、十分に対応可能なことがわかった。

5-3 語彙の特殊性の比較

最後に、語彙の特殊性について確認する。特殊性とは、他の素材には出現しないが、当該素材には出現する度合いがどの程度かと定義する。

この分析は、筆者の手元にある表1や表3で扱った素材のみについて行った。表2や表4の素材は詳細が不明で入手も困難なためである。

分析には Web 上で使用する Text-based Range(http://www.lex tutor.ca/range/range_text/)というソフトウェアを用いた。この結果を表計算ソフトにて加工し、つぎの結果を得た。

まず、英語 I、英語 II の教科書、「男はつらいよ」の英語字幕のいずれか一つにでも出現するレマは3,030 個であった。レマでこれだけの個数であるから、実際の語彙はより多様であるはずである。

また、英語 I 教科書に出現するのが1,905 レマ、英語 II 教科書が2,421 レマ、「男はつらいよ」が772 レマである。なお、これらの数値は表1のものとは異なるが、ソフトウェアの違いによると思われる。

これらの数値の大小を検討するため、その素材内の語数との割合を算出する。もちろん、レマと語数(tokens)は異なった指標であるが、目安にはなる。その結果は、英語 I が6.2[%]、英語 II が5.8[%]、「男はつらいよ」が13.7[%]となる。この結果から、「男はつらいよ」には、総語数の割にレマの数は多い傾向にありそうなことになる。しかし、Guiraud 値が生まれたのと同じ考えでいくと、そもそも延べ語数に差がありすぎるので、必ずしも断言はできない。

なお、日本映画の772 レマ中、英語の検定教科書には出現しないのが149 レマ(付録)と20[%]程度を占めている。

つぎに、表5にその「男はつらいよ」独自のレマの難度を分析したものを示す。

表5 「男はつらいよ」
独自レマの難度

GSL1[%]	4.0
GSL2[%]	22.8
GSL3[%]	4.0
その他[%]	69.1

この表から、独自のレマは、GSL2 および GSL1 から3以外のものが多いことがわかる。

以上、5-1 から5-3 まで、語彙の多様性、難度、特殊性を調べてきた。まず、多様性については、今回の日本映画の英語字幕については、語彙の多様な範囲には分布しておらず、狭い範囲に集中していることが明らかになった。難度については、GSL1 および2の範囲が多く、GSL3 が少ない状態であり、比較的やさしい語彙が多く使用されていることがわかった。最後に、特殊性については、独自なレマについて、GSL のリストにないものが7割程度を占めることがわかった。

6. 考 察

では、リサーチクエスチョンに答えながら、結果について考えてみる。

まず、リサーチクエスチョン(1)の多様性であるが、各素材間での多様性の差は存在する。ただし、数倍の開きとまでは行かない。しかし、英語の検定教科書の語彙と比較して、狭い範囲に分布しているのは確かであるので、この日本映画の英語字幕を理解し

ようと思えば、教科書の語彙を理解するのよりも、狭い範囲の語彙を習得すれば、目的を果たせる可能性が高い。

つぎに、リサーチクエスチョン(2)の難度の差に進む。GSL3 レベルの語彙が日本映画には極端に少ない。検定教科書の場合の 20[%]前後に過ぎない。教科書ではある程度難度の高い語彙が出現するが、今回の日本映画にははるかに少ない割合でしか出現しない。

もちろん、英語 I および英語 II はオーラルコミュニケーションのみを対象としているわけではないので、会話表現が多い映画の字幕との比較で、一概には言えないかもしれない。しかし、表 4 にある英語 I とオーラルコミュニケーション I の教科書の場合でも、GSL3 の占める割合は 1.3[%]と、やはり「男はつらいよ」の 4 倍以上ある。

やはり、「男はつらいよ」の英語字幕には難度の高い語彙が少ない。この英語字幕を理解するのに GSL3 レベルの語彙にあまり注力する必要はない。

リサーチクエスチョン(3)の特殊性であるが、この映画の 2 割程度は検定教科書 10 冊に出現しないのである。この事実を踏まえれば、教科書に出てくる語彙を習得すればよいということにならないことがわかる。

7. ま と め

3 つのリサーチクエスチョンに回答していくなかで、「男はつらいよ」の英語字幕の理解に必要なことはつぎの 3 点である。狭い範囲のよく出現する語彙を習得するように心掛け、GSL3 レベルの語彙にはそれほど注力する必要はなく、独自の語彙に力を入れることが重要である。

とくに最初に書いた狭い範囲の語彙の繰り返し使用があることがわかった。6 回、16 回、あるいは 20 回と様々な研究結果が紹介されている(Zahar et al., 2001)が、複数回出現する語彙には偶発的な習得も期待可能である。また、映画は好みの作品であれば、2 時間程度で 1 回楽しむことができ、繰り返し見るのも苦にならない場合も多い。その結果、映画は偶発的な語彙習得も起こりやすい(古樋, 2009)。

つぎに、本研究の限界を明らかにしておく。まず、分析対象の日本映画が 1 本に過ぎず、延べ語数が他の素材よりもはるかに少ない点が問題である。

分析本数を増やすには、先に記したようになかなか大変な作業ではあるが、取り組んでいきたい。

また、今後は、外国映画についても、同様の分析を試みてみたい。

本研究を実際の教育の場にどのように応用できる

のかについても検討し、映画の英語字幕を効率的に活用し、容易に語彙習得が行えるような方法を創造していきたい。

* 本稿は、映画英語教育学会第 15 回全国大会(2009 年 6 月 20 日於 専修大学生田校舎)で行った口頭発表に加筆・修正を加えたものである。

参 考 文 献

- Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, 52(2), 323-363.
- Barcroft, J. (2003). Effects of questions about word meaning during L2 spanish lexical learning. *The Modern Language Journal*, 87(4), 546-561.
- Brouwer, C. (2003). Word searches in NNS-NS interaction: Opportunities for language learning?. *The Modern Language Journal*, 87(4), 534-545.
- Chikamatsu, N. (2006). Developmental word recognition: A study of L1 English readers of L2 Japanese. *The Modern Language Journal*, 90(1), 67-85.
- Daalen-Kapteijns, M. van, Elshout-Mohr, M., & de Gropper, K. (2001). Deriving the meaning of unknown words from multiple contexts. *Language Learning*, 51(1), 145-181.
- De Groot, A. M. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language Learning*, 56(3), 463-506.
- Fukink, R. G., Blok, H., & de Gropper, K. (2001). Deriving word meaning from written context: A multicomponential skill. *Language Learning*, 51(3), 477-496.
- Fukink, R. G., Hulstijn, J., & Simis, A. (2005). Does training in second-language word recognition skills affect reading comprehension? An experimental study. *The Modern Language Journal*, 89(1), 54-75.
- 古樋直己. (2009). 偶発的語彙習得と英語力・語の頻度・コンテキストの関係 —英語字幕付き邦画を用いた場合—, 『英語教育研究』, 14, 29-40.
- 石川慎一郎. (2008). 『英語コーパスと言語教育』, 大修館書店.
- Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *The Modern Language Journal*, 88(3), 416-432.
- 門田修平(編著). (2003), 『英語のメンタルレキシコン』, 松柏社.
- Lee, S. (2007). Effects of Textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning*, 57(1), 87-118.
- Min, H. (2008). EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning*, 58(1), 73-115.
- 文部科学省. (1999a). : 中学校学習指導要領第 9 節外国語,

- http://www.mext.go.jp/b_menu/shuppan/sonota/990301/03122602/010.htm
- 文部科学省. (1999b). : 高等学校学習指導要領第 8 節外国語,
http://www.mext.go.jp/b_menu/shuppan/sonota/990301/03122603/009.htm
- Nassaji, H. (2003). Higher-level and lower level text processing skills in advanced ESL reading comprehension. *The Modern Language Journal*, 87(2), 261-276.
- Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learner's lexical inferencing strategy use and success. *The Modern Language Journal*, 90(3), 387-401.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*, Cambridge University Press.
- Rodríguez, M., & Sadoki, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classrooms. *Language Learning*, 50(2), 385-412.
- Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning*, 57(2), 165-199
- Sagara, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *The Modern Language Journal*, 90(2), 228-243.
- 植松茂男. (2004). DVD 映画教材利用時の英語字幕が英語学習に与える影響について, メディア教育研究, 1, 107-114.
- 山田洋次. (2008). 「男はつらいよ HD リマスター版」, 松竹.
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review*, 15, 541-572.
- * 参考文献の記載様式は, 英語教育関係学会で標準とされる「APA 形式」による.

付録 「男はつらいよ」英語字幕のみに現れる語彙(レマ)149 個

abode	dowry	hypnotise	rubbish
acquaint	dragon	indebted	sacrifice
adequate	drawback	inn	salesman
apology	due	insist	sandal
arrow	dummy	insult	scum
ashamed	dump	kid	sentiment
bankrupt	dung	lame	shack
bargain	dust	loaf	shall
beer	economy	lousy	sir
bet	edit	lout	sire
birthplace	electron	mad	smart
blossom	electronic	manure	sneeze
boss	elegant	meddle	sob
brat	employ	mesmerise	sort
buddy	enterprise	neon	spoil
bug	errand	nut	stain
bump	expel	octopus	stairs
bust	fart	odd	steak
butter	fiancee	oriental	stove
catnip	formal	overtime	stuff
chase	fur	pamper	suicide
choke	gangster	pardon	suitor
complaints	gape	peddle	surname
conscience	geisha	personnel	swear
corpse	ghost	phony	tag
crap	glance	pierce	tax
creak	gory	pillow	textile
damn	groom	plump	tofu
dank	guest	poke	tonight
dare	gulp	praise	tramp
delinquent	guy	publication	ugly
deliver	hag	pun	urine
demon	hangover	pup	wistful
depress	headache	quarrel	wit
desperate	honeymoon	rant	wrist
dim	hook	rave	
discount	hum	repent	
disown	humble	reshuffle	