

# 文章の難易度測定方法に関する研究

——「やさいちチェッカー」の「硬さ」について——

岩 田 一 成

1. はじめに
2. 文章の難易度を測る
  2. 1 リーダビリティ
  2. 2 「やさいちチェッカー」で判定したい文書の種類
3. 文章ジャンルの比較
4. 語彙密度と文名詞密度
  4. 1 語彙密度
  4. 2 文名詞密度
  4. 3 分析手続き
5. 分析結果と考察
6. おわりに

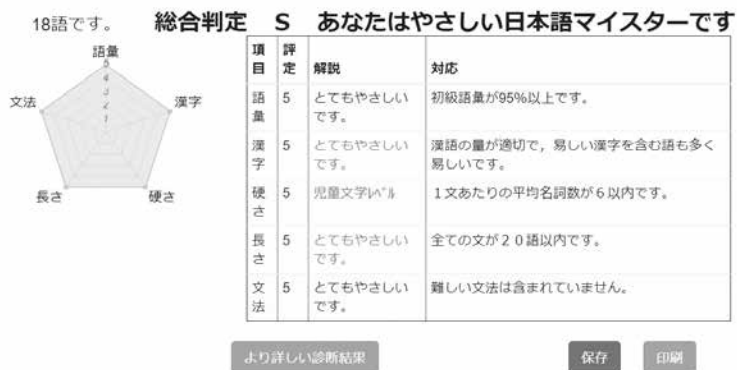
筆者の関わる研究チームは文章の難易度を診断する Web ツールとして「やさいちチェッカー」を公開している。ここでは、複数の指標で文章を評価するのだが、その一つに「硬さ」という項目がある。本論文では、その「硬さ」について論じる。「やさいちチェッカー」は公用文の難易度判定をその目的としているため、実質語の詰め込み具合を測定することが有効ではないかと考えた。公用文は、法律文の影響を受けており、たくさんの情報を詰め込んで一文を形成するという構造を保持しているからである。そこで、語彙密度、文名詞密度という二種類の測定方法を比較した。語彙密度は、節あたりの実質語の数であり、文名詞密度は文当たりの名詞の数である。両者を使って各種文章を診断してみると、文名詞密度が診断ツールには適しているという結論になった。その理由は、計算がシンプルであること、「白書」「法律」といった硬い文書で数値が高く跳ね上がること、異なったジャンルの文章間で数値のふり幅が大きいことの三点である。

## 1. はじめに

2019年に外国人労働者の受け入れが始まり、外国人材の受け入れに関する制度の整備が進んでいる。同年に出された『外国人材の受入れ・共生のための総合的対応策（改訂）』（外国人材の受入れ・共生に関する関係閣僚会議）には「やさしい日本語」の活用が明記されるようになった。それに対応する形で、2020年に『在留支援のためのやさしい日本語ガイドライン』（出入国在留管理庁・文化庁）が公開された<sup>1</sup>。

「やさしい日本語」の普及には、文章の難易度を測る目安が必要である。「この文章は難しい／やさしい」という評価を各執筆者の主観に委ねずに数値化することができれば、わかりやすさの基準が共有されやすくなるからである。上記ガイドラインには、文章の難易度が判定できるいくつかのツールが紹介されており、その中に筆者が関わっている「やさしちチェッカー」がある<sup>2</sup>。その「やさしちチェッカー」で用いられている「硬さ」という指標について論じるのが本論文の目的である。

### (1) 「やさしちチェッカー」の画像



「やさしちチェッカー」は指定の場所に文章を入力、またはコピーして診断を押すと、(1)の写真のような結果がでる。写真を見るとわかるように、語彙・漢字・硬さ・長さ・文法の5つの指標がレーダーチャートで出るようになっている。それぞれの指標で5段階評価を行い、総合判定も表示する。総合判定は各指標5段階評価の合計点を基にAからEで評価するが、おまけとして

S「あなたはやさしい日本語マイスターです」というものもある（写真参照）。解説欄のコメントにマウスカーソルを合わせると説明がポップアップし、その評定が出された根拠がわかる。また、解説のリンクをクリックすることで入力欄の該当箇所を着色表示することができる。例えば、漢字の解説部分をクリックすると、入力した本文の難解漢字が色付けされることになる。このレーダーチャートがバランスよく広がっていくと「やさしい日本語」に近づくことになる。

本稿では、この5つの指標の一つ「硬さ」という項目についてその計算のしくみを論じる<sup>3</sup>。まず、公用文の難易度判定に使う「やさしにちチェッカー」の目的を踏まえて、語彙密度、文名詞密度という測定方法を紹介する。その上で、両者を使ってさまざまな文章を数値化し、より適した指標はどちらかということ議論する。結論から言えば、「やさしにちチェッカー」は文名詞密度を採用しているが、文名詞密度が公用文の判定に向いている理由は最後に論じる。

## 2. 文章の難易度を測る

### 2. 1 リーダビリティ

文章の難易度を数値化する試みは、リーダビリティ研究としてアメリカでは1940年代から行われている（柴崎2014）。リーダビリティという用語を使わない場合もあるが、日本語でも様々なものが考案されている。ここでは、ウェブ上で公開されている以下の3つを紹介したい。ここに取り上げたものはほんの一部である。

#### (2) 本節で紹介する文章診断サイト

「リーダビティー・リサーチ・ラボ」

(<http://readability.nagaokaut.ac.jp/readability/>)

「日本語文章難易度判別システム」

(<http://jreadability.net/>)

「リーディングチュウ太」

(<http://language.tiu.ac.jp/>)

この分野の草分け的存在である「リーダビティー・リサーチ・ラボ」は、学年配当漢字、文字種、文節数、述語数などで判定している。学年配当、文節

などという用語からわかるように、学校教育現場の教員に広く使われているということである（柴崎2014）。

日本語教育関係者向けに設計されているものもある。「日本語文章難易度判別システム」は、初級の前半・後半、中級の前半・後半、上級の前半・後半と6段階で文章の難易度を判定している。ここで使われているリーダビリティは計算式  $[X = \{平均文長 * -0.056\} + \{漢語率 * -0.126\} + \{和語率 * -0.042\} + \{動詞率 * -0.145\} + \{助詞率 * -0.044\} + 11.724]$  で求められる。李（2016）は、この計算式に至る経緯を論じており、複数の公式を試して人間による主観判定に一番近いものを採用したというプロセスを公開している。

同じく日本語教育関係者向けの「リーディングチュウ太」は、旧日本語能力試験で公開されていた単語ごとの級を利用している。結果は単語ごとに1級、2級などの級が付与され（現在はそれをN1、N2と呼び変えて表示）、それらの総合で文章の難易度を判定している。級判定は、日本語教育関係者になじみが深いため、このサイトの知名度は高く多くの人に使われている。このサイトは、文章診断のほかにも様々な機能があり、川村（2009）にそれらの使い方が示されている。

ここで紹介したものは、公開されているものの一部であるが、目的に合わせてさまざまなものがあることを確認したい。そして、柴崎氏の論文では以下のように指摘している。

「このように文章の難易度を決定するための式は数多くあるが、数多くあるということはすなわち、どれが最適とは言えないということであり、それぞれに長短がある。従って、読み手の年齢や使用目的によって式を使い分けるのが賢明な方法であると言えよう。（柴崎2014：p52）」

本稿では、これらの文章診断サイトとは異なった目的で文章を判定するにはどういった手法があり、どれが適しているのかを考えていきたい。

## 2. 2 「やさち checker」で判定したい文書の種類

さまざまな目的で日本語の難易度を測る取り組みが公開されてきていることをここまで見てきた。本稿が扱う「やさち checker」は、国や自治体が発行する公用文の難易度判定をその目的としている。

公用文の難解さを考える際、これまで語彙の問題としてとらえる傾向があっ

た。例えば1970年代は、「可及的速やかに、善処する、前向きに検討する、遺漏のないよう…」といった役所言葉が批判の対象となり、各自治体は削減に取り組んだ。2000年代に入ると、「バックオフィス、アウトソーシング」といった外来語が問題視されるようになった。ところがお役所言葉をなくしても外来語を減らしても、文章構造自体が難解であれば読みやすくはならない。

岩田（2016）では語彙レベルの問題ではなく、談話レベルの問題として公用文の読みにくさを論じている。そこでは、長い文章、間接的な説明、法律文の借用といった談話レベルの特徴を指摘している。例えば以下の例文を見ていただきたい。

### (3)

国民年金に加入している人や保険料を納めた期間と免除された期間を合わせて25年以上ある人が亡くなったとき、その人によって生計を維持されていた子のある配偶者または子が受けられます（子は、18歳に到達する年度の末日までの子または、20歳未満で1級2級の障がい状態の子に限られます）。

ただし、老齢基礎年金の受給資格のない方がなくなったときは、死亡日の前々月より前の被保険者期間において、保険料を納めなかった期間が3分の1をこえないことが必要です。（平成38年3月までは、死亡日の前々月までの1年間に保険料の滞納がなければよいことになっています）

（N市『遺族基礎年金』、岩田2016：27より）

この文章は、お役所言葉や外来語を減らしても簡単にはならない。法律文をそのまま借用した文体で書かれており、さらに談話全体が整理されていないこともあり難解になっている。このように、公用文の難易度を判定するにはそこに潜んでいる法律文のニュアンスに反応できるような方法がいいと言えそうである。上に岩田（2016）で公用文の読みにくさの特徴として法律文の借用を挙げたが、これは文言をそのままコピーしている（3）のような難解文書を指す。一方、長い文章、間接的な説明などという他の特徴も、多かれ少なかれ法律文の影響を受けており、これは公用文全体に関わる特徴であると言える。

## 3. 文章ジャンルの比較

ここでは、「Yahoo!知恵袋」（以後知恵袋）「Yahoo!ブログ」（以後ブログ）「白書」「法律」といった各ジャンルから実例を紹介したい。「知恵袋」「プロ

グ」などは話し言葉に近く比較的やわらかい文章と言われるものだが、「白書」「法律」は反対に固い文章である。中納言<sup>4</sup>を用いて『現代日本語書き言葉均衡コーパス』(以下BCCWJ)を検索してみた。何か統一した基準で例文を抽出すべく、比較的広いジャンルで使われている「生活」という語をキーワードにした。中納言で検索し、比較的意味が取りやすい例を掲載している。ここではジャンル間の文章の違いを大雑把に把握するのが目的である。

#### (4) 「知恵袋」

(前文脈：お前は女じゃない、お前みたいな女を相手にする男がいるはずがない。楽しそうに笑いながら言っていました。)

そんなことを言いながら、男の人から電話がかかってきたりするとだらしのない生活してるから、そんなことになるんだと怒鳴られて大変でした。

〈サンプルID OC09\_04657 知恵袋 2005〉

#### (5) 「ブログ」

(前文脈：ところで、JSBA(日本スノーボード協会)から、雪坊主と一緒に更新手続きのお知らせがありましたよね。08/09年度から自動引き落としもはじまるそうで、振り込み手数料もないので、そっちで手続きしようかと思っています。)

が、生活費以外はネット銀行にしているんですよ。

〈サンプルID OY15\_03286 ブログ 2008〉

#### (6) 「白書」

対中国ODA 日本の対中国ODAは、中国沿海部のインフラ整備、環境対策、保健・医療などの基礎生活分野の改善、人材育成など中国経済の安定的発展に貢献し、中国の改革・開放政策を維持・促進させる上で大きな役割を果たしてきました。

〈サンプルID OW6X\_00117 政府開発援助(OA)白書 2005〉

#### (7) 「法律」

3 公安委員会は、次条第一項に規定する犯罪被害者等早期援助団体その他の犯罪被害等を早期に軽減するとともに、犯罪被害者等が再び平穏な生活を営むことができるよう支援することを目的とする民間の団体(第五項におい

て「犯罪被害者等早期援助団体等」という。)の自主的な活動の促進を図るため、必要な助言、指導その他の措置を講ずるように努めなければならない。(サンプルID OL1X\_00034 犯罪被害者等給付金の支給等による犯罪被害者等の支援に関する法律 1980)

各ジャンルの文章を見ると、「知恵袋」「ブログ」は比較的一文が短い、前文脈を掲載しないと意味がわからない(そのため前文脈を括弧の中に掲載している)。つまり流れで文章を理解することになる。一方、「白書」「法律」は一文が長い、その一文に情報が凝縮されており、(4)(5)に見られるような前文脈への依存は少ない。こういった文は、名詞や動詞といった実質語がたくさん文中に埋め込まれ情報が圧縮されていることがわかる。BCCWJを使って自・他動詞を分析している中俣(2020)では、「法律」の文章は他動詞の比率が非常に高いことを指摘しており、実質語が多くなる原因の一つであると言える。こういったことから、文における実質語の詰め込み具合がどうも文章の硬さに関わっている可能性がある。

#### 4. 語彙密度と文名詞密度

文中にある実質語の数を数えて各ジャンルにおける文章の特徴を数値化しようという取り組みは、Halliday(1989)、樺島(1990)、佐野ほか(2009)に見られる。単位を文レベルにするのか、節レベルにするのか、実質語の内訳をどこまでとするのか、など細かい点で異なるが、実質語の詰まり具合を数値化することの重要性は早くから議論されていたことがわかる。ここでは、語彙密度と文名詞密度という方法を紹介したい。なお、本稿では、実質語の詰まり具合を数値化して測るものを、文章の硬さと呼ぶ。

##### 4.1 語彙密度

Halliday(1989)が提案した語彙密度とは、実質語の密度を数値化したものである。日本語の分析もすでに大規模コーパスを用いて行われている(佐野ほか2009)。もともとHallidayの議論は話し言葉と書き言葉の違いを考える際の指標として提案されている。書き言葉は実質語の割合が高く、一方、話し言葉は機能語の割合が相対的に高くなるという特徴に注目しており、語彙密度の定義を引用すると「語彙密度は、節当たりの実質語の数で測りうる(LEXICAL DENSITY will be measured as the number of lexical items per clause.)」

(Halliday 1989 : 67) となっている。

その後、Halliday (1993a) では、書き言葉の中にもバリエーションがあることを指摘し、科学的な文書では一般の書き言葉よりも際立って高い語彙密度を持つことなどを指摘している。

ここからは少し具体例を示しながら考えたい。以下に2種類の文章を示す。両者を比べてみると、AよりもBのほうが硬い印象を受けるのではないだろうか。

(8)

- A クヌギの木は秋になると、葉が落ちて積もります。そこでカブトムシが卵を産みます。
- B クヌギの木は秋、落ち葉が堆積することでカブトムシの産卵場を提供します。

それぞれの文の語彙密度を計算してみると、以下のようになる。実質語には下線を引き、節の切れ目にはスラッシュを入れてある。そこで、実質語の総数を節の数で割れば語彙密度が出てくることになる。なお、実質語や節の認定方法については4.3で詳しく述べる。

(9) 語彙密度の計算

- A 語彙密度 2.25 (実質語 9、節 4)

クヌギの木は秋になると、葉が落ちて積もります。そこでカブトムシが卵を産みます。／

- B 語彙密度 9 (実質語 9、節 1)

クヌギの木は秋、落ち葉が堆積することでカブトムシの産卵場を提供します。／

このように、語彙密度を使うと文の硬さが数値化できるようになる。そしてAとBの違いを比較することができるのである。

#### 4. 2 文名詞密度

ここでもう一つの方法を紹介したい。文名詞密度とは一文当たりの名詞の数を数えるというシンプルなものである。これは、語彙密度をアレンジしたもので、単純な計算式で語彙密度の近似値を出せないかと考え筆者が考案した。た



だし、文章中の名詞を数えるという発想は、樺島（1990）にすでに見られ<sup>5</sup>、アイデア自体は新しいものではない。（9）の例で具体的に紹介すると以下のようになる。名詞に下線を引き、文末はスラッシュで区切っている。

(10) 文名詞密度の計算

A 文名詞密度3（名詞6、文2）

クヌギの木は秋になると、葉が落ちて積もります。／そこでカブトムシが卵を産みます。

B 文名詞密度9（名詞9、文1）

クヌギの木は秋、落ち葉が堆積することでカブトムシの産卵場を提供します。

ここでは文名詞密度を比較対象として計算する理由を述べたい。文の実質語において出現頻度が高いのは名詞なので、名詞だけを数えたらいいのではないかという発想が出発点である。（9）Bで実質語は9あることを見たが、（10）Bを見ればそのすべてが名詞であることがわかる<sup>6</sup>。

また、名詞化（nominalization）が、話し言葉と書き言葉の区別に関わっているという指摘（Chafe1985）や、文章の種類に関わっているという指摘をあげたい。名詞は硬い文章において大事な役割を果たしているのである。Halliday（1993b）は科学文書の分析において、シュレツペグレル（2017）は各種教科書の分析において、名詞化がアカデミックな文体の特徴となっていると指摘している。そこでは、ある文で新しい情報を提示し、次に続く文の主題が既出情報の再提示になる流れで、論理が展開されるとしているが、具体的に例で示すと以下のようになる。

(11) 秋に枯葉がたくさん落ちて積もる。落ち葉の堆積はカブトムシに住処を提供する。

下線部は前の文の「枯葉がたくさん落ちて積もる」という既出情報を主題として再提示していることになる。この名詞化という現象は英語の議論で提案されているものだが、日本語の場合、名詞化には漢語化が同時に起こり、一層文章の硬さに貢献することになる。（11）で言うと、「堆積」がそれにあたる。

名詞化に関連する議論として、文法的比喩（Halliday1989、1998）というも

のがある。これも、もともとは書き言葉と話し言葉の違いを分析するために提案された概念であるが、アカデミックな文章の特徴を説明する際にも用いられる。例えば、人間の言葉は人を含む「物事」は名詞、動作・行為などの「出来事」は動詞で表すという原則が日常言語では適用されている。Halliday (1989) の例を和訳して示すと「アナウンスの後、人々が拍手をした」では、「人々」という名詞に拍手という動作が動詞で表されている。一方、ややアカデミックな文では「拍手がアナウンスに続いた」となり、本来動作である「拍手」が名詞になっている。このあるべき品詞と単語のペアがずれることを文法的比喩と呼んでいる。

#### 4. 3 分析手続き

分析は、『現代日本語書き言葉均衡コーパス (DVD版)』(以下BCCWJ) を使って、ジャンル別に語彙密度と文名詞密度を計算した。コーパスを形態素解析にかけ、計算に必要な要素を抽出した。

語彙密度の計算方法は佐野ほか (2009) に従っている。実質語は名詞、動詞、形容詞、副詞などを検索し分子とした。分母としてカウントする節は、連用節、並列節などを含む一定範囲の節を基準として決めてその対象とした<sup>7</sup>。指定の節を自動認定するにあたり、「節境界検出ルール」(以後CBAP) を利用させていただいた。CBAPは、丸山ほか (2004) で説明がされているが、文章から節を切り分けてタグ付けをしてくれるツールである。これにより指定した節だけをカウントすることができるようになる。

文名詞密度は、計算がシンプルである。分子には名詞を取り、分母は句点の数でカウントした。名詞のカウントは、UNDICによる形態素解析で、名詞-普通名詞-副詞可能のような、名詞から始まる品詞を対象としている<sup>8</sup>。節で切り分ける必要がないためCBAPは使っていない。

#### (12) 使用ツール

形態素解析エンジン MeCab0.98

日本語辞書 ipadic2.7.0 (MeCa用)、UNIDIC2.1.2 (MeCab用)

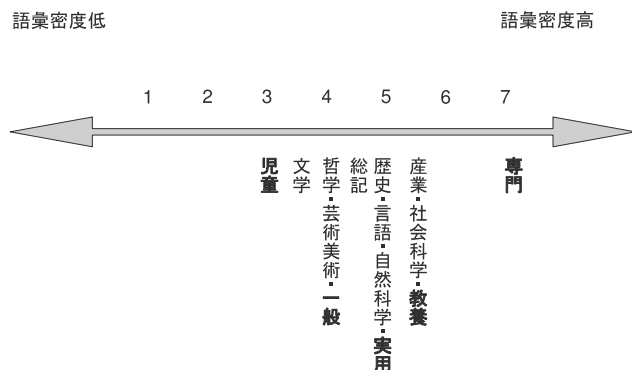
節認定 「節境界検出ルール」(CBAP\_1.0.pl)

\* CBAPを使用し、節数を求める際にipadic辞書を使用。節数以外の語数や名詞実質語数はUNIDICで形態素解析した結果を使用。

## 5. 分析結果と考察

本稿は、BCCWJを分析している佐野ほか（2009）と同じ基準で語彙密度を計算している。佐野ほか（2009）では、BCCWJを「児童、文学、…専門」のように区分して、それらの語彙密度を示している。

### (13) BCCWJの各分野と語彙密度（佐野ほか2009：50）



本稿は、この区分を踏襲しつつ、「白書」や「法律」なども加えて硬い文書の数値を出している。また、「知恵袋」「ブログ」といった比較的やわらかいものも分析対象とすることで、数値の幅を示そうと試みた。

語彙密度と文名詞密度の計算結果は以下のとおりである。この数値は岩田ほか（2015）で一部を公開しているが、全体像を示すのは初めてである。（14）にあるOC（知恵袋）やLBPB（LB：図書館、PB：出版・書籍を合算したもの）というのはBCCWJのジャンルである。「文学」、「自然科学」というのは、LBPBの細分類項目である。「児童」や「専門家」というジャンルはないのだが、LBPBというジャンルの書籍で「児童」や「専門」といった対象者指定がなされた書籍群を指している。

#### (14) 語彙密度と文名詞密度の比較

	知恵袋 OC	児童 (対象) LBPB	ブログ OY	文学 LBPB	専門 (対象) LBPB	自然 科学 LBPB	白書 OW	法律 OL
語彙 密度	3.0	3.3	3.8	3.6	8	7.1	9.9	9.6
文名詞 密度	3.8	4.2	4.6	4.9	11	8.5	15.5	18.7

結果を見れば、語彙密度は3～9.6に分布し、文名詞密度は3.8～18.7に分布していることがわかる。

文名詞密度の特徴は、「白書」や「法律」でその数値が跳ね上がるという点である。この特徴は非常に重要で、公用文の中にある法律文的な要素を測るには向いている可能性がある。名詞と句点を数えるだけなので、計算式がシンプルだということも大事なポイントである。また、文章判定に使うには、振り幅が大きい方が利用しやすいため、「やさちチェッカー」の「硬さ」という指標には文名詞密度を採用している。3.8～18.7の幅を5つに区分することで以下のように判定結果を出している（中島・岩田2019）。

#### (15) やさちチェッカーの「硬さ」に関する判定基準

- 1：文名詞密度が15超なら法律文レベル
- 2：文名詞密度が12超なら専門書と法律文の中間レベル
- 3：文名詞密度が9超なら専門書レベル
- 4：文名詞密度が6超なら児童文学と専門書の中間レベル
- 5：文名詞密度が6以内なら児童文学レベル

語彙密度や文名詞密度は、節（文）当たりの実質語の詰め込み具合を見ているだけなので、語彙自体の難易度が考慮されていない<sup>9</sup>。そこから文（節）の構造を見ているだけでは文章の難易度とずれるのではないかと、という疑問が沸いてくる。これに関しては佐野（2016）が興味深い議論を展開しており、語

彙密度とそこに出現する単語には関係があるというものである。例えば、「挽肉」と「調達」を比べたら、明らかに前者は語彙密度の低い文章に出現し、後者は高い文章に出現するということを指摘している。つまり、語彙密度を測れば、そこで使われている語彙の難易度もある程度連動して予測できるということになる。

## 6. おわりに

本稿では、文章の硬さを判定するために、語彙密度と文名詞密度の2種類の方法を比較してみた。「やさしちチェッカー」は公用文の難易度を測るためのものである。公用文には法律文に近いものがあること、「白書」や「法律文」は情報がたくさん詰め込まれていること、こういった特徴を考えると、語彙密度や文名詞密度が有効であると考えた。結果を見ると、計算がシンプルであること、文名詞密度は「白書」や「法律」で数値が跳ね上がること、ふり幅が大きいことがわかり「やさしちチェッカー」により向いていると判断した。

本稿は、文名詞密度が語彙密度よりも優れていると主張するものではない。本来は文よりも細かい節レベルで分析する語彙密度のほうがより緻密な分析に向いており、細かいジャンル間で違いを見たりするには有効であることが予測できる。状況に応じて使い分けるべきであろう。

今後は、現場のニーズに合わせて様々なテキストを分析していく研究が可能である。例えば、学校の教科書は国語よりも社会のほうが語彙密度も文名詞密度も高いのだが、そういった研究は教育現場への貢献が期待できる。

最後に柴崎（2014）を引用して締めくくりたい。

「リーダビリティ研究は真理の探究ではない。それは万人とは行かずともなるべく多くの人に違和感なく受け入れられる文章難度の基準を設定することであり、その目的は社会での応用である。(62)」

本研究を評価するのは、「やさしちチェッカー」の利用者各位であり、この文章診断システムが多くの人に受け入れられることを願っている。

## 謝辞

本稿の執筆にあたり、中島明則氏にBCCWJを分析してデータの処理を行っていただいた。また、株式会社ATR-Promotionsには「節境界検出ルール」

(CBAP\_1.0.pl) の使用を許可していただいた。さらに本研究は以下の補助金（共に日本学術振興会科学研究補助金による基盤研究（A））を受けて行っている。関わってくださったすべてのみなさまに感謝申し上げます。

- ・「やさしい日本語を用いたグローバルコミュニケーション社会実現のための総合的研究」（課題番号22242013）
- ・「やさしい日本語を用いた言語的少数者に対する言語保障の枠組み策定のための総合的研究」（課題番号25244022）

## 注

- 1 『外国人材の受入れ・共生のための総合的対応策（改訂）』  
<http://www.moj.go.jp/content/001311603.pdf>  
『在留支援のためのやさしい日本語ガイドライン』  
<http://www.moj.go.jp/content/001327230.pdf>
- 2 「やさしにちチェッカー」 <http://www.4414uj.sakura.ne.jp/Yasanichil/nsindan/>
- 3 5つの指標全体の説明については岩田ほか（2015）を見られたい。
- 4 国立国語研究所が運営するコーパス検索アプリケーション  
<https://chunagon.ninjal.ac.jp/bccwj-nt/search>
- 5 日常会話、小説、社説、新聞、俳句などをデータとして、自立語に占める名詞の割合を出している。そこでは、名詞の割合が高いほど文の長さが長くなる傾向を指摘している。
- 6 「提供する」については動詞の一部なのだが、機械処理上サ変接続が可能な名詞となる。
- 7 分子となる実質語の詳細な内訳は佐野ほか（2009：32）に一覧が載っているが、分母となる節の種類は同（21）に一部が掲載されているだけである。分析対象の節をすべて知りたい方は、代表執筆者佐野大樹氏のresearchmapから全種類がダウンロードできるようにになっている。  
[https://researchmap.jp/multidatabases/multidatabase\\_contents/detail/228720/89b778118cf9fe26dc4716bd73c04874?frame\\_id=486573](https://researchmap.jp/multidatabases/multidatabase_contents/detail/228720/89b778118cf9fe26dc4716bd73c04874?frame_id=486573)
- 8 対象となる名詞の詳細は中島・岩田（2019）に一覧がある。
- 9 「やさしにちチェッカー」では「硬さ」とは別に「漢字」や「語彙」の難易度を表示している。（1）「やさしにちチェッカー」の画像を参照されたい。

## 参考文献

- 岩田一成（2016）『読み手に伝わる公用文——<やさしい日本語>の視点から』大修館書店
- 岩田一成・森篤嗣・松下達彦（2015）『「やさしにちチェッカー」－公的文書の難易度を指標化するシステムの開発－』『日本語教育学会秋季大会予稿集』407-408
- 樺島忠夫（1990）『日本語のスタイルブック（新装版）』大修館書店

- 川村よし子 (2009) 『チュウ太の虎の巻－日本語教育のためのインターネット活用術』 くろしお出版
- 佐野大樹 (2016) 「語彙密度から見た語彙シラバス」 森篤嗣編 『ニーズを踏まえた語彙シラバス』 くろしお出版 79-93
- 佐野大樹・丸山岳彦・山崎誠・柏野和佳子・秋元祐哉・稲益佐知子・田中弥生・大矢内夢子 (2009) 『語彙密度を利用した『現代日本語書き言葉均衡コーパス』テキスト分類の試み』 特定領域研究「日本語コーパス」平成20年度研究成果報告書 (web公開)
- 柴崎秀子 (2014) 「リーダビリティ研究と「やさしい日本語」」『日本語教育』158, 49-65
- シュレツペグレル・メアリー (2017) 『学校教育の言語 機能言語学の視点』ひつじ書房
- 中島明則・岩田一成 (2019) 「書き換え支援システム 公用文をやさしい日本語に」庵功雄ほか編『「やさしい日本語」と多文化共生』ココ出版 161-172
- 中俣尚己 (2020) 「日本語母語話者は本当に自動詞を好むのか」江田すみれ・堀恵子編『自動詞と他動詞の教え方を考える』くろしお出版 25-42
- 丸山岳彦・柏岡秀紀・熊野正・田中英輝 (2004) 「日本語節境界検出プログラムCBAPの開発と評価」『自然言語処理』11-3, 39-68
- 李在鎬 (2016) 「日本語教育のための文章難易度に関する研究」『早稲田日本語教育学』21, 1-16
- Halliday, M.A.K. (1989). *Spoken and written language*. Oxford University Press
- Halliday, M.A.K. (1993a). Some grammatical problems in scientific English. In M.A.K.Halliday&J.R.Martin (Eds.), *Writing science: Literacy and discursive power* (69-85). Routledge
- Halliday, M.A.K. (1993b). The Analysis of Scientific Texts in English and Chinese. In M.A.K.Halliday&J.R.Martin (Eds.), *Writing science: Literacy and discursive power* (124-132). Routledge
- Halliday, M.A.K. (1998). Things and relations Regrammaticising experience as technical knowledge. In J.R.Martin & R.Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science* (185-235). Routledge
- Chafe, W.L. (1985). Linguistic differences produced by differences between speaking and writing. In D.R. Olson, N.Torrance, & A.Hildyard (Eds.), *Literacy, Language, and Learning: The nature and consequences of reading and writing* (105-123). Cambridge University Press

