

中学校用英語教科書に含まれる学習語彙の計量的分析 ——品詞別・ランク別プロフィールの作成——

伊 東 治 己・山 田 あゆみ

(奈良教育大学英語教室)

(平成3年4月23日受理)

本調査の主な目的は、中学校用英語教科書に含まれる英語学習語彙をパーソナル・コンピュータを使って分析し、その品詞別およびランク別プロフィールを明らかにすることである。コンピュータ入力の対象になった英語テキストは、次の通りである⁽¹⁾。

- 1) EE : *Everyday English* (中教出版)
- 2) NC : *New Crown* (三省堂)
- 3) NH : *New Horizon* (東京書籍)
- 4) OW : *One World* (教育出版)
- 5) SS : *Sunshine* (開隆堂)
- 6) TE : *Total English* (秀文出版)
- 7) HI : *Elementary, Intermediate and Advanced Stories for Reproduction : American Series*, by L. A. Hill (Oxford University Press)

以上7種類の英語テキストのうち、最初の6種類は平成2年度版中学校用英語教科書で、最後はその比較対照用のテキストである。また、品詞別およびランク別プロフィール作成の対象になったのは、中学校用英語教科書のNC、NH、SSと比較用テキストのHIの4種類計12冊である。

1. 先行研究について

外国語教育における学習語彙の研究には、大きく二つの流れが存在している。ひとつは、学習者に「どんな語彙を教えるべきか」という問いかけに対する答を追及する方向であり、もうひとつは、学習者に「どんな語彙が教えられているか」という問いかけに対する答を追及する方向である。前者においては、既存の各種語彙統計の結果を参考にしたり、あるいは、独自に目標言語のサンプルを広範囲に集め、その中に含まれる語彙を頻度、範囲、有用度、適用度、学習容易度(McKey 1965: 176-190)など様々な観点から分析し、学習者にとって重要と思われる基本語を選定していくという研究が主流になっている。この方向の研究としては、H. E. Palmer (1931)の*Second Interim Report on Vocabulary Selection*いわゆる「パーマーの3千語」や、M. West (1936)の約2千語からなる*A General Service List of English Words*、あるいは、C. K. Ogden (1968)の850語からなる*Basic English*が代表的である。日本人研究者によるものとしては、1300語の中学校基本語彙と4800語の高等学校基本語彙を選定した全英連(1963; 1981)、約4千語の基本語彙を選定したJACET(1983)、718語からなる口語英語の基本語彙を選定した清川(1987)のほか、多くの研究がある。中学校学習指導要領に見られる必修語リストもこの流れの中に位置付けられるであろう。

後者、つまり、「どんな語彙が教えられているか」という問いかけに対する答を追及する方向

においては、教科書を中心とした教材に含まれている語彙の実態を、異語数・総語数などの観点から明らかにしたり、自国の教科書に含まれる語彙と諸外国の教科書に含まれる語彙を比較・検討することによって、自国の教科書に含まれる語彙の性格を明らかにして行こうとする研究が主流になっている。この方向での我が国における研究としては、当時の中学校用12種と高校用25種に含まれる語彙について、その使用頻度と重要度を示した速川（1966）の研究が草分け的存在になっている。比較的最近のものとしては、中学校用5種と、高校用27種の英語教科書の中での語彙の使用状況を調査した淀縄（1983）、中学校用教科書、高校用教科書、大学教養課程用テキスト、大学専門課程用テキスト、および英字新聞に含まれる語彙を、姿語総頻数、異姿語数、相対新出異姿語数とその比率、新語密度、総語数と異姿語数の比率など、多角的な観点から比較・検討した金田（1981）、教科書改訂の度に、中学校用全種に含まれる語彙について、それぞれが使用されている用例をすべて文単位でコンコーダンス風にまとめている垣田（1977）、西ドイツ・フランス・ソ連の中学校用英語教科書に含まれている語彙を比較分析した教科書研究センター（1984）、単語の使用頻度と分布という観点から日本とソ連の中学校用英語教科書を比較した縫部（1984；1985）、単語の使用頻度に加えて、食事／飲食・乗物／輸送・職業／身分といった分野別の単語の使用状況という観点から日本とフランスの中学校用英語教科書に含まれる語彙を比較した三浦（1985）、などの研究が目にとまった。

今回筆者達が行った語彙調査は、以上述べた学習語彙研究の二つの流れの中の後者、つまり、「どんな語彙が教えられているか」という問いかけに対する答を追及する研究の一環として位置付けられるものである。その主な目的は、既に述べたように、日本の中学校用英語教科書に含まれる語彙について、品詞別およびランク別に異語数や総語数を調べ、その品詞別・ランク別プロフィールを明らかにすることである。

2. PCQR による基礎データの作成

2.1 PCQR について

これは、筆者（伊東）の前任校である和歌山大学教育学部の英語教育研究室において、同大学の他の研究者の協力を得て開発された英文用例検索プログラムである（伊東 1989）。中学校用英語教科書に含まれる全ての単語について、その用例を検索することがその主要なねらいになっているが、英語データであれば、いかなるものでも用例検索のターゲットにすることができる極めて便利なプログラムである。プログラム自体はベーシック言語で書かれているが、MS-DOS 上で動き、かつ、対話形式になっているので、NEC 9801 シリーズのパソコンがあれば、簡単に利用できるようになっている。なお、このプログラムは一般に公開しているもので、利用希望者には無料でコピーしている。

2.2 英文データの入力

学習語彙の品詞別・ランク別プロフィール作成のための基礎データとして、まず、PCQR 中の「英文データ入力プログラム」を使って、各対象テキストに含まれる英文を一文ずつコンピュータに入力し、文データ・ファイルを作成した。使用したコンピュータは、NEC の PC 9801VX である。入力の対象になったのは、中学校用英語教科書の場合、各主要レッスンの本文とターゲット・センテンスで、練習問題やまとめに当たるもの、あるいは発展教材や応用教材のような副次的部分の英文は入力されなかった。比較教材の HI の場合は、各レッスンのストーリーの部分の英

文だけである。入力はすべて大文字で行われた。*thirty-five* のようにハイフンで繋がれた単語は、そのままでは *thirty* と *five* の二語として処理されるので、もし全体を一語として処理したい場合には、THIRTY'FIVE のように入力した。また *New York* のように、分かれてはいるが全体を一語として処理したい場合にも、NEW'YORK のように入力した。また、*I'm* や *don't* のような縮約形については、基本的には、I AM や DO NOT のように分割して入力した。ただし、*Let's sing a song.* の *let's* や *This is mine, isn't it?* の *isn't* などは、その性格上、分割せずそのままの形で入力した。

2.3 コンコーダンスの作成

各対象テキストに含まれている英文が文データ・ファイルの中に入力された段階で、PCQR の中の「コンコーダンス作成プログラム」を使って、各文データ・ファイルに含まれている異語をすべて抽出し、その使用回数を計算し、かつ、アルファベット順に並べ替える作業を行った。この結果、各対象テキストについて異語の使用頻度リストが単語データ・ファイルとして作成された。この頻度リストの中では、コンピュータ処理の性格上、意味が同じでも形の違うものはすべて異語として計算され、処理される。例えば、*desk* と *desks* や、*study* と *studies* などとはすべて異語として処理される。逆に、形容詞の *hard* と副詞の *hard* のような同音異義語は、同一の単語として処理されている。

各対象テキストについて、文データ・ファイルと単語データ・ファイルが作成された段階で、PCQR の中の「コンコーダンス検索プログラム」を実行すると、それぞれの異語について、当該テキストの中でのその用例がすべて検索できるようになる。具体的には、コンピュータのキーボードから任意の単語（例えば STUDY）を入力すると、当該テキスト（この場合は NH）の中でのその用例がすべて一文ずつ、コンコーダンスの形で学年と出現ページを伴って下のように表示・印刷される。このコンコーダンスの検索結果は、後の「検索用語彙データ」の入力の段階で、各異語についてその品詞を決定したり、*hard* のような同音異義語の場合に品詞ごとの頻度を入力する時などにおいて大いに活用されることになる。

検索している単語は STUDY です

total ==> 13

| | | | |
|----|---|----|---|
| 1 | 1 | 53 | DO YOU STUDY IT AT SCHOOL? |
| 2 | 2 | 38 | THEN YOU HAVE TO STUDY SCIENCE VERY HARD. |
| 3 | 2 | 38 | I HAVE TO STUDY SCIENCE. |
| 4 | 2 | 38 | MIKE HAS TO STUDY SCIENCE. |
| 5 | 2 | 39 | I AM GOING TO STUDY SCIENCE EVERY DAY, KATHY. |
| 6 | 2 | 39 | WHY DO YOU HAVE TO STUDY SCIENCE SO HARD? |
| 7 | 2 | 39 | YOU DO NOT HAVE TO STUDY SO HARD. |
| 8 | 2 | 39 | I HAVE TO STUDY SCIENCE NOW. |
| 9 | 2 | 39 | DO YOU HAVE TO STUDY SCIENCE NOW? |
| 10 | 2 | 39 | NO, I DO NOT (HAVE TO STUDY SCIENCE NOW). |
| 11 | 3 | 7 | I HAVE WANTED TO STUDY IT SINCE I LEARNED ABOUT IT IN JAPAN. |
| 12 | 3 | 59 | MY FATHER ONCE WENT TO SHANGHAI TO STUDY COMPUTER SCIENCE, BUT THE FIRST THING THAT HE HAD TO LEARN WAS THE DIALECT SPOKEN THERE. |
| 13 | 3 | 59 | I HAVE JUST BEGUN TO STUDY JAPANESE. |

3. 検索用語彙データの作成

3.1 語彙データの構成

本調査の主目的は、4種の対象テキストに含まれる語彙の品詞別・ランク別プロフィールを明らかにすることにある。そのために、対象テキストに含まれる各異語についてその品詞とランクおよび使用頻度を調べ、市販のデータ・ベース・ソフト（NINJA3/PRO）を使って次のような形でデータ化した。

以下、このデータの具体的な作成手順を紹介する。なお、この作業の対象になったテキストは、NC、NH、SS、HIの4種計12冊である。

| No. | 単語 | 品詞 | ランク | 頻度 |
|-----|-------------|----|-----|----------|
| 1 | A | T | 1 | 218 |
| 2 | ABLE | J | 1 | 3 |
| 3 | ABOUT | P | 1 | 33 |
| 4 | ABOUT | D | 1 | 13 |
| 5 | ABOVE | P | 2 | 1 |
| 6 | ACID | J | 10 | 7 |
| 7 | ADIRONDACK | N | 11 | 1 |
| 8 | ADIRONDACKS | N | 11 | 1 |
| 9 | AFRICA | N | 5 | 2 |
| 10 | AFTER | P | 1 | 7 |
| 11 | AFTER | C | 1 | 2 |
| 12 | AFTERNOON | N | 3 | 2 |
| 13 | AGAIN | D | 1 | 5 |
| 14 | AGAINST | P | 1 | 1 |
| 15 | AGO | D | 1 | 9 |
| 16 | AHEAD | D | 2 | 1 |
| 17 | AIR | N | 1 | 2 |
| 18 | AKIO | N | 11 | 1 |
| 19 | ALBUM | N | 11 | 1 |
| 20 | ALL | Q | 1 | 25 |
| 21 | ALMOST | D | 1 | 2 |
| 22 | ALONE | D | 2 | 2 |
| 23 | ALREADY | D | 1 | 1 |
| 24 | ALSO | D | 1 | 9 |
| 25 | ALTAMIRA | N | 11 | 3 (以下省略) |

3.2 品詞と頻度の入力

英語学習語彙の品詞別プロフィールを作成する上でまず第一に問題になるのが、品詞の分類法である。従来の学校文法では、英語の語彙は伝統的に8つの品詞に分類されて来たが、その一方で、その8品詞の分類の枠組みに納まりきれない単語が多く存在するのも事実である。さらには、伝統的な品詞分類の用語（例えば名詞や動詞）を使わずに品詞の類型を提示する試みもなされて

きている。本調査では、この点を考慮しつつ、かつ、現代アメリカ英語の用例を数多く收拾しているいわゆるブラウン・コーパス (Brown Corpus) に見られる品詞分類を参考にして (Francis and Kučera 1982)、次に示すような13の品詞を設定した。なお、各品詞についている記号はコンピュータ入力に際して利用した符号である。

| | | | | |
|--------|--------|--------|--------|---------|
| N: 名詞 | J: 形容詞 | D: 副詞 | V: 動詞 | B: Be動詞 |
| A: 助動詞 | R: 代名詞 | W: 疑問詞 | P: 前置詞 | C: 接続詞 |
| Q: 数量詞 | T: 冠詞 | M: その他 | | |

「動詞」に含まれるのは、いわゆる一般動詞のみである。*give in* の *in* などは「副詞」に含まれる。「助動詞」には、法助動詞の他、*have* や *do* の助動詞用法も含まれている。「数量詞」には、*ten* や *twenty* のような数詞の他、*many* や *some* や *few* のような数に関係のある単語も含まれている。最後の「その他」の中には、*O. K.* や *good-bye* のような挨拶ことばやいわゆる感嘆詞の部類が含まれている。

品詞と使用頻度の入力においては、PCQR によって作成された単語使用頻度リストを参照した。ただ、この頻度リストにおいては、それぞれの異語について品詞は明示されておらず、かつ、*hard* のような同音異義語の場合や *work* のように名詞としても動詞としても使用される単語については、品詞が未分化の状態での頻度が計算されているので、そのままの形では利用できない。品詞ごとの使用頻度を確定するためには、PCQR 中の「コンコーダンス検索プログラム」を使って問題となっている単語の用例を丹念に調べることが必要になってくる。例えば、NH における *work* のコンコーダンスは次のようになっている。

検索している単語は WORK です

total ==> 8

- | | | | |
|---|---|----|---|
| 1 | 2 | 50 | MY FATHER USUALLY GIVES ME THREE DOLLARS FOR THE WORK. |
| 2 | 2 | 50 | WELL, IN AMERICA MANY PARENTS PAY THEIR CHILDREN FOR WORK LIKE THAT. |
| 3 | 2 | 84 | ALL OF US SHOULD WORK TOGETHER TO STOP ACID RAIN. |
| 4 | 3 | 70 | I CAN AT LEAST DO SOME OF YOUR WORK," CHIYOKO SAID. |
| 5 | 3 | 70 | I WILL DO THE WORK BY MYSELF. |
| 6 | 3 | 71 | ITARU FOUND THAT HIS WORK WAS VERY DIFFICULT. |
| 7 | 3 | 71 | SOON CHIYOKO LEARNED TO DO A PART OF HER HUSBAND'S WORK AND BEGAN TO HELP HIM. |
| 8 | 3 | 73 | TODAY FEW PEOPLE KNOW ABOUT THE WORK WHICH MR. AND MRS. NONAKA DID ON MT. FUJI. |

これをもとに、NH での *work* の使用頻度 (8 例) は、名詞としての用法 (7 例) と動詞としての用法 (1 例) に分けて入力される。これと同じ作業が、各対象テキストの中で必要と思われるすべての単語に関して繰り返された。

3.3 ランクの入力

各異語のランクの入力に際しては、ブラウン・コーパス (Francis and Kučera 1982) のランク・リストを参照した。このランク・リスト(リストA) は、単語の使用頻度だけでなく、その使用範囲をも考慮して作成されたものであり、1 位から6000位までの単語がランク順に列記してある。さら

に特徴的なことは、次に示すように、独自の分類法に基づく品詞名が各単語について明記しており、いわゆる同音異義語の場合には、それぞれの品詞ごとにランクを示す数値が示してある点である。

〈リスト A〉

| 順位 | 単語 | 品詞 | 実頻度 | 調整頻度 |
|----|------|------------|-------|----------|
| 1 | the | article | 69975 | 69792.94 |
| 2 | be | verb | 39175 | 39109.95 |
| 3 | of | prep. | 36432 | 35786.01 |
| 4 | and | co. conj. | 28872 | 28821.11 |
| 5 | a | article | 23073 | 22984.95 |
| 6 | in | prep. | 20870 | 20685.17 |
| 7 | he | pers. pro. | 19427 | 17280.77 |
| 8 | to | inf. mark. | 15025 | 14990.82 |
| 9 | have | verb | 12458 | 12192.06 |
| 10 | to | prep. | 11165 | 11129.57 |
| 11 | it | pronoun | 10942 | 10836.51 |
| 12 | for | prep. | 8996 | 8899.55 |
| 13 | they | pers. pro. | 8284 | 8162.08 |
| 14 | with | prep. | 7286 | 7267.37 |
| 15 | I | pers. pro. | 8387 | 6885.48 |
| 16 | not | neg. adv. | 6976 | 6739.48 |
| 17 | that | sub. conj. | 6468 | 6373.68 |
| 18 | on | prep. | 6183 | 6151.18 |
| 19 | as | sub. conj. | 6029 | 5982.09 |
| 20 | at | prep. | 5377 | 5317.20 |
| 21 | by | prep. | 5246 | 5066.04 |
| 22 | this | sing. det. | 5145 | 5064.57 |
| 23 | we | pers. pro. | 4865 | 4699.87 |
| 24 | she | pers. pro. | 6039 | 4378.51 |
| 25 | from | prep. | 4371 | 4358.51 |

(以下省略)

〈リスト B〉

| 単語 | 順位 | 品詞 |
|------------|------|------------|
| a | 5 | article |
| a. m. | 2763 | adverb |
| abandon | 1967 | verb |
| abide | 5051 | verb |
| ability | 1170 | noun |
| able | 419 | adjective |
| aboard | 4561 | adverb |
| abolition | 5859 | noun |
| about | 71 | prep. |
| about | 147 | adv./part. |
| above | 501 | prep. |
| above | 4695 | adjective |
| above | 1383 | adverb |
| abroad | 2318 | adverb |
| abrupt | 3688 | adjective |
| abruptly | 5360 | adverb |
| absence | 1713 | noun |
| absent | 3283 | adjective |
| absolute | 3079 | adjective |
| absolutely | 4621 | qualifier |
| absolutely | 5572 | adverb |
| absorb | 2153 | verb |

(以下省略)

[備考] 実頻度とは、コーパス全体での総出現回数の中で、調整頻度とは、分野ごとの出現状況を考慮して算出された頻度のことである。単語のランク付けは後者の調整頻度を基準にして行われている。

ただ、このブラウン・コーパスのリストは、上記(リスト A)のように、ランク順に単語が列記してあるので、このままの形では本調査におけるランク入力の参考資料として使用するには不都合である。そこで、今回の調査では、上に示すように、ブラウン・コーパスのランク・リストをアルファベット順に並べ替えた「ブラウン・コーパス・アルファベット順ランク・リスト」(リスト B)を参考資料として利用した。この「アルファベット順ランク・リスト」は、前述の PCQR と同様、筆者(伊東)が和歌山大学教育学部の英語教育研究室において同大学の他の研究者の協力を得て作成したものである。

ランクの具体的数値の入力に際しては、ブラウン・コーパスの数値をそのまま利用せず、ランクを次のような段階に分けた上で、各段階の数値を入力した。

| | |
|------------------|---------------------|
| ランク 1 : 1位～ 500位 | ランク 7 : 3001位～3500位 |
| 2 : 501位～1000位 | 8 : 3501位～4000位 |
| 3 : 1001位～1500位 | 9 : 4001位～4500位 |
| 4 : 1501位～2000位 | 10 : 4501位～5000位 |
| 5 : 2001位～2500位 | 11 : 5001位以上 |
| 6 : 2501位～3000位 | |

4・データの分析

4.1 全体的傾向

データを分析するにあたってまず初めに、対象テキスト全体の大まかな傾向を調べた。下の表1は、平成2年度版中学校用英語教科書6種類（EE、NC、NH、OW、SS、TE）と、その比較対照テキストであるHIにおける総文数、総語数、異語数、総語数に対する異語数の割合（いわゆる Type-Token Ratio で以下 T/T と略す）⁽²⁾をまとめたものである。なおここで示す語数には英単語のほか数字なども含まれている。

表1 全般的語彙傾向

| 教科書 | EE | NC | NH | OW | SS | TE | HI | 平均 |
|-----|------|------|------|------|------|------|-------|------|
| 総文数 | 1468 | 1164 | 1223 | 1524 | 1362 | 1381 | 913 | 1354 |
| 総語数 | 9688 | 7327 | 7857 | 9159 | 8686 | 7894 | 13788 | 8438 |
| 異語数 | 1385 | 1152 | 1189 | 1285 | 1296 | 1130 | 1585 | 1240 |
| T/T | 7.0 | 6.4 | 6.3 | 7.1 | 6.7 | 7.0 | 8.7 | 6.8 |

上の表から次のような特徴が見られる。

① 7種類のテキストの中では、HIが総文数において最も少くなっているのに対して、総語数においては最も多くなっている。このことから、HIは長い文を多用する傾向にあると言える。

② 中学校用教科書を比較してみると、NCの総文数、総語数が最も少なくなっている。このことから、NCに含まれる言語材料は中学校用教科書の中で最も少なくなっているということが言える。逆に言語材料が比較的多く含まれているのは、EE、OWである。ただし、本研究においては入力の対象となったのは、主要レッスンの本文とターゲット・センテンスだけであり、副次的教材の部分は入力の対象外とされているので、一概に全テキストを同一の尺度で比較するのは問題かもしれない。なお、参考までに入力の対象になった主要レッスン（リーディング用レッスンも含む）の数は、EEが42、NCが37、NHが40、OWが41、SSが36、TEが34となっている。

③ 総語数に対する異語数の割合（T/T）、つまり、各異語の平均使用回数を比較してみると、HIの値が最も高くなっている。このことから、HIは中学校用教科書と比べて、同じ種類の語彙を多用する傾向にあると言える。中学校用教科書の中では、OW、EE、TEの値が高くなっている。なお、現行の中学校学習指導要領によれば、3年間で扱われるべき新語の数が900～1050語となっているのに、本調査における異語数はどれも、規定の上限を越えている。これは既に紹介したよ

うに、コンピュータによる異語数の数え方が、学習指導要領での教え方と異なっているためである。

以上、対象テキスト全部について、その中に含まれる学習語彙の全体的使用傾向を見てきたが、これだけでは学習語彙について具体的なイメージが掴みにくい。そこで、次にどのような品詞の単語がどれくらいの割合で使用されているのか、またどのランク（重要度）に位置する単語がどのくらいの割合で使われているのかを調べるために、分析の対象を中学校用英語教科書6種のうち、比較的広く使用されているNC、NH、SSの3種と比較対照テキストであるHIにしぼって、それぞれに含まれる学習語彙を品詞別かつランク別に分析し、それぞれの品詞別・ランク別プロフィールを明らかにしていくことにする。

4.2 品詞別分析

ここでは、学習語彙の使用頻度を品詞別に分けて分析していくことにする。まず初めに、総語数で数えた場合の各品詞別使用頻度の割合（％）を比較してみると、次のグラフのような結果が現われた。各グラフには、上で設定した13の品詞ごとの使用頻度だけでなく、内容語と機能語別の使用頻度も示してある。内容語と機能語の区別については後程触れるが、以下のグラフにおいては名詞、形容詞、副詞、（一般）動詞の4品詞に属する単語を内容語とし、それ以外をすべて機能語として使用頻度が算出されている。参考資料としてブラウン・コーパス（BC）における品詞別使用頻度も示しておく⁽³⁾。各グラフの見出し語は、それぞれ、名（名詞）、形（形容詞）、副（副詞）、動（一般動詞）、Be（Be動詞）、助（助動詞）、代（代名詞）、疑（疑問詞）、前（前置詞）、接（接続詞）、冠（冠詞）、数（数量詞）、他（その他）、内（内容語）、機（機能語）と解釈していただきたい。なお、以下のグラフの作成に当たっては、市販の表計算ソフト Lotus 1-2-3を利用した。

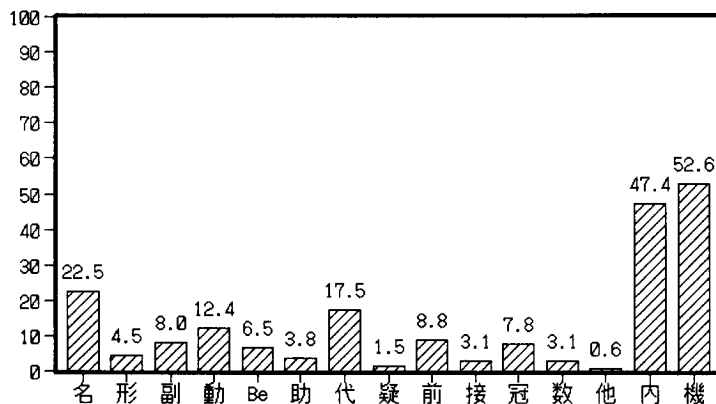


図1 品詞別総語数の割合 (%) (NC)

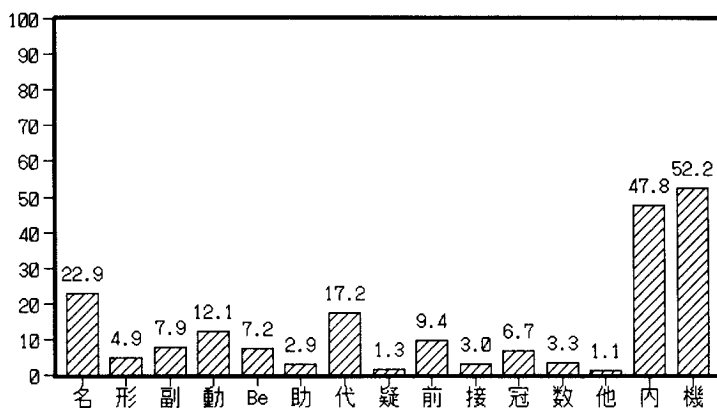


図2 品詞別総語数の割合 (%) (NH)

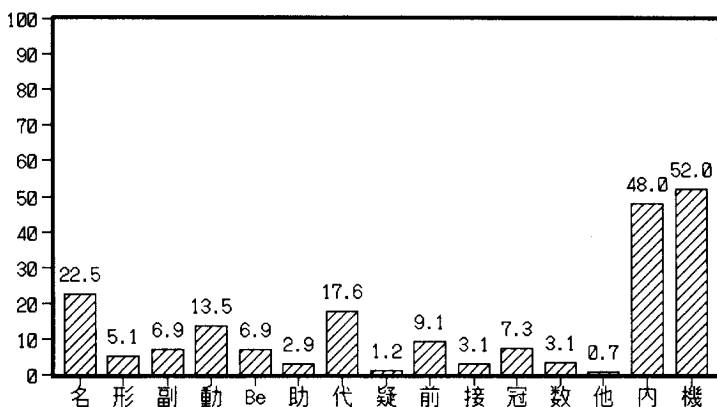


図3 品詞別総語数の割合 (%) (SS)

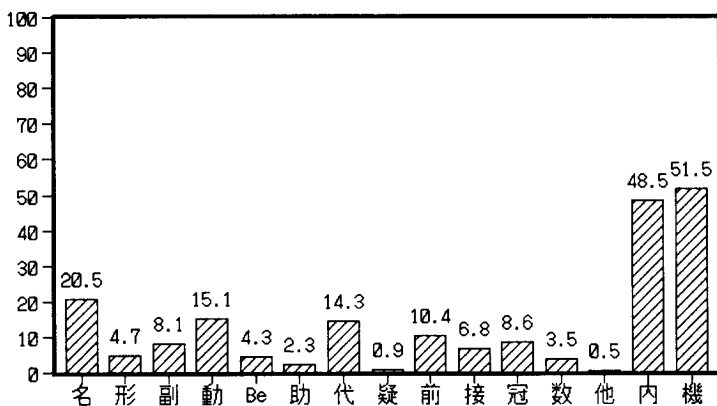


図4 品詞別総語数の割合 (%) (HI)

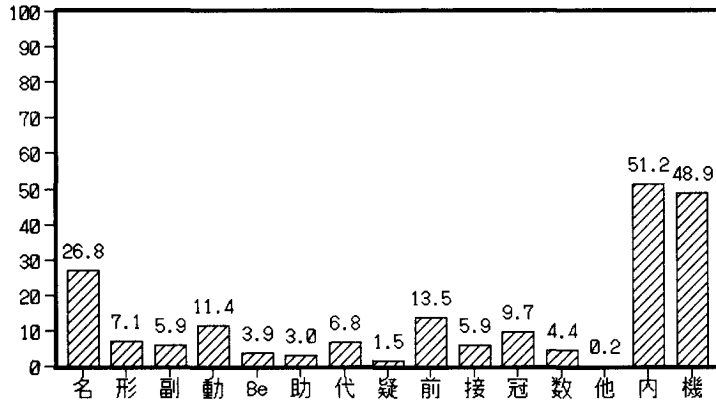


図5 品詞別総語数の割合 (%) (BC)

以上のグラフより次のような特徴が見られる。

① 全体的に見て、どのテキストにおいても共通して名詞の使用頻度の割合が一番高くなっている。この理由として、英語でのコミュニケーションにおいて、名詞の果たす役割が非常に大きいことが挙げられよう。これは、幼児のコミュニケーションが、まず名詞の羅列から始まって、次第に複雑な文章へと移行していくという現象からも明らかであろう。また文構成上の特徴を考えてみても、一文に含まれる名詞の割合が他の品詞に比べて高いということからも裏付けられる。さらに、ブラウン・コーパスでも同様に名詞が全体に占める割合が大きくなっており、この傾向は英語一般の特徴となっているようである。

② 5種類のグラフを比較してみた場合、名詞の次に多いのは、中学校用教科書では代名詞、次いで一般動詞となっているのに対し、HIでは2番目に一般動詞、次いで代名詞となっている。一方、ブラウン・コーパスにおいては、2番目に前置詞、3番目に一般動詞がきており、代名詞は6番目である。一般動詞の使用頻度が代名詞の使用頻度より高いということはHIと共通であるが、代名詞の使用頻度はかなり低く、中学校用教科書の半分以下になっている。以上のことから、学習を意識して作られたテキストの場合は代名詞が多く使用される傾向にあり、中学校用教科書では特にその傾向が顕著になっていると言える。

③ 一般動詞とBe動詞の割合を比較してみると、中学校用教科書においては、Be動詞の割合が一般動詞の割合の約2分の1になっているが、HIではBe動詞の割合が一般動詞の割合の約3分の1になっている。ブラウン・コーパスにおいてもHIと同様、Be動詞の割合が一般動詞の割合の約3分の1になっている。つまり、中学校用教科書では動詞の中でのBe動詞の割合が比較的高くなっている。文型で言えば、S+V+C(第2文型)の使用が比較的多くなっていると言える。

④ 接続詞の割合を見てみると、HIの方が、中学校用教科書より、約2倍程高くなっている。このことから、HIでは、重文あるいは複文が多く使われ、中学校用教科書では、単文が比較的多く使われる傾向にあると言える。なお、HIの接続詞の割合は、ブラウン・コーパスにおける接続詞の割合に比較的近くなっており、その方が自然体に近いと言えるであろう。

⑤ 形容詞と副詞の関係に着目してみると、今回分析の対象にした4種類とも、形容詞よりは副詞の使用頻度の方が高くなっている。一方、ブラウン・コーパスにおいては、形容詞の使用頻

度の方が副詞の使用頻度より高くなっており、逆の傾向を示している。このことは、学習者用のテキストでは、形容詞の使用頻度が比較的少なくなる傾向にあることを示唆している。内容のあるコミュニケーションを行うためには、形容詞は大切であるので、もっと形容詞の使用頻度を増やしてもよいのではないかとと思われる。

⑥ 内容語と機能語の区別は言語研究において重要な区別とみなされてきたが (cf. Fries 1953)、その分け方は必ずしも一定してない嫌いがある。本調査においては、内容語が開放群で機能語が閉鎖群であるという基本的性格に着目し、名詞、形容詞、副詞、一般動詞の部類に含まれるものを内容語とし、それ以外の Be 動詞、助動詞、代名詞、疑問詞、前置詞、接続詞、冠詞、数量詞、その他の部類に含まれるものを機能語とした。内容語と機能語をこのように区別した上で、4種類のテキストにおける両者の割合を比較してみると、いずれのテキストにおいても、大体、半々の数値になっているものの、内容語よりも機能語の出現率の方が若干多くなっている。逆に、ブラウン・コーパスでは機能語の頻度よりも内容語の頻度の方が若干高くなっており、この逆転現象も学習者用教材のひとつの特徴とみなされるかもしれない。

次に異語数の観点から各品詞別傾向を探ると、次のグラフのような結果が見られた。

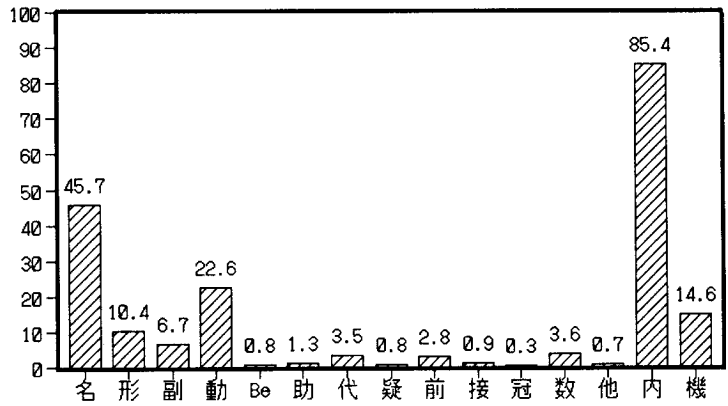


図6 品詞別異語数の割合 (%) (NC)

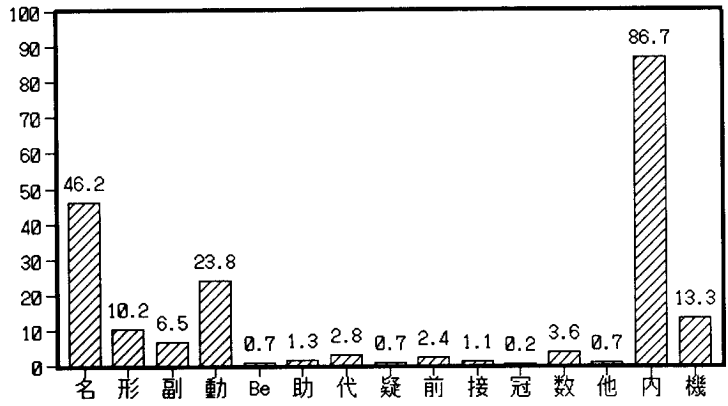


図7 品詞別異語数の割合 (%) (NH)

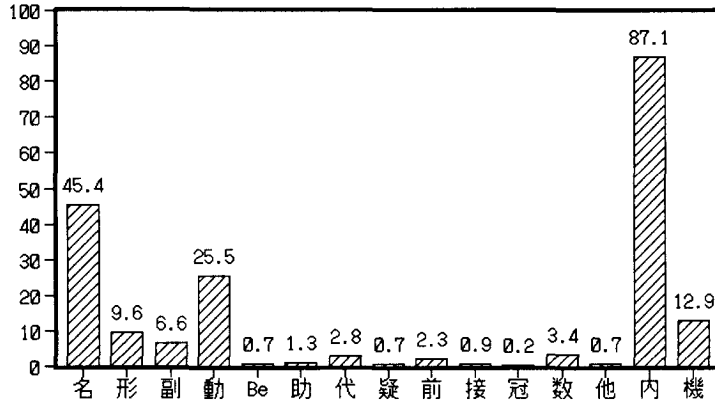


図8 品詞別異語数の割合 (%) (SS)

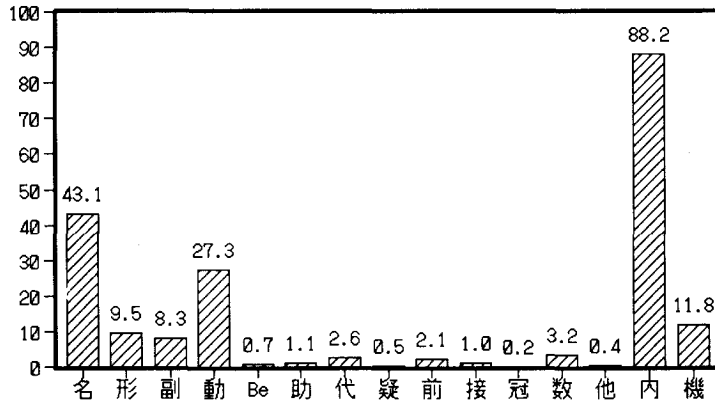


図9 品詞別異語数の割合 (%) (HI)

表2 品詞別総語数に対する異語数の割合 (T/T)

| T/T | 名 | 形 | 副 | 動 | Be | 助 | 代 | 疑 |
|-----|-----|-----|-----|-----|------|------|------|------|
| NC | 3.0 | 2.7 | 7.4 | 3.4 | 47.5 | 18.3 | 31.1 | 12.1 |
| NH | 3.2 | 3.1 | 7.9 | 3.3 | 71 | 14.2 | 39.7 | 13.1 |
| SS | 3.2 | 3.5 | 6.8 | 3.5 | 66.4 | 15 | 41.2 | 11.6 |
| HI | 4.0 | 4.1 | 8.3 | 4.6 | 54.2 | 17.3 | 45.7 | 15.9 |

| T/T | 前 | 接 | 冠 | 数 | 他 | 内 | 機 | 全体 |
|-----|------|------|-------|-----|------|-----|------|-----|
| NC | 19.5 | 20.3 | 189.7 | 5.2 | 5.1 | 3.4 | 22.2 | 6.2 |
| NH | 25.5 | 18.3 | 176 | 5.9 | 10.5 | 3.5 | 25.2 | 6.4 |
| SS | 26.1 | 22.1 | 212 | 6 | 7.1 | 3.6 | 26.3 | 6.5 |
| HI | 40.9 | 54.9 | 393.7 | 9.2 | 11.0 | 4.6 | 36.6 | 8.4 |

以上のグラフと表より次のような傾向が見られる。なお、ブラウン・コーパスに見られる品詞別使用頻度の割合は、総語数での比率であるため、ここでは比較の対象とはしなかった。

① 品詞別に異語数の割合を比較してみると、いずれの場合にも第一位が名詞で全体の5割近くを占めており、次いで一般動詞、形容詞、副詞の順となっている。

② 内容語と機能語別の異語数の割合を比較してみると、テキスト4種類とも、内容語が全異語数の85%以上、機能語が15%未満で、圧倒的に内容語の割合が高くなっている。これは、内容語が解放群であり、機能語が閉鎖群であるという性格上、当然のことであろう。その中でも特にHIにおいて内容語の割合が高くなっているのが特徴的である。

③ 表2に示されているT/Tの数値を比較すると、すでに指摘したように、HIの数値が中学校用教科書と比べて全般的に高くなっている。特に、冠詞、接続詞、前置詞、代名詞など閉鎖群である機能語においてT/Tの値が高くなっているのが目立つ。とりわけ、HIの冠詞と接続詞のT/Tの割合は中学校用教科書と比較して極端に高くなっている。これは、HIにおいては冠詞と接続詞が中学校用教科書と比べて多く使用されていることを示しており、HIの大きな特徴の一つとなっている。

④ 中学校用教科書の中でT/Tの数値を比較してみると、Be動詞と代名詞において、NCの割合が他の2種類の教科書と比べて一段と低くなっているのが分かる。Be動詞も代名詞もどちらも閉鎖群であることから、このことは、NCにおいて両者の絶対的使用回数が他と比べて少ないことを物語っている。NCのひとつの特徴がそこに現われていると言えよう。

4.3 ランク別分析

ここでは、4種類の対象テキストに含まれる語彙をその重要度つまりランク別に分類し、それぞれのテキストの傾向を分析していくことにする。

まず初めに、ランク別総語数の割合(%)を比較してみると、次のグラフのような結果が現われた。なお、ランクの指標となっている1~11の数字の意味については、前述のランクの区分表を参照していただきたい。

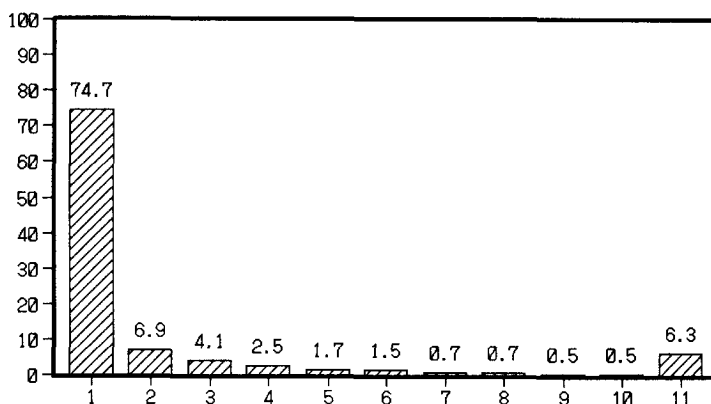


図10 ランク別総語数の割合(%)〈NC〉

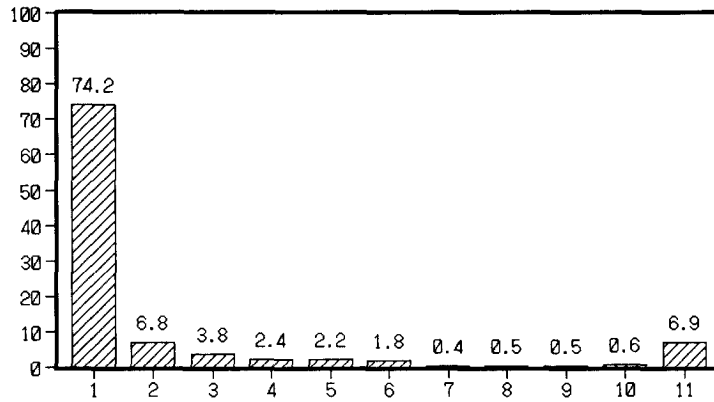


図11 ランク別総語数の割合 (%) (NH)

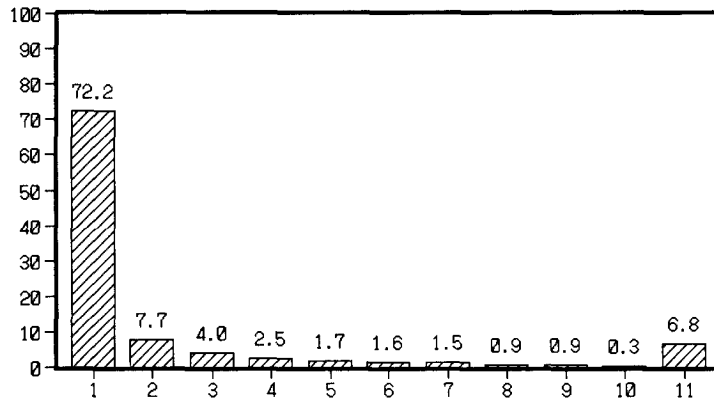


図12 ランク別総語数の割合 (%) (SS)

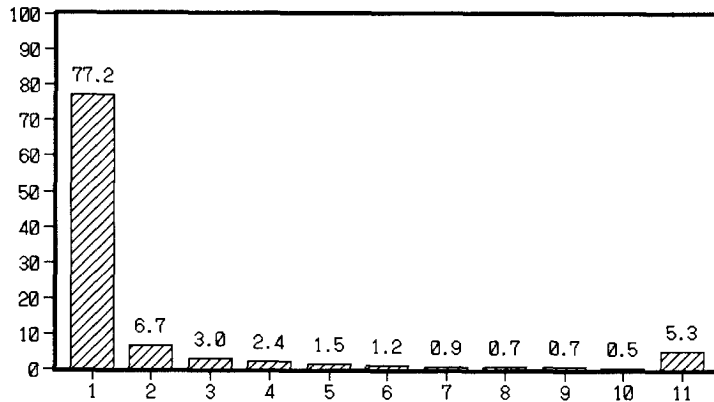


図13 ランク別総語数の割合 (%) (HI)

以上のグラフより次のような傾向が見られる。

① 全体的に見て、ランク1の総語数の割合が圧倒的に高く、平均すると75%を占めている。特に高いのはHIで、77.2%となっている。このことから、いずれのテキストにおいても、全使用語彙数の大半を、ランク1（最初の500語）から取り上げる傾向にあると言える。

② ランク2からは一挙に数値が減少しているのが大きな特徴であり、ランク10まで着実に減少している。ただしNHだけは、わずかな上昇傾向も見られる。

③ HIは、中学校用教科書に比べて、ランク11の数値が低い。つまり、HIでは、最初の5千語に含まれない語彙の使用頻度が中学校用教科書に比べて低くなっていると言える。

次に、上の図に示されているランク別総語数の割合を累積比に換算して比較してみよう。下のグラフは、今回調査した4種類のテキストにおける総語数のランク別累積数をパーセントで表すと同時に、日常英語に使われる語彙を対象にしたThorndike(TD)の調査から得られた割合 (cf. 小川 1982) を添えたものである。

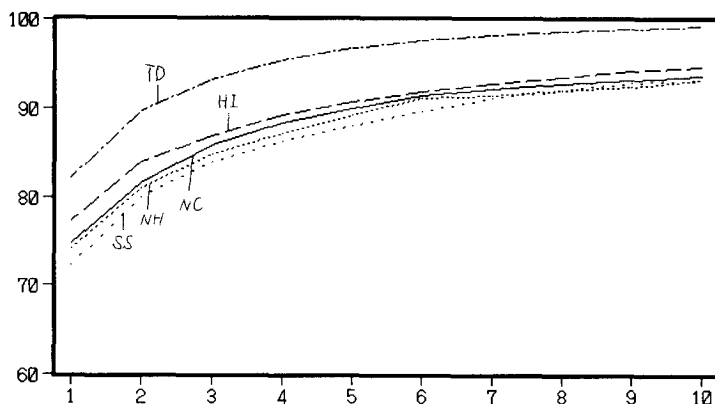


図14 ランク別総語数累積率 (%)

上のグラフから次のような特徴が見られる。

① ランク1とランク2までの語彙が、平均して、全体のほぼ80%を占めている。特に、HIでは、ランク2までの累積数は全体の83.9%といったように、非常に高い数値を示している。これは、今回調査したテキストに使用されている語彙の8割あるいはそれ以上が少なくとも1千語レベル内にあることを示している。つまり、ランク・リストの最初の1千語を完全にマスターしていれば、教科書の中の語彙は8割以上が理解できることになる。頻度の高さが、語彙選定にあたって最も重要な基準であるとは必ずしも言えないが、少なくとも、今回の調査したテキストにおいてはそれが語彙選定の重要な指標になっていると言えるであろう。

② ランク1からランク10までの総語数の累積比率(%)の上昇カーブはいずれのテキストにおいても、大体類似した傾向を示しているが、その中でも、HIの上昇カーブが他のテキストの上昇カーブと比べて早くから急勾配になっている。一方、中学校用教科書の場合、グラフの上昇カーブはNCが最も急で、SSが最も緩くなっている。

③ 4種類のテキストの累積比率(%)をThorndikeの累積比率と比較してみると、全体的に

その上昇カーブは類似した傾向にあると言える。ただ、4種類のテキストの上昇カーブの方が、Thorndikeのものより全体的にやや緩やかになっている。分析の対象となった語彙の母集団の性格や大きさの違いを考えると、Thorndikeの結果と今回の調査の結果を同じ尺度で比較することには若干の危険性を伴うが、上のグラフから判断する限り、4種類のテキストの中ではHIのランク別総語数の累積比率がThorndikeのもの、つまり、日常英語の実態に最も近いと言える。

次に、各ランクごとの異語数の割合(%)を調べてみると次のような結果が現われた。

表3 ランク別総語数に対する異語数の割合(T/T)

| T/T | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 全体 |
|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| NC | 11.6 | 2.8 | 2.8 | 2.2 | 2.2 | 2.6 | 2.2 | 2.5 | 1.9 | 3.7 | 2.6 | 6.2 |
| NH | 12.1 | 3.2 | 2.8 | 2.2 | 2.7 | 2.9 | 1.5 | 2.3 | 2.2 | 3.4 | 2.7 | 6.4 |
| SS | 13.1 | 3.2 | 2.6 | 2.8 | 2.1 | 3.3 | 2.7 | 2.7 | 2.7 | 2.5 | 2.9 | 6.5 |
| HI | 18.9 | 3.8 | 2.7 | 2.9 | 2.2 | 2.3 | 1.8 | 2.3 | 2.1 | 2.3 | 3.2 | 8.4 |

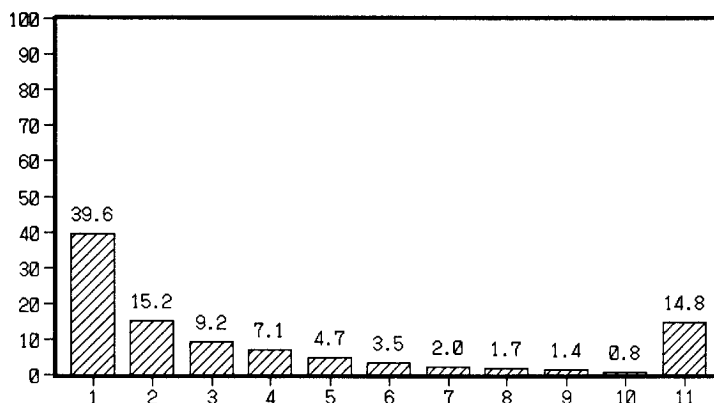


図15 ランク別異語数の割合(%)〈NC〉

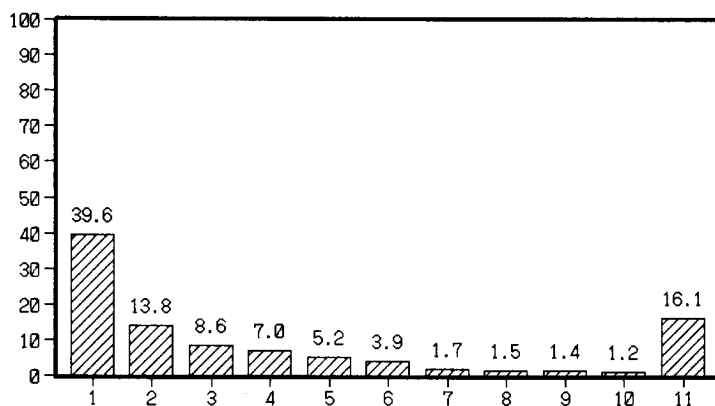


図16 ランク別異語数の割合(%)〈NH〉

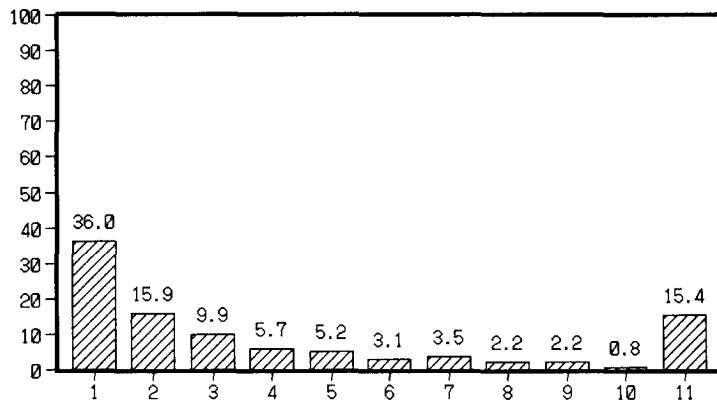


図17 ランク別異語数の割合 (%) (SS)

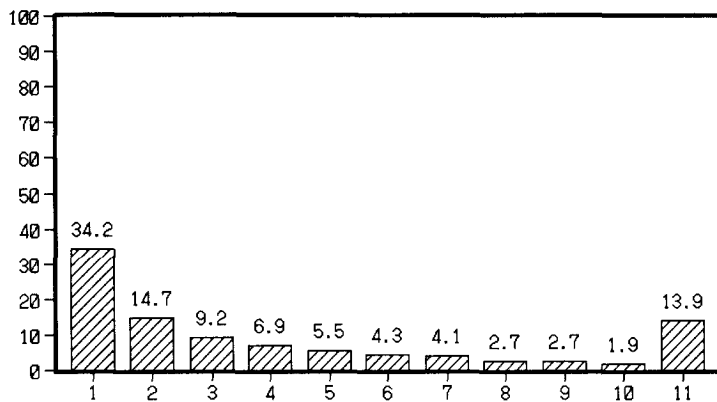


図18 ランク別異語数の割合 (%) (HI)

以上のグラフをもとに、異語数で数えた場合のランク別語彙使用傾向を分析すると次のような特徴が見られる。

① 全体的にみて、総語数の場合と同様、ランク1からランク10までの異語数の減少率はいずれのテキストにおいても類似した傾向にある。つまり、いずれの場合もランク1の値が一番高くランク1からランク2にかけて減少幅が最も大きく、ランク2からランク10にかけては穏やかな勾配で着実に異語数の割合が減少している。ただし、カーブの曲線は総語数の場合ほど急勾配にはなっていない。例えば、ランク1における総語数の割合は全体の7割を占めているのに対し、ランク1の異語数の割合は3割強で、ランク1とランク2を合わせても5割前後に過ぎない。

② ランク1の異語数の割合 (%) をテキスト間で比較してみると、NHとNCの値が同値で最も高く、次いでSSとなり、最後にHIとなっている。ところが、ランク1の総語数の割合を比べてみると、HIの数値が最も高くなっている。つまり、ランク1に関して言えば、HIは、異語数の割合では他の3種類のテキストよりも低くなっているが、総語数の割合では最も高くなっている。このことは、HIが、ランク1に含まれる簡単な語彙を中学校用教科書以上に多く反復して使用する傾向にあることを示している。

③ ランク 2 (1 千語レベル) までに含まれる異語数の割合 (%) を比較してみると、HI に比べて中学校用教科書の方がやや高くなっている。その中でも最も高い数値を示しているのが NC (54.8%) で、一番低いのは SS (51.9%) であるが、いずれも 50% を越えている。このことから、中学校用教科書では、比較的基本的な語彙 (ランク 1、ランク 2) で、全異語数の過半数を消化していると言える。

④ 中学校用教科書においては、3 種類ともランク 6 (3 千語レベル) までに含まれる異語数の割合が全体のほぼ 8 割を占めている。このことから、最初の 3 千語レベルの単語を覚えていれば、テキストに含まれる全異語数の 8 割がわかることになる。

5. まとめ

中学校用英語教科書 3 種 (NC、NH、SS) とその比較対象テキスト (HI) に含まれる学習語彙について、その品詞別・ランク別プロフィールを明らかにしてきたが、最後に両者の相違点を簡単にまとめてみよう。

中学校用教科書と HI との間の最も顕著な相違は、おそらく、ランク別分析において、ランク 1 およびランク 2 における HI の総語数の割合が中学校用教科書の数値と比べて高く、逆に、ランク 11 つまり最初の 5 千語以外の単語においては HI の数値の方が中学校用教科書のものより低くなっているという点であろう。つまり、HI においては比較的簡単な単語が中学校用教科書以上に多く使用されているのである。この事実は、一般的に認められている HI のストーリーのおもしろさが、決してそれが中学校用教科書に見られるような厳しい語彙制限から解放されていることに由来するものではないということを示している。また、見方を変えれば、よく指摘される中学校用教科書の味気なさは、決して学習指導要領によって押し付けられている語彙制限のためではないとも言える。

次に顕著な相違点は、HI における総語数に対する異語数の割合 (T/T) が中学校用教科書と比べて高くなっている点であろう。このことは、ひとつの単語の出現回数が HI のにおいては中学校用教科書においてよりも多くなっていることを示している。学習者にしてみれば、それだけ同じ単語を目にする機会が多いのである。当然、その単語の保持率も高くなるであろう。一方、中学校用教科書においては、ひとつの単語の出現回数が比較的少なく、それだけ学習者の頭の中に残る確率も少なくなるであろう。

品詞別に中学校用教科書と HI を比べてみると、一般動詞や冠詞については HI の割合の方が高くなっている反面、代名詞や Be 動詞については中学校用教科書の割合の方が高くなっているという点も留意すべき重要な相違点であろう。特に、冠詞が少なく、代名詞の使用が多いという中学校用教科書の語彙的特徴は、その中に含まれている英語表現の具体性が全般的に低くなっていると言うことを示唆している。この事実と、HI は中学校用教科書に比べて内容語に富んでいるという事実を考え合わせれば、両者の相違は一層はっきりしてくる。冠詞は、日本人学習者にとって特に学習困難な文法項目のひとつである。冠詞の出現回数が少ないというこの中学校用教科書の語彙的特徴は、日本人学習者にとっての冠詞の学習困難度を軽減する方向には向かっていない。

また、中学校用教科書においても HI においても、形容詞の種類とその使用回数が比較的少なかったという事実も注目に値する。内容豊かなコミュニケーションにとって形容詞は、名詞と同

様に大切な存在である。コミュニケーションを目指した教科書づくりが試みられている中で、この点は是非考慮に入れられるべきであろう。

今回の調査は、この種の研究としては、限られた言語材料を対象にした極めて規模の小さい研究であった。当然、上で述べられてきたことも、そういう前提の下で理解されなければならない。今後、分析の対象を広げることによって、より緻密な分析結果を報告できるように研究を継続して行きたい。

注

- (1) コンピュータへのデータ入力に協力してくれた奈良教育大学英語科の大学院生野村和代さんと同学生菱木香織さんに感謝の意を表したい。
- (2) いわゆる Type-Token Ratio を算出する場合には、異語数を総語数で割るのが一般的なようであるが、ここではその逆に、総語数を異語数で割った後の数値を T/T の値として表示している。故に、本稿での T/T は、各異語の平均出現回数を示していることになる。そこで、もしこの T/T の値が大きければ、それだけ同じ単語を頻繁に使用していることになる。
- (3) Francis and Kučera (1982) は、分析の対象となった約100万語を全部で22の品詞類に分けて、その使用頻度を算出している。本稿では、22の品詞類を「その他」も含めて13の品詞類に圧縮した後に、その使用頻度を換算した。なお、ブラウン・コーパスでの品詞分類と本稿での品詞分類は、完全には一致していないので、その点を考慮の上で、グラフの数値を比較していただきたい。

引用文献

- Francis, W. N. and H. Kučera (1982), *Frequency Analysis of English Usage: Lexicon and Grammar*, Houghton Mifflin Company.
- 速川 浩 (1966), 『教科書に現れた英語単語の研究』大修館書店.
- 伊東治己 (1989), 「パーソナル・コンピュータによる英語教科書用例検索システムの開発」『和歌山大学教育学部紀要 (教育科学)』第38集, 117-127.
- JACET 教材研究委員会 (1983), 『英語講読用教科書のあり方』についてのアンケート調査報告—『JACET 基本語第2次案』を中心に— 大学英語教育学会.
- 垣田直己他 (1977), 『電子計算機による英語教科書の使用語彙総覧—中学校編』溪水社.
- 金田正也 (1981), 「英語学習必要語量の事例研究 (中間報告)」『名古屋学院大学外国語教育紀要』No. 4, 25-37.
- 清川英男 (1987), 「口語英語の基本語リスト作成の試み (最終報告)」『和洋女子大学英文学会誌』No 21, 43-62.
- 教科書研究センター (編) (1984), 『教科書からみた教育課程の国際比較6・英語科編』ぎょうせい.
- Mackey, W. F. (1965), *Language Teaching Analysis*, Longman.
- 三浦省五 (1985), 「英語教科書の語彙に関する研究—日本とフランスの英語教科書の比較を中心として—」『中国地区英語教育学会研究紀要』No. 15, 97-106.
- 縫部義憲 (1984), 「コンピューターによる英語教科書の語彙統計(1)—ソ連の英語教科書の語彙調査—」『鳥取大学教育学部研究報告 (教育科学)』第26巻, 225-254.
- 縫部義憲 (1985), 「コンピューターによる英語教科書の語彙統計(2)—日本の英語教科書の語彙調査—」『鳥取大学教育学部研究報告 (教育科学)』第27巻第1号, 195-209.
- 小川芳夫 (編) (1982), 『英語教授法辞典 (新版)』三省堂.
- Ogden, C. K. (1968), *Basic English: International Second Language*, Harcourt, Brace & World, Inc.
- Palmer, H. E. (1931), *Second Interim Report on Vocabulary Selection*. 語学教育研究所編『英語教授法事典』開拓社, 1962, 281-371に収録.
- West, M. (1936), *A General Service List of English Words, with Semantic Frequencies and Supplementary Word List for the Writing of Popular Science and Technology*, Longman, 1953².

淀縄光洋（1983），『高校英語語彙の実態と学習語彙の在り方』都立教育研究所。

全英連（1963），『中学英単語活用集』南雲堂。『現代英語教育講座 5・英語の語彙』研究社（1966）に収録。

全英連（1981），『高校基本英単語活用集』研究社。

〈参考資料①〉

表4 品詞別統計数値

| 教材 | 視点 | 名詞 | 形容詞 | 副詞 | 動詞 | Be動詞 | 助動詞 | 代名詞 | 疑問詞 | 前置詞 | 接続詞 | 冠詞 | 数量詞 | その他 | 内容語 | 機能語 | 全体 |
|----|-----|------|------|------|------|------|------|------|------|------|------|-------|-----|------|------|------|-------|
| NC | 異語数 | 540 | 123 | 79 | 267 | 10 | 15 | 41 | 9 | 33 | 11 | 3 | 43 | 8 | 1009 | 173 | 1182 |
| | 総語数 | 1643 | 328 | 583 | 901 | 475 | 274 | 1275 | 109 | 644 | 223 | 569 | 224 | 41 | 3455 | 3834 | 7289 |
| | T/T | 3.0 | 2.7 | 7.4 | 3.4 | 47.5 | 18.3 | 31.1 | 12.1 | 19.5 | 20.3 | 189.7 | 5.2 | 5.1 | 3.4 | 22.2 | 6.2 |
| | 異語% | 45.7 | 10.4 | 6.7 | 22.6 | 0.8 | 1.3 | 3.5 | 0.8 | 2.8 | 0.9 | 0.3 | 3.6 | 0.7 | 85.4 | 14.6 | 100 |
| | 総語% | 22.5 | 4.5 | 8.0 | 12.4 | 6.5 | 3.8 | 17.5 | 1.5 | 8.8 | 3.1 | 7.8 | 3.1 | 0.6 | 47.4 | 52.6 | 100 |
| NH | 異語数 | 564 | 124 | 79 | 291 | 8 | 17 | 34 | 8 | 29 | 13 | 3 | 44 | 8 | 1058 | 164 | 1222 |
| | 総語数 | 1795 | 382 | 622 | 952 | 568 | 228 | 1350 | 105 | 740 | 283 | 528 | 260 | 84 | 3751 | 4101 | 7852 |
| | T/T | 3.2 | 3.1 | 7.9 | 3.3 | 71.0 | 13.4 | 39.7 | 13.1 | 25.5 | 18.3 | 176.0 | 5.9 | 10.5 | 3.5 | 25.0 | 6.4 |
| | 異語% | 46.2 | 10.1 | 6.5 | 23.8 | 0.7 | 1.4 | 2.8 | 0.7 | 2.4 | 1.1 | 0.2 | 3.6 | 0.7 | 86.6 | 13.4 | 100 |
| | 総語% | 22.9 | 4.9 | 7.9 | 12.1 | 7.2 | 2.9 | 17.2 | 1.3 | 9.4 | 3.0 | 6.7 | 3.3 | 1.1 | 47.8 | 52.2 | 100 |
| SS | 異語数 | 601 | 127 | 87 | 338 | 9 | 17 | 37 | 9 | 30 | 12 | 3 | 45 | 9 | 1153 | 171 | 1324 |
| | 総語数 | 1947 | 442 | 595 | 1172 | 598 | 255 | 1525 | 104 | 784 | 265 | 636 | 272 | 64 | 4156 | 4503 | 8659 |
| | T/T | 3.2 | 3.5 | 6.8 | 3.5 | 66.4 | 15.0 | 41.2 | 11.6 | 26.1 | 22.1 | 212.0 | 6.0 | 7.1 | 3.6 | 26.3 | 6.5 |
| | 異語% | 45.4 | 9.6 | 6.6 | 25.5 | 0.7 | 1.3 | 2.8 | 0.7 | 2.3 | 0.9 | 0.2 | 3.4 | 0.7 | 87.1 | 12.9 | 100 |
| | 総語% | 22.5 | 5.1 | 6.9 | 13.5 | 6.9 | 2.9 | 17.6 | 1.2 | 9.1 | 3.1 | 7.3 | 3.1 | 0.7 | 48.0 | 52.0 | 100 |
| HI | 異語数 | 709 | 157 | 136 | 450 | 11 | 18 | 43 | 8 | 35 | 17 | 3 | 53 | 6 | 1452 | 194 | 1646 |
| | 総語数 | 2828 | 641 | 1122 | 2086 | 596 | 312 | 1963 | 127 | 1432 | 934 | 1181 | 485 | 66 | 6677 | 7096 | 13773 |
| | T/T | 4.0 | 4.1 | 8.3 | 4.6 | 54.2 | 17.3 | 45.7 | 15.9 | 40.9 | 54.9 | 393.7 | 9.2 | 11.0 | 4.6 | 36.6 | 8.4 |
| | 異語% | 43.1 | 9.5 | 8.3 | 27.3 | 0.7 | 1.1 | 2.6 | 0.5 | 2.1 | 1.0 | 0.2 | 3.2 | 0.4 | 88.2 | 11.8 | 100 |
| | 総語% | 20.5 | 4.7 | 8.1 | 15.1 | 4.3 | 2.3 | 14.3 | 0.9 | 10.4 | 6.8 | 8.6 | 3.5 | 0.5 | 48.5 | 51.5 | 100 |

中学校用英語教科書に含まれる学習語彙の計量的分析

〈参考資料②〉

表5 ランク別統計数値

| 教材 | 視点 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 全体 |
|-----|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| NC | 異語数 | 468 | 180 | 109 | 84 | 55 | 41 | 24 | 20 | 17 | 9 | 175 | 1182 |
| | 総語数 | 5446 | 503 | 302 | 183 | 121 | 108 | 52 | 49 | 33 | 33 | 459 | 7289 |
| | T/T | 11.9 | 2.8 | 2.8 | 2.2 | 2.2 | 2.6 | 2.2 | 2.5 | 1.9 | 3.7 | 2.6 | 6.2 |
| | 異語% | 39.6 | 15.2 | 9.2 | 7.1 | 4.7 | 3.5 | 2.0 | 1.7 | 1.4 | 0.8 | 14.8 | 100 |
| | 総語% | 74.7 | 6.9 | 4.1 | 2.5 | 1.7 | 1.5 | 0.7 | 0.7 | 0.5 | 0.5 | 6.3 | 100 |
| | 異累積 | 468 | 648 | 757 | 841 | 896 | 937 | 961 | 981 | 998 | 1007 | 1182 | 1182 |
| | 総累積 | 5446 | 5949 | 6251 | 6434 | 6555 | 6663 | 6715 | 6764 | 6797 | 6830 | 7289 | 7289 |
| | 異累% | 39.6 | 54.8 | 64.0 | 71.2 | 75.8 | 79.3 | 81.3 | 83.0 | 84.4 | 85.2 | 100.0 | 100 |
| 総累% | 74.7 | 81.6 | 85.8 | 88.3 | 89.9 | 91.4 | 92.1 | 92.8 | 93.3 | 93.7 | 100.0 | 100 | |
| NH | 異語数 | 483 | 169 | 105 | 86 | 63 | 48 | 21 | 18 | 17 | 15 | 196 | 1221 |
| | 総語数 | 5824 | 537 | 296 | 185 | 169 | 140 | 32 | 42 | 38 | 51 | 538 | 7852 |
| | T/T | 12.1 | 3.2 | 2.8 | 2.2 | 2.7 | 2.9 | 1.5 | 2.3 | 2.2 | 3.4 | 2.7 | 6.4 |
| | 異語% | 39.6 | 13.8 | 8.6 | 7.0 | 5.2 | 3.9 | 1.7 | 1.5 | 1.4 | 1.2 | 16.1 | 100 |
| | 総語% | 74.2 | 6.8 | 3.8 | 2.4 | 2.2 | 1.8 | 0.4 | 0.5 | 0.5 | 0.6 | 6.9 | 100 |
| | 異累積 | 483 | 652 | 757 | 843 | 906 | 954 | 975 | 993 | 1010 | 1025 | 1221 | 1221 |
| | 総累積 | 5824 | 6361 | 6657 | 6842 | 7011 | 7151 | 7183 | 7225 | 7263 | 7314 | 7852 | 7852 |
| | 異累% | 39.6 | 53.4 | 62.0 | 69.0 | 74.2 | 78.1 | 79.9 | 81.3 | 82.7 | 83.9 | 100.0 | 100 |
| 総累% | 74.2 | 81.0 | 84.8 | 87.1 | 89.3 | 91.1 | 91.5 | 92.0 | 92.5 | 93.1 | 100.0 | 100 | |
| SS | 異語数 | 477 | 210 | 131 | 76 | 69 | 41 | 47 | 29 | 29 | 11 | 204 | 1324 |
| | 総語数 | 6250 | 665 | 346 | 214 | 148 | 137 | 128 | 79 | 77 | 28 | 587 | 8659 |
| | T/T | 13.1 | 3.2 | 2.6 | 2.8 | 2.1 | 3.3 | 2.7 | 2.7 | 2.7 | 2.5 | 2.9 | 6.5 |
| | 異語% | 36.0 | 15.9 | 9.9 | 5.7 | 5.2 | 3.1 | 3.5 | 2.2 | 2.2 | 0.8 | 15.4 | 100 |
| | 総語% | 72.2 | 7.7 | 4.0 | 2.5 | 1.7 | 1.6 | 1.5 | 0.9 | 0.9 | 0.3 | 6.8 | 100 |
| | 異累積 | 477 | 687 | 818 | 894 | 963 | 1004 | 1051 | 1080 | 1109 | 1120 | 1324 | 1324 |
| | 総累積 | 6250 | 6915 | 7261 | 7475 | 7623 | 7760 | 7888 | 7967 | 8044 | 8072 | 8659 | 8659 |
| | 異累% | 36.0 | 51.9 | 61.8 | 67.5 | 72.7 | 75.8 | 79.4 | 81.6 | 83.8 | 84.6 | 100.0 | 100 |
| 総累% | 72.2 | 79.9 | 83.9 | 86.3 | 88.0 | 89.6 | 91.1 | 92.0 | 92.9 | 93.2 | 100.0 | 100 | |
| HI | 異語数 | 563 | 242 | 151 | 113 | 91 | 70 | 67 | 44 | 44 | 32 | 229 | 1646 |
| | 総語数 | 10636 | 917 | 415 | 326 | 203 | 159 | 121 | 103 | 91 | 75 | 727 | 13773 |
| | T/T | 18.9 | 3.8 | 2.7 | 2.9 | 2.2 | 2.3 | 1.8 | 2.3 | 2.1 | 2.3 | 3.2 | 8.4 |
| | 異語% | 34.2 | 14.7 | 9.2 | 6.9 | 5.5 | 4.3 | 4.1 | 2.7 | 2.7 | 1.9 | 13.9 | 100 |
| | 総語% | 77.2 | 6.7 | 3.0 | 2.4 | 1.5 | 1.2 | 0.9 | 0.7 | 0.7 | 0.5 | 5.3 | 100 |
| | 異累積 | 563 | 805 | 956 | 1069 | 1160 | 1230 | 1297 | 1341 | 1385 | 1417 | 1646 | 1646 |
| | 総累積 | 10636 | 11553 | 11968 | 12294 | 12497 | 12656 | 12777 | 12880 | 12971 | 13064 | 13773 | 13773 |
| | 異累% | 34.2 | 48.9 | 58.1 | 64.9 | 70.5 | 74.7 | 78.8 | 81.5 | 84.1 | 86.1 | 100.0 | 100 |
| 総累% | 77.2 | 83.9 | 86.9 | 89.3 | 90.7 | 91.9 | 92.8 | 93.5 | 94.2 | 94.7 | 100.0 | 100 | |

A Computational Analysis of the Vocabulary in the English Textbooks for Japanese Junior High School

Harumi ITO and Ayumi YAMADA

(Department of English, Nara University of Education, Nara 630, Japan)

(Received April 23, 1991)

The purpose of this paper is to investigate the nature of the vocabulary contained in English textbooks for Japanese junior high school through an effective and extensive use of the personal computer. Specifically, the paper aims at constructing vocabulary profiles for each target text by counting the number of unique words and the total occurrence of words in each of the 13 different parts of speech and also in each of the 11 different ranks of importance. For this purpose, out of the current six different English textbooks for Japanese junior high school, the three most widely used textbooks were selected as the target texts for the detailed analysis. One EFL textbook was also selected out of many to make a comparison possible between the Japanese junior high school textbooks and EFL textbooks coming from abroad. The vocabulary data contained in the four target texts was processed by the personal computer. As a result, the alphabetical list of unique lexical words with tags for proper parts of speech and ranks of importance, and with frequencies of occurrence, was produced for each target text. The data contained in this alphabetical list of words was further processed by the computer, producing graphic representations of the vocabulary profiles covering the 13 parts of speech and the 11 ranks of importance. These obtained profiles clearly indicate several vocabulary characteristics peculiar to the English textbooks for Japanese junior school, including a more frequent appearance of copulas and pronouns and a less frequent appearance of the articles, compared with the EFL textbook also analyzed in this study.