

機能語と実質語の連続性に関する プロトタイプ論的考察

山内博之

1. はじめに

言語は語によって構成され、語には機能語と実質語の2種類があると考えられている。英語で言えば、機能語は「function word」であり、実質語は「content word」である。そのためか、実質語のことを内容語と言うこともある。

日本語学においては、機能語と実質語（内容語）は対立するものであるととらえられている。たとえば、寺村（1982）では、語の中で、関係の意味を担うものが機能語であり、辞書の意味を担うものが実質語であると述べられている。辞書の意味を担う複数の実質語が、機能語によって関係づけられることによって文が構成されるというのが、寺村（1982）における文のとらえ方である。

語を機能語と実質語の2つに分けるという考え方は、日本語学のみでなく、一般言語学においても広く認められている。亀井・河野・千野（1996）においては、他の語との統語的・文法的関係を示す語が機能語であり、語彙の意味（lexical meaning）を持つ語が内容語であると記されている。さらに、亀井・河野・千野（1996）では、代名詞・前置詞・接続詞・助動詞・冠詞が機能語に属し、名詞・形容詞・動詞・副詞が内容語に属するという事も記されている。寺村（1982）における機能語と実質語の二分法は、このような一般言語学における語の考え方を踏襲したものであると考えることができる。

小さい言語単位が接続して大きい単位を形成する時のルールが文法であるので、機能語とは、まさに文法研究の中心的な要素となるものである。また、語彙とは語の集まりのことであるが、それは往々にして、ある領域における語の集まりであったり、ある観点から見た語の集まりであったりするので、それらは辞書的な意味を担っている必要がある。つまり、語彙研究の中心的な要素と

(2)

なるのは実質語だということである。文法と語彙との対立関係は、機能語と実質語との対立関係と概ねパラレルであると言える。

第二言語としての日本語の習得研究においても、文法習得と語彙習得は対立する概念としてとらえられている。たとえば、野田・迫田・渋谷・小林 (2001) 『日本語学習者の文法習得』という論文集があるが、これは、そのタイトルのとおり、文法習得の研究に焦点を絞った一書である。一方、2003年に開催された第14回第二言語習得研究会全国大会においては、「語彙の習得」というテーマでパネルディスカッションが行なわれた。その内容は金澤 (2004) に詳しくまとめられているが、このパネルディスカッションでは、4名の登壇者により、日本語の語彙習得に関する知見が述べられた。このように、第二言語としての日本語の習得研究においても、文法と語彙、つまり、機能語と実質語は、基本的に対立するものとしてとらえられている。

以上のように、一般言語学においても、また、日本語学や第二言語としての日本語の習得研究においても、機能語と実質語は、基本的に対立するものとしてとらえられている。確かに、機能語と実質語は異なるものであり、それは、「を」という格助詞と「机」という名詞を見比べてみれば明らかであろう。格助詞「を」は主に名詞と動詞の関係的意味を表すものであり、辞書的意味を有しているとは考えにくい。一方、名詞「机」は、「desk」という辞書的意味を有しているが、他の語との関係的意味を表しているとは考えにくい。つまり、格助詞「を」と名詞「机」は、同質のものであるとは考えにくく、異なるタイプの語であると考えの方が明らかに妥当であろう。

では、接続詞「それに」と副詞「さらに」の違いはどうだろうか。亀井・河野・千野 (1996) によれば、接続詞は機能語に属し、副詞は実質語 (内容語) に属するとのことなので、接続詞「それに」と副詞「さらに」の間には、格助詞「を」と名詞「机」の間にあると同様の違いがあっても不思議ではない。しかし、実際には、接続詞「それに」と副詞「さらに」は、かなり似た語であるという印象がもたれるのではないか。次の (1) (2) を見ていただきたい。

(1) 時間が^()ないし、^()お金もない。

(2) たくさんの建物が破壊され、^()多くの人の命までもが奪われた。

上記の (1) (2) の ^() には、「それに」も「さらに」も入り得るように思われるし、どちらを入れても、文意に大きな差があるようには感じられない。また、「それに」「さらに」が辞書的意味と関係的意味のどちらを担っているの

かと問われると、なかなか答えにくい。接続詞「それに」と副詞「さらに」の間には、格助詞「を」と名詞「机」の間にあるような大きな違いは感じられない。

本稿では、機能語と実質語が、語を二分する分類であるとは考えず、連続的に並んでいるものであると考え、そのありようを実証的に示すことを目的とする。つまり、語を二分法でとらえるのではなく、プロトタイプ論的にとらえ、その実態を示すことを目的とするということである。まず、第2章では、BCCWJ（現代日本語書き言葉均衡コーパス）を用いて、主に日本語学的な側面から、機能語と実質語の連続性を観察する。そして、次に第3章では、KY コーパスを用いて、主に第二言語習得研究的な側面から、機能語と実質語の連続性を観察する。続く第4章では、「助詞」「動詞」「名詞」の3種の品詞にターゲットを絞り、再度 KY コーパスを用いて機能語と実質語の連続性を観察する。最後の第5章では、全体のまとめと日本語教育への提言を行なう。

なお、本稿では、実質語と内容語を同一のものであると考えるが、寺村（1982）に倣って、実質語という呼び方を用いることにする。

2. BCCWJ から見る語彙と文法の連続性

この章では、BCCWJ を用いて、主に日本語学的な側面から、機能語と実質語の連続性を観察する。

BCCWJ は1億語コーパスと呼ばれることもあるが、その名のとおりに、BCCWJ には約1億の形態素が含まれている。国立国語研究所のホームページでは、それら1億の形態素の品詞情報が公開されているので、この章では、それらを利用して分析を行なう。

次の表1は、BCCWJ の各品詞の延べ語数と異なり語数をまとめたものである。ただし、品詞の立て方については、筆者が若干の修正を加えた。国立国語研究所のホームページでは、「形容詞」の他に「形状詞」という品詞が立てられているのだが、表1では、これら2つをまとめて「形容詞」とした。また、同様に、「接頭辞」と「接尾辞」という2つの品詞も、表1では、両方をまとめて「接辞」とした。

表1の延べ語数の列を見ると、BCCWJ には、全部で104,612,423の形態素が、つまり、1億を少し超える数の形態素が含まれていることがわかる。これら約1億の形態素の中で際立って数が多いのが、助詞と名詞である。そして、それに続くのが助動詞と動詞である。これら4つの品詞のみが、延べ語数で1千万

(4)

表1 品詞別に見た BCCWJ の延べ語数と異なり語数

| 品詞 | 延べ語数 | 異なり語数 | 延べ語数／異なり語数 |
|-----|-------------|---------|------------|
| 助詞 | 31,428,580 | 139 | 226,105 |
| 助動詞 | 10,279,970 | 71 | 144,788 |
| 連体詞 | 997,276 | 48 | 20,777 |
| 接続詞 | 481,094 | 35 | 13,746 |
| 代名詞 | 1,516,372 | 121 | 12,532 |
| 接辞 | 4,215,052 | 1,047 | 4,026 |
| 動詞 | 14,148,216 | 9,540 | 1,483 |
| 形容詞 | 2,902,230 | 2,483 | 1,169 |
| 副詞 | 1,830,329 | 3,071 | 596 |
| 感動詞 | 161,716 | 371 | 436 |
| 名詞 | 36,651,588 | 168,210 | 218 |
| 合計 | 104,612,423 | 185,136 | |

語を超えている。

表1の右端の列には、延べ語数を異なり語数で除した値を載せた。たとえば、最上段の助詞の行を見ていただくと、延べ語数である31,428,580を異なり語数である139で割ることによって得た値である226,105が、右端の列に記載されていることがわかる。表1は、最上段に助詞があり、次に、助動詞、連体詞と続き、最下段が名詞となっているが、この順序は「延べ語数／異なり語数」の値の大から小への順序である。

「延べ語数／異なり語数」の値が最も大きいのが助詞の226,105であるが、これは、BCCWJの中で、1種類の助詞が平均して226,105回も使用されていることを示している。逆に、「延べ語数／異なり語数」の値が最も小さいのが名詞の218であるが、これは、1種類の名詞が平均して218回しか使用されていないことを示している。

この「延べ語数／異なり語数」の値が大きいということ、つまり、異なり語数に対する延べ語数の割合が大きいということは、まさに機能語の特徴と言っているのではないかと。逆に、「延べ語数／異なり語数」の値が小さいということ、つまり、異なり語数に対する延べ語数の割合が小さいということは、まさに実質語の特徴と言っているのではないかと。表1における「延べ語数／異なり語数」の値を、最上段の助詞から最下段の名詞まで見ていくと、概ね連続的に変化しているように見える。表1に見られる11種類の品詞は、最も機能語的な助詞

から、最も実質語的な名詞まで、概ね連続的に並んでいると言ってもいいのではないだろうか。

極めて大雑把な見方ではあるが、「延べ語数／異なり語数」の数字の桁数で表1の品詞を分類してみると、次の(3)のようになる。

- (3) ①6桁：助詞・助動詞
- ②5桁：連体詞・接続詞・代名詞
- ③4桁：接辞・動詞・形容詞
- ④3桁：副詞・感動詞・名詞

上記の(3)では、①6桁の助詞・助動詞から④3桁の副詞・感動詞・名詞までが連続して並んでいる。しかし、表1に戻って、①から④までのそれぞれのグループ間の「延べ語数／異なり語数」の数値の差を細かく見ていくと、①6桁の助詞・助動詞と②5桁の連体詞・接続詞・代名詞の間には、やや大きな差があるように感じられる。機能語と実質語の連続性に関して表1から読みとれることを、この章の結論として述べると、次の(4)のようになる。

- (4) BCCWJにおける11種類の品詞は、最も機能語らしい助詞から最も実質語らしい名詞まで、概ね連続的に並んではいるが、助詞・助動詞とその他の品詞との間には、やや大きな差があるように感じられる。

3. KY コーパスから見た語彙と文法の連続性

この章では、KY コーパスを用いて、主に、第二言語としての日本語の習得という側面から、機能語と実質語の連続性を観察する。具体的には、まず、前章の表1と同様の表を、KY コーパスを用いて作成する。そして、それを眺め、前章の表1と比較して考察を行なう。

KY コーパスは、一般にはテキストファイルで配布されているが、筆者は、形態素解析済みのKY コーパスを早稲田大学の李在鎬氏からいただくことができたので、李氏による形態素解析済みのKY コーパスを用いて表1に相当する表を作成することにした。

表1との比較を目的とした表を作るためには、品詞名を表1と同じにする必要がある。李氏は茶釜を用いてKY コーパスの形態素解析を行っていたので、茶釜の品詞タグを、次の表2のように変換することにした。

(6)

表2 茶釜タグの変換

| 茶釜のタグ | 変換したタグ |
|--|--------|
| 「フィルター」 | フィルター |
| 「感動詞」 | 感動詞 |
| 「形容詞 - 自立」「形容詞 - 非自立」「名詞 - ナイ形容詞語幹」「名詞 - 形容動詞語幹」 | 形容詞 |
| 「助詞 - 格助詞 - 一般」「助詞 - 格助詞 - 引用」「助詞 - 格助詞 - 連語」「助詞 - 係助詞」「助詞 - 終助詞」「助詞 - 接続助詞」「助詞 - 副詞化」「助詞 - 副助詞」「助詞 - 副助詞／並立助詞／終助詞」「助詞 - 並立助詞」「助詞 - 連体化」 | 助詞 |
| 「助動詞」「名詞 - 特殊 - 助動詞語幹」 | 助動詞 |
| 「接続詞」 | 接続詞 |
| 「形容詞 - 接尾」「接頭詞 - 数接続」「接頭詞 - 名詞接続」「動詞 - 接尾」「名詞 - 接尾 - サ変接続」「名詞 - 接尾 - 一般」「名詞 - 接尾 - 形容動詞語幹」「名詞 - 接尾 - 助数詞」「名詞 - 接尾 - 助動詞語幹」「名詞 - 接尾 - 人名」「名詞 - 接尾 - 地域」「名詞 - 接尾 - 特殊」「名詞 - 接尾 - 副詞可能」 | 接辞 |
| 「動詞 - 自立」「動詞 - 非自立」 | 動詞 |
| 「副詞 - 一般」「副詞 - 助詞類接続」 | 副詞 |
| 「名詞 - サ変接続」「名詞 - 一般」「名詞 - 固有名詞 - 一般」「名詞 - 固有名詞 - 人名 - 一般」「名詞 - 固有名詞 - 人名 - 姓」「名詞 - 固有名詞 - 人名 - 名」「名詞 - 固有名詞 - 組織」「名詞 - 固有名詞 - 地域 - 一般」「名詞 - 固有名詞 - 地域 - 国」「名詞 - 数」「名詞 - 非自立 - 一般」「名詞 - 非自立 - 形容動詞語幹」「名詞 - 非自立 - 助動詞語幹」「名詞 - 非自立 - 副詞可能」「名詞 - 副詞可能」 | 名詞 |
| 「名詞 - 代名詞 - 一般」 | 代名詞 |
| 「連体詞」 | 連体詞 |

上記の表2のようにして茶釜の品詞タグを変換し、次の表3を作成した。KY コーパスには90人分のOPIデータが収録されているので、表3は、OPIの被験者90人分の発話データに含まれた形態素について、延べ語数、異なり語数と、延べ語数を異なり語数で除した値を、品詞別にまとめた表ということになる。

表3の延べ語数の列を見ると、KY コーパスには166,043の形態素が含まれていることがわかるが、その中で際立って数が多いのが助詞と名詞であり、そして、それに続くのが助動詞と動詞である。延べ語数の多さが際立っている品詞は、BCCWJにおいても、また、KY コーパスにおいても、助詞、名詞、助

表3 品詞別に見た KY コーパスの延べ語数と異なり語数

| 品詞 | 延べ語数 | 異なり語数 | 延べ語数／異なり語数 |
|------|---------|-------|------------|
| 助動詞 | 20,430 | 48 | 426 |
| 助詞 | 45,745 | 151 | 303 |
| フィラー | 9,301 | 157 | 59 |
| 代名詞 | 3,822 | 69 | 55 |
| 連体詞 | 2,347 | 45 | 52 |
| 感動詞 | 7,902 | 172 | 46 |
| 接続詞 | 2,678 | 84 | 32 |
| 副詞 | 7,949 | 375 | 21 |
| 動詞 | 18,236 | 924 | 20 |
| 接辞 | 4,419 | 323 | 14 |
| 形容詞 | 5,930 | 439 | 14 |
| 名詞 | 37,284 | 4,455 | 8 |
| 合計 | 166,043 | 7,242 | |

動詞、動詞の4者であった。

次に、右端の「延べ語数／異なり語数」を見ていただきたい。前章の表1では、数値の桁数で品詞をグルーピングしたが、表3では、助動詞と助詞が3桁で、その他はすべて2桁であるので、この方法だと2グループにしか分類できないことになってしまう。そこで、「延べ語数／異なり語数」が2桁の品詞を、さらに2つのグループに分けることにした。表3に現われた12種類の品詞の「延べ語数／異なり語数」の平均値を計算してみたら23であったので、2桁のグループを、「延べ語数／異なり語数」の値が23より大きいフィラー、代名詞、連体詞、感動詞、接続詞と、「延べ語数／異なり語数」の値が23より小さい副詞、動詞、接辞、形容詞、名詞とに分けることにした。この分類を示したものが、次の(5)である。

- (5) ① 3桁 : 助動詞・助詞
 ② 2桁平均以上: フィラー・代名詞・連体詞・感動詞・接続詞
 ③ 2桁平均以下: 副詞・動詞・接辞・形容詞・名詞

①の3桁のグループに属するのは助動詞・助詞であり、この2つが最も機能語らしいグループに属するというのは、表1と同じ、つまり、BCCWJと同じ

である。ただし、表1と表3を見比べてみると、「延べ語数／異なり語数」の数値において、助詞と助動詞の順位が入れ替わっていることがわかる。表1のBCCWJにおいては助詞が1位で助動詞が2位であったが、表3のKYコーパスでは助動詞が1位で助詞が2位となっている。

この理由は、おそらく、BCCWJが書き言葉で、一方、KYコーパスが話し言葉であるということにあるのではないかと思われる。書き言葉であるBCCWJは、データのほとんどが普通体の日本語であり、一方、話し言葉であるKYコーパスは、データのほとんどが丁寧体の日本語である。まったく同じ内容の文であっても、普通体を丁寧体に変換するだけで助動詞の数は増えていく。

たとえば、「食べる」という述語を考えた場合、助動詞が付加されていないので、助動詞の数は0である。しかし、その丁寧形である「食べます」は、助動詞「食べる」の連用形に助動詞「ます」が付加されているので、助動詞の数は1となる。「食べない」は、動詞「食べる」の連用形に助動詞「ない」が付加されており、助動詞の数は1であるが、その丁寧形である「食べません」は、動詞「食べる」の連用形に、助動詞「ませ」と助動詞「ん」が付加されているので、助動詞の数は2となる。このように、普通体の文を丁寧体に変換すると、一般に、助動詞の数は増えていく。そのため、ほとんどのデータが普通体の日本語であると思われるBCCWJよりも、ほとんどのデータが丁寧体であると思われるKYコーパスの方が、助動詞の出現回数が相対的に多くなっているのであろう。

助動詞の出現回数の多寡は、そのデータが書き言葉であるのか話し言葉であるのかということに大きく影響を受けるものと思われるが、表1と表3の品詞においては、フィラーと感動詞の出現回数も、書き言葉であるのか話し言葉であるのかということに影響を受けていることが予想される。そもそも、書き言葉にはフィラーはないであろうし、「はい」や「おはよう」などの感動詞も、書き言葉ではそう多くは現れないであろう。

本稿では、第2章においては、BCCWJをデータとして用いることによって、主に日本語学的な側面から機能語と実質語の連続性を探ることを目的とし、第3章においては、KYコーパスを用いることによって、主に第二言語としての日本語の習得という側面から機能語と実質語の連続性を探りたいと考えている。つまり、BCCWJは母語話者コーパスの代表として選んだものであり、KYコーパスは学習者コーパスの代表として選んだものである。したがって、書き言葉であるのか話し言葉であるのかという変数に影響を及ぼされたく

表4 BCCWJ と KY コーパスの品詞の連続性（フィラー、感動詞を除く）

| 分類 | BCCWJ | KY コーパス |
|------|--------------|------------------|
| 機能語的 | ①助詞・助動詞 | ①助詞・助動詞 |
| 中間 | ②連体詞・接続詞・代名詞 | ②連体詞・接続詞・代名詞 |
| 実質語的 | ③接辞・動詞・形容詞 | ③接辞・動詞・形容詞・副詞・名詞 |
| | ④副詞・名詞 | |

はない。そこで、フィラーと感動詞を除いて、表1と表3を比較することにする。

表4は、「延べ語数／異なり語数」の数値によって行なった、最も機能語らしい品詞から最も実質語らしい品詞までの、表1と表3における分類のありようを、1つの表にまとめて示したものである。

(3)は、表1の考察の結果をまとめたものであるが、(3)における①と②を、表4においては、それぞれ、「機能語的」「中間」とし、③と④はまとめて「実質語的」とした。一方、(5)は、表3の考察の結果をまとめたものであるが、(5)における①②③を、表4においては、それぞれ「機能語的」「中間」「実質語的」とした。

このようにして、(3)の①～④、及び、(5)の①～③を再分類した表4を見ると、「機能語的」「中間」「実質語的」というそれぞれの分類に属している品詞は、BCCWJとKYコーパスで、完全に一致していることがわかる。BCCWJとKYコーパスのいずれにおいても、助詞・助動詞は「機能語的」であり、連体詞・接続詞・代名詞は「中間」であり、接辞・動詞・形容詞・副詞・名詞は「実質語的」である。

BCCWJとKYコーパスのいずれにおいても、助詞・助動詞と他の品詞との間に、やや大きな不連続が認められはするものの、全体的には、BCCWJとKYコーパスのいずれにおいても、つまり、日本語学的な側面と第二言語としての日本語の習得という側面のいずれにおいても、機能語のプロトタイプである助詞・助動詞から、実質語のプロトタイプである名詞まで、概ね滑らかに連続しているのではないかと思うのだが、いかがであろうか。

4. 助詞・動詞・名詞の連続性

この章では、助詞、動詞、名詞という3種類の品詞に絞って、機能語から実質語への連続性を見ていく。第2章の表1においても、第3章の表3において

も、延べ語数が顕著に多かったのが助詞、助動詞、動詞、名詞の4者であった。しかし、この章では、助動詞は除いて分析することにした。その理由は形態素解析にある。形態素解析を行なうと、「のだ」や「ようだ」のような助動詞は1つの助動詞として認定されず、「の」と「だ」、「よう」と「だ」という2つの形態素に分割され、その結果、「の」と「よう」は助動詞だとは認定されなくなってしまう。「のだ」も「ようだ」も日本語にとって非常に重要な助動詞であるので、これらをうまくとらえることができていない「助動詞」というカテゴリーに焦点を当てて分析することは、あまり望ましくないことであろうと考えた。これが、助動詞を除いて、助詞、動詞、名詞のみを分析対象にする理由である。

分析にはKYコーパスを使用する。BCCWJを使用してもよいのだが、分析結果を日本語教育に還元することを考え、ここではKYコーパスを用いることにした。

前章の分析でもKYコーパスを用いたが、前章では、表2を見ればわかるように、「動詞-自立」「動詞-非自立」という2種類の茶釜のタグを統合して「動詞」という品詞を新たに設定し、分析を行なった。名詞についても同様に、「名詞-一般」「名詞-サ変接続」「名詞-非自立-一般」などの茶釜のタグを統合して「名詞」という品詞を設定し、分析を行なった。そのため、前章での分析に使用した「動詞」及び「名詞」には、日本語教師や日本語学者が一般には想起しないような語までが「動詞」「名詞」という品詞に含まれてしまっていた。そこで、この章では、動詞らしい動詞と名詞らしい名詞のみを分析対象にできるよう、前章で設定した「動詞」「名詞」というカテゴリーは使用せず、今一度茶釜タグに立ち返って、茶釜において「動詞-自立」及び「名詞-一般」というタグが付されている形態素のみを、それぞれ「動詞」及び「名詞」と考えて分析を行なうことにした。なお、「助詞」については、前章の分析で使用したものと同一ものを用いる。

「助詞」「動詞-自立」「名詞-一般」のそれぞれの形態素について、KYコーパスにおける使用回数をまとめたものが、次の表5である。

表5を見ると、KYコーパスの90人分のOPIデータの中で、500回以上使用されている形態素が、助詞には15語、動詞-自立には5語、名詞-一般には1語あることがわかる。500回以上使用された形態素の数は、助詞が最も多く、次に動詞-自立、その次が名詞-一般という順になっている。次に、使用回数が100～499回の形態素数を見ると、どの品詞もだいたい同じような数字ではあるが、動詞-自立が最も多くなっている。そして、使用回数が50～99回の

表5 KY コーパスにおける助詞・動詞・名詞の使用回数

| 使用回数 | 助詞 | 動詞 - 自立 | 名詞 - 一般 |
|-------------|-----|---------|---------|
| 500 回以上 | 15 | 5 | 1 |
| 100 ~ 499 回 | 18 | 20 | 18 |
| 50 ~ 99 回 | 9 | 16 | 27 |
| 50 回未満 | 109 | 823 | 2,674 |
| 合計 | 151 | 864 | 2,720 |

形態素数と 50 回未満の形態素数を見ると、名詞 - 一般が最も多くなっている。

前章及び前々章の分析では、それぞれの品詞が、最も機能語的な助詞・助動詞から最も実質語的な名詞まで概ね連続的に並んでいることが明らかになったが、上記の表5を見ても、最も機能語的な助詞から、中間的な動詞 - 自立を経て、最も実質語的な名詞 - 一般へと、数値が連続的に並んでいることがわかるのではないかと思う。

表5は、助詞、動詞 - 自立、名詞 - 一般のそれぞれについて、使用回数の多い形態素がどの程度あるのかを示したものであるが、次の表6は、それぞれの品詞について、具体的にどの形態素が多く使用されているのかを示したものである。それぞれの品詞について、使用回数の上位 15 位までの形態素を表6に記載した。なお、形態素の直後の () の中の数字は、その形態素の使用回数を示している。

表6を見ると、まず、助詞の使用回数が非常に多いことがわかる。上位1位から3位の「は」「の」「て」は、使用回数が5,000回を超えている。一方、名詞 - 一般は、第1位の「人」でも使用回数が597回であり、助詞15位の「よ」よりも少ない使用回数となっている。

このように見ると、動詞 - 自立は、やはり助詞と名詞 - 一般の中間に位置していると言える。動詞第1位の「する」は、様々なサ変動詞を作り得るのであるため、実質的な第1位とは言いにくい。実質的には、第2位の「ある」が、最も多く使用されている動詞であろう。ただし、使用回数を見ると、第3位の「思う」、第4位の「言う」と大きな差はない。使用回数は、「ある」が1,170回、「思う」が1,108回、「言う」が1,016回であり、これらは、助詞「けど」「とか」とほぼ同程度の使用回数である。動詞の中でのスターである「ある」「思う」「言う」は、スターぞろいの助詞の中に入っても、決して見劣りすることはないであろう。

表6 KY コーパスにおける助詞・動詞・名詞の使用回数ベスト15

| 順位 | 助詞 | 動詞 - 自立 | 名詞 - 一般 |
|-----|------------|------------|-----------|
| 1位 | は (5,267) | する (2,441) | 人 (597) |
| 2位 | の (5,225) | ある (1,170) | 先生 (385) |
| 3位 | て (5,038) | 思う (1,108) | 自分 (384) |
| 4位 | に (3,673) | 言う (1,016) | 日本語 (323) |
| 5位 | が (3,322) | なる (626) | 大学 (251) |
| 6位 | と (3,173) | 行く (494) | 友達 (219) |
| 7位 | ね (2,274) | わかる (436) | 学校 (211) |
| 8位 | か (2,177) | 見る (396) | 日本人 (191) |
| 9位 | を (1,939) | 来る (373) | 家 (148) |
| 10位 | で (1,735) | やる (324) | 子供 (145) |
| 11位 | から (1,631) | いる (298) | 学生 (138) |
| 12位 | も (1,612) | できる (294) | テレビ (135) |
| 13位 | けど (1,158) | 違う (238) | 国 (132) |
| 14位 | とか (1,033) | 住む (237) | 本 (114) |
| 15位 | よ (745) | 作る (175) | 経済 (113) |

5. まとめ

第2章から第4章までの分析において、助詞・助動詞とその他の品詞との間にはある程度の不連続が認められるもの、機能語から実質語への連続性は概ね示すことができたのではないかと思う。語は、機能語と実質語に二分されるものではなく、最も機能語らしいものから最も実質語らしいものまで連続して並んでいるととらえる方がいいのではないか、つまり、二分法よりもプロトタイプ論的な見方を採用する方がいいのではないか、ということである。

次に、日本語教育への示唆についてであるが、今後は、動詞に注目することが大切ではないかと思う。現在の日本語教育、特に初級の日本語教育を支配しているのは文法シラバスである。文法シラバスというのは、実際のところ、助詞・助動詞シラバスであると言ってよいだろう。第2章及び第3章の分析によると、延べ語数において顕著な品詞は、助詞、助動詞、動詞、名詞の4者であった。これらのうち、助詞と助動詞はすでに文法シラバスとして日本語教育の現場にしっかり入り込んでいるが、動詞と名詞は未消化なままである。

助詞と助動詞がシラバスになりやすい理由は、「延べ語数／異なり語数」の

大きさにある。助詞と助動詞は、表1と表3のいずれにおいても「延べ語数／異なり語数」の値が非常に大きかった。つまり、1つの助詞か助動詞を覚えれば、それを何度も繰り返し使うことができるということである。このような、コストパフォーマンスのよい品詞はシラバスになりやすい。一方、名詞は、表1と表3のいずれにおいても「延べ語数／異なり語数」の値が最小であった。つまり、ある名詞を覚えても、それを使うチャンスはあまりないということである。また、名詞は、人によって使うものが異なりやすい。趣味や職業などが違えば、まったく異なる名詞を使うかもしれない。シラバスとは、少なくともそのコースの学習者全員に共通すべきものである。人によって使用する語が異なる可能性のある品詞は、やはりシラバスにはしにくい。

このように考えていくと、文法シラバスの次に我々が作成すべきシラバスは、動詞シラバスであると言える。動詞は、「延べ語数／異なり語数」の値が助詞、助動詞よりは小さいが、名詞よりは大きい。また、表6を見ても、「ある」「思う」「言う」など、助詞と同じぐらい頻繁に使用されている動詞があることもわかる。今後は、動詞が日本語教育の鍵を握るのではないだろうか。

参考文献

- (1) 金澤裕之 (2004) 「第14回第二言語習得研究会 (全国大会) パネルディスカッション『語彙の習得』」 『第二言語としての日本語の習得研究』 7号
- (2) 亀井孝・河野六郎・千野栄一 (編著) (1996) 『言語学大辞典 第6巻 述語編』 三省堂
- (3) 寺村秀夫 (1982) 『日本語のシンタクスと意味 第I巻』 くろしお出版
- (4) 野田尚史・迫田久美子・渋谷勝己・小林典子 (2001) 『日本語学習者の文法習得』 大修館書店

謝辞

この論文は、2018年12月22日に行なわれた「第四回学習者コーパス・ワークショップ&シンポジウム『第二言語習得における語彙の役割』」での発表内容に、若干の修正を加えてまとめたものである。発表の機会を与えてくださった迫田久美子先生、ならびに、シンポジウムの席上で有益なコメントをくださったみなさまに期して感謝いたします。