

「やさしい日本語」を用いた公文書の書き換え ——多文化共生と日本語教育文法の接点を求めて

庵 功雄
岩田一成
森 篤嗣

1. はじめに

『在留外国人統計』（入管協会 2009）によると、平成 20 年度末の外国人登録者数は 221 万 7426 人で 10 年前に比べて 46.6% 増加しているとされている。地域に定住する外国人（生活者としての外国人）が増えるにしたがって、その人たちへの補償教育（cf. 山田 2002）をどうするかということが問題となってくる。このような背景から、「多文化共生」は現在の日本語教育におけるキーワードの一つとなっている⁽¹⁾。ただし、外国人側だけに日本語習得の負担を押し付けるのではなく、日本人側も「相手に伝わりやすい日本語とは何か」と常に自問していく必要があるだろう。本稿では、その取り組みの一環として「やさしい日本語」を用いて公文書を書き換えるという作業を行う。最終的な日日翻訳の機械化を目標として、今回の作業から見えてきた課題を論じる。

2. 多文化共生と日本語教育

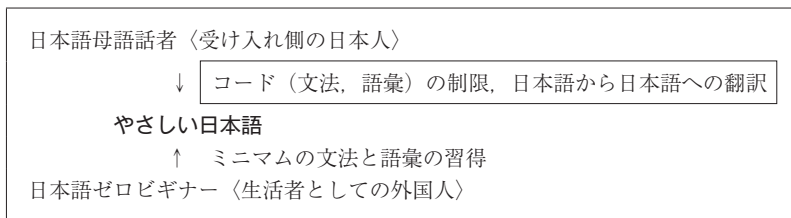
本稿で行った公文書の書き換えは、「日日ほんやくコンニャクプロジェクト（HK プロジェクト）」の一部として位置づけられる。本節では、そ

のプロジェクトにおける本論考の位置付けと、「やさしい日本語」の必要性について述べる。

2-1. 日日ほん訳コンニャクプロジェクト

HK プロジェクトは次の図のように「やさしい日本語」を媒介とした外国人と日本人のコミュニケーションを目指している。ここで示されているように、「やさしい日本語」は受け入れ側の日本人と生活者としての外国人の相互が学ぶものである。

図1 地域日本語教育における「やさしい日本語」の位置づけ



HK プロジェクトの具体的な目標は、「やさしい日本語」変換作業を機械化することと、地域日本語教育で使用可能な日本語教材を提供すること(『にはんごこれだけ!1』⁽²⁾を2010年出版している)の2点に分けられる。本稿は変換作業の機械化(図の囲み線の部分)の前段階として、公文書を書き換えるという作業を行い、その結果を基に文法面と語彙面に分けて、書き換え基準を提示する。そして作業から見てきた公文書に関する問題点を明らかにし、語彙の制限についても見通しを述べる。

2-2. 「やさしい日本語」の必要性

多言語表示に関しては田中(2009)で標準モデル4言語というのが指摘されており、日本語・中国語・韓国朝鮮語・英語の4言語が普及しつつあることを指摘している。この4言語だけでいいのかという議論もあるが、公文書でも同様にすべての情報を4言語で表そうとすると非常に大きなコ

ストがかかる。公文書の一部は外国語に訳されているものもあるが、今後ポルトガル語、インドネシア語など外国人住民の増加に合わせて言語を増やしていけば限りなく翻訳が必要になる。また、いくら翻訳を増やしても、全外国人住民の母語ですべての情報を提供することは不可能であろう。もう一つ確認すべきことは、英語の限界である。現在英語による外国人住民への情報提供が多くなされているが、国立国語研究所の調査結果によると英語ができる人は44%に過ぎず日本語の62.6%を下回るという報告がされている(岩田2010)。つまり、伝達効率を考えるなら英語より日本語が適しているということになる。このように多言語化の限界、伝達効率の高さなどから、日本語という選択肢が重要になってくる。さらにその日本語をやさしくすることでコミュニケーションを更に円滑化できるのではないだろうか。

3. 「やさしい日本語」に関する先行研究と本稿の立場

3-1. 先行研究1：簡約日本語

外国人が学習する際の便宜を考え、日本語のコード(文法、語彙)を制限するという最初の試みのひとつは野元菊雄氏による「簡約日本語」である(野元1990, 野元・川又・義本1991)。この試みの基本的発想は本稿も同じであるが、その応用が逐語訳的書き換えの域を出なかったことなどで厳しい批判にさらされることになってしまった。当時朝日新聞(1988. 2. 26夕刊)に掲載された簡約日本語の例である。

(1) 〈書き換え前〉 まず北風が強く吹き始めた。しかし北風が強く吹けば吹くほど、旅人はマントにくるまるのだった。遂(つい)に北風は、彼からマントを脱がせるのをあきらめた。

→ 〈簡約日本語〉 まず北の風が強く吹き始めました。しかし北の風が強く吹きますと吹きますほど、旅行をします人は、上に着ますもの

を強く体につけました。とうとう北の風は彼から上に着ますものを脱ぎさせますことをやめませんとなりませんでした。

「やさしい日本語」を考える際、「変な日本語を使う変な人、日本語を乱す人」といった風評が立たない（御園生・前田 2007：37）ものを目指す必要がある。そのためには、逐語訳の発想ではなく大胆な意識も必要になるであろう。またプレゼンテーションの際、文学作品を扱うことは避けるべきであろう。

3-2. 先行研究 2：減災のための「やさしい日本語」（減災 EJ）

「やさしい日本語」という名称からすれば本稿に先立つのが、災害時の外国人への情報提供の方法をやさしいものとするという目的から始まった「やさしい日本語」（以下、「減災 EJ⁽³⁾」と表記する）である（佐藤 2004, 2007, 御園生・前田 2007 など）。「減災 EJ」はその有効性についても検証されており（松田他 2000）、ニュースをやさしくすることで内容に関する質問の正答率が 30% から 90% に上がるという指摘をしている。そこではニュースという音声情報を使っているので、ポーズ、スピード、繰り返しといった読み方に関する配慮も挙げられている。

結論から言えば、「やさしい日本語」と「減災 EJ」は共通の目的意識を持っていると言える。ただ、「減災 EJ」では、災害場面を想定している点、「やさしい日本語」を日本語能力試験 3 級レベルと想定している点（水野 2006）⁽⁴⁾、松田他（2000）の検証作業に端的に現れるように音声言語を主な対象としている点の特徴であり、本稿の立場と異なる。

3-3. HK プロジェクトにおける「やさしい日本語」

公文書をやさしくするということは、書き言葉を扱うため「減災 EJ」のようにポーズ、スピード、繰り返しなどを調整して伝達することはできない。また、震災場面だけではないので内容のバリエーションが様々なも

のになりうる。よって、公文書のための「やさしい日本語」は言語そのものをやさしくし、どんな内容にも対応できるようにできるかぎり客観的に規格されたものでなければならない。またそのレベルに関しても慎重に決定されるべきである。

「やさしい日本語」は、1. で見た「簡約日本語」のようにコードを制限しすぎると大きな批判を浴びることとなる。しかし、2. の「減災 EJ」のように初級を 3 級レベルに想定するのはハードルが高すぎると考える。なぜなら留学・就学ビザの外国人以外は基本的に学習時間の確保が難しく、3 級レベルが必要とする 300 時間の学習を満たすことは難しいからである。

3 級レベルは初級レベルとも言い換えられるが、そもそも初級文法というものは、構文・文型に関わる項目が網羅的にリストアップされているため、内容が非常に多くなっている。小林（2009：28）では、「初級（3・4 級レベル）では、助詞や活用などの日本語の基本的な文型に関わる要素をひと通り教える。→中級（2 級レベル）以降では、複合辞や機能語を教える。」といった文法教育観が指摘されている。こういった現状を踏まえ、すでに庵（2009）では一機能一形式を基準に初級文法を軽くする試みが提案されている。2-1. 節で示した図に「ミニマムの文法」とあるが、それが庵の提案する文法（STEP1 レベル、STEP2 レベル）であり、いわゆる初級文法と比べると項目は大幅に減らされている。以下に引用する通りであるが、STEP1 レベルと STEP2 レベルで初級という位置付けがされており、これらは話し言葉が想定されている（以後、まとめてこの一覧を指す時は「STEP1・2」と呼ぶ）。

ステップ 1 文法項目	
動詞文	…ます／ません …ました／ませんでした
名詞文	…です／じゃないです …でした／じゃなかったです
形容詞文	…です／くないです …かったです／くなかったです
	王さんは主婦ですか？ はい、そうです。いいえ、違います。
〈応答〉	昨日、会社に行きましたか？ はい（行きました）。いいえ（行きませんでした）

	～ので (理由)・なので, ～ために／～ように／～ための (目的)
その他	～んです (どうして…んですか? ～からです.)
理解レベル	
モダリティ (対人)	～てもいいです (許可), ～てはいけません (禁止), ～ましょう (勧誘), ～たほうがいいです (当為), ～なさい (命令)
その他	昨日買った本 (はこれです.) (名詞修飾), 田中さんが来るか (どうか) (教えてください.) (名詞化)

従来の初級文法と比べると使役・受け身・可能・尊敬といった動詞の変化が全く扱われていない点, 推量の「らしい・ようだ・そうだ」も「と思います」で一本化されている点, 「と・ば・たら・なら」が「たら」で一本化されている点, 多くの文法項目が圧縮されており, 軽くなっていることが伺える。

3-4. 本研究の目的と特徴

地域で生活していく際に必要な情報が記された公文書⁽⁵⁾が「やさしい日本語」で書かれていれば, 外国人が被る不利益はかなり軽減されると思われる⁽⁶⁾. 日本語をやさしくすることは決して外国人だけのためではなく, 日本人の高齢者なども利益を受けることである. このように情報弱者に対しても平等に情報を伝えることができるようにすることがHKプロジェクトの最終目的であるが, その第一歩として本稿では「やさしい日本語」を用いた書き換えを試行し, そこから書き換え基準を提示する. 同様の動きとして, 森 (2006) は「四日市市パブリックコメント手続き条例」を例として書き換えを行い, 福生市 (2009) も生活便利帳を「やさしい日本語」で書いている. どちらも単独の文書を書き換えているのみであり, 複数の文書の見出しから書き換え基準を抽出しているというものではない. ここまでの先行研究との違いを考えると, 本稿は以下のように特徴付けられる.

- i 文法・語彙に関する書き換え基準を公文書の変換作業から帰納的に抽出する
- ii 書き換え基準は、機械化に向けて可能な限り客観的に記述する
- iii 「やさしい日本語」を文法に関しては3級レベルより低く設定する
(庵 2009 参照)

本稿に先立ち、2009年5月29日に広島市安佐南区沼田出張所にてちらし(本稿で言う公文書)をすべて収集し(集収88件)、その中から以下の10件を書き換えるの対象として選んだ。その際、知らないと生活に不利益を被る情報であるかどうかを基準とした。

- | |
|--|
| <ul style="list-style-type: none">① みんなで備えよう新型インフルエンザ② テレビについて大切なお知らせです(地デジについて)③ 定額給付金の給付が始まります④ 景気の後退・雇用情勢の悪化などにより生活に困っている人へ⑤ 広島県の最低賃金⑥ 広島県県営住宅入居者募集⑦ 臨時職員(看護補助者)募集案内⑧ 市民相談窓口のご案内⑨ 新しく地域の一員となられるみなさまへ(町内会・自治会のお知らせ)⑩ 広島市暴力被害相談センターご案内 |
|--|

4. 書き換え結果と書き換え基準

公文書の書き換えを「やさしい日本語」で行う場合、文法についてはすでに庵(2009)でミニマル文法が提示されており、それに基づいて書き換えればよい。しかし、語彙をどこまで制限するかはこれからの検討課題である。先述したとおり、減災EJでは、「やさしい日本語」を日本語能力試験3級レベル(300時間の学習)と想定している(水野2006)が、本稿では3-3.に述べたように、地域日本語教育が中心となる「やさしい日本語」では、300時間の学習を前提にするのはハードルが高すぎる⁽⁷⁾。しか

し、その3級レベルですら、国際交流基金・日本国際教育支援協会(2002)に掲載されている1,409語のリストを含む1,500語ということになり、その語彙数はかなり制限される。例えば、前節で示した公的文書の原文と書き換え文の語彙を、3級1,500語のリストでチェックしてみる。四角囲みは固有名詞、斜体が3級語彙にあるもの、下線が3級語彙にないものである。

- (2) 広島市では、市民相談センターにおいて、以下に掲げる事例のような「どこに相談したらいいかわからない。」という日常生活における法的な困りごとについて、……
- 広島市の市民相談センターでは、日常生活で法律に関係する問題(下に例を示します)について相談を受けます。

書き換え前、書き換え後ともに3級の1,500語に入らない語彙が多く見られる。例えば、「法的な」は「法律」に書き換えることで3級の1,500語をクリアしているが、一方で「事例」から「例」、「掲げる」から「示す」は、直観的にはやさしく書き換えられているように思うが、3級の1,500語はクリアできていないことになる。また、「日常」は書き換えられていないが、3級の1,500語の範囲外である。このように、例ではあるが「例」、「示す」、「日常」などが公的文書の書き換えに使えないというのは、非常に不自然な日本語になる恐れがある。こういった点に注意しながら本稿の書き換えにおいては、まず直観的に3人の執筆者がやさしいであろうと考える語彙に書き換えを行った。今後語彙をどのように規定していくかについては次の節で述べる。本節では文法と語彙に分けて、書き換え基準を提示する。

4-1. 文法的な書き換え基準

ここでは公文書を書き換えるに当たり、文法面での書き換え基準を示す。

なお、文法の基準となるのは庵（2009）で示された文法項目である。3-3. に一覧を載せてあるのでそれを見ればわかるが、敬語はそもそもこの項目に入っていないので、埼玉県（2006）にあるような「敬語を使わない」という書き換え基準ははじめから不要である。以後、矢印の先に指示されているものが我々の書き換え例である。

以後、現段階で提示できる文法上の書き換え基準を示す。これがすべてであると主張するものではないが、こういった基準を積み重ねていかねばならない。

i 複合述部→副詞+述部

佐藤（2004）では複合動詞は誤解を招きやすいという指摘が紹介されている。今回の書き換え作業では複合動詞の例は少なく、複合述部とした。書き換え案の方が日本語としてすわりが悪くなるのは事実であるが、ここで確認したいことは、「副詞+述部」という構造に書き換えると、辞書を使って読むことが可能になるという点である。「ておく」などの補助動詞を辞書で調べるには、「おく」が本動詞に付加された成分であることを理解しないといけない。こういった知識が無くても、「副詞+述部」であればより簡単に辞書を引けるはずである。

(3) 備蓄しておきましょう。→前に買いましょう。

(4) 家賃が高すぎる。通勤に時間がかかりすぎる。→ひじょうに高い、
かかる

ii 言いさし・体言止め→述部まで明示

文が完結しない表現は、我々は欠けている部分を補いながら解釈するわけである。しかし、必ずしも短い文が親切であるとは言えないであろう。述部まではっきり明示することで情報の内容を誤解なく伝えられるようになると思う。この点に関しては、本当に読みやすくなるのかという視点

で、検証作業も必要になってくるであろう。

- (5) 住みよいまちを！→住みよいまちを作りましょう！
- (6) すぐに 110 番通報！→すぐに 110 番に通報しましょう！
- (7) 割増賃金を支給→給料を少し多く払います

iii 排他文→非排他文

埼玉県（2006）、水野（2006）では「二重否定文を避ける」というルールが示されているが、厳密には否定を二重に行っているわけではないので排他文とする。上記先行研究と指摘内容は共有する部分が多い。ある一定の範囲を指定して、それ以外を否定するという表現をとることで、その範囲を明確に伝えようとするものである。

- (8) 定期募集以外の申込みには必要ありません。→定期募集の申し込みに必要です。
- (9) 高炉によらない製鉄業等を除く→高炉を使う仕事だけ
- (10) 18 歳未満又は 65 歳以上の労働者には適用されません。→18 歳から 64 歳までの人にあてはまります。

iv 長文→ナンバリング・箇条書き

松田（1999）、森（2009）にも同様の指摘があるが、1 文が長い場合は、ナンバリングや箇条書きで細かく分解していくと読みやすくなると考える。

- (11) 接触感染とは、例えば、感染した人の咳などが付着した手で、ドアノブなどに触れた後に、他の人がそこに触れ、かつ、その手で自分の口や鼻などに触れることによって感染することです。
→「接触感染」は、次のようなことです。
 - 1) インフルエンザにかかった人の咳が手につきます。

- 2) その手でドアのノブなどに触ります。
 - 3) 他の人がそのノブなどに触ります。そして、その手で自分の口や鼻に触ります。
- このようにして、インフルエンザにかかることが「接触感染」です。

v 連体修飾（外の関係，非制限的連体修飾）は解体する⁽⁸⁾

連体修飾については，埼玉県（2006），松田（1999）にも同様の指摘があるが，公文書において，すべての連体修飾を解体することは実質上不可能である。本稿では内の関係（寺村 1975）については認める。しかし以下のように外の関係の⁽¹²⁾，非制限的連体修飾と呼ばれる⁽¹³⁾のような例は，非常にセンテンスが長くなってしまい理解の妨げになっているのではないかと考える。外の関係は，⁽¹²⁾の例のように修飾部分を括弧に入れてしまえば簡潔に表現できる。その際，すべてが必要な情報かを吟味する必要がある。「どこに相談したらいいのかわからない」といった部分は削除してしまっても大きな問題はないであろうと考える。この削除については5節で詳しく取り上げる。⁽¹³⁾のような非制限的連体修飾の例は連体修飾部分を取ってしまっても本質的な意味は変わらないので，取ってしまった修飾部分を別のセンテンスとして後ろに書けば解体が可能である。

- (12) 広島市では，市民相談センターにおいて，以下に掲げる事例のような「どこに相談したらいいかわからない。」という日常生活における法的な困りごとについて…

→広島市では，市民相談センターにおいて，日常生活で法律に関係する問題（下に例を示します）について…

- (13) 誰もが安全で快適に暮らせるまち。このための身近な組織としてご近所のつながりのできた町内会・自治会があります。

→誰もが安全で快適に暮らせるまち。このために町内会・自治会があります。自治会は身近な組織としてご近所のつながりできています。

これらの書き換えを行ってみると、基本的に庵（2009）の「STEP1・2」で書き換えは可能である。ただ、公文書の細かいニュアンスを伝えるには、もう少し文法項目を加えてもよさそうである。これについて次節で述べる。

4-2. 文法書き換えに必要な追加項目⁽⁹⁾

ここまで文法に関する書き換え基準について論じてきたが、少し視点を変えて、庵（2009）が提示する「STEP1・2」文法項目の検証作業を行いたい。まずは書き換え例の一つ挙げる。

- (14) なお、募集する住宅名、受付日時、受付機関などについては、別冊「県営住宅募集一覧」をご覧ください。
→募集する住宅の名前、受付の日と時間、受けをする場所は、別冊「県営住宅募集一覧」を見てください。

下線部の「などについて」は「など」「について」両方とも「STEP1・2」の範囲外であり、書き換え例では省いてしまっている。若干ニュアンスの違いはあるものの、このように省略してしまっても本質的な意味は変わらない。しかし、両形式は公文書に頻繁に使われるものであり、また細かいニュアンスを伝えるためにも、文法項目として追加してもいいのではないだろうか。

- (15) 申込みは、1世帯につき1戸のみ申し込むことができます。
→1申込みは、1世帯が1戸（2戸、3戸はダメです）申し込むことができます。

このような書き換えは、下線部「のみ」が「STEP1・2」の範囲外であるために書き換え例は長くなってしまっている。「のみ」は汎用性が低い

が「だけ」であれば、追加してもいいのではないかと考える。

このように、公文書に頻出する、こまかいニュアンスを加える、書き換えると長くなってしまおうといった文法項目は、書き言葉専用の文法項目としてSTEP R (eading) というものを立て、今後追加していく作業が必要になる。「STEP1・2」が話し言葉を想定しているものであることを考慮すれば、この追加項目の必要性は、ある意味当然の要求ということになる。今回の書き換えで必要性が指摘できるのは以下の項目である。

～である、～してみる、～ことになる、～について、～までに、～だけ、～前に、～後に、～場合、や、と

4-3. 語彙の書き換え基準と「方針」

今回の書き換え作業を行うにあたっては、「程度」を「ほど」に、「利用する」を「使う」に書き換えるなど、漢語を和語に開くと簡単になるのではないかという書き換え基準を考慮しつつ進めた。今回の書き換え作業の最終目的が機械翻訳であることから、ルール化を進めることは必須であるが、語彙によっては一律に漢語を和語に開いてしまうと公的文書の書き換えとして、余計にわかりにくくなる恐れもある。6節で詳しく述べるが、これらは厳密な書き換え基準と言うよりも、書き換え方針という意識により、今後人海戦術で書き換えを進めていき、蓄積したデータから、現実的な書き換え基準を抽出するという手順を考えている。今回の書き換え作業から見えてきたいいくつかの書き換え方針について言及しておく。

i 固有名詞

固有名詞もできるだけ書き換える。このとき、書き換えた後に（ ）で固有名詞を残しておく。また、固有名詞に限らず、（ ）の使用ルールは「原語を（ ）で括る」に統一する。例：書類には、あなたの家賃の支払いなどを県に保証する人（連帯保証人）2人の印が必要です。

ii 複合名詞

漢語やカタカナ語の長い複合語は、積極的に分割して書き換える。

iii 語彙を開くときに文法が使える場合

例えば、「減免」の書き換え案として「安くしたりタダにしたり」のように、文法項目を用いて書き換えられる場合には、庵（2009）で示された文法項目を用いる。

iv 助数詞

助数詞に関しては、汎用性が高いと考えられるので、おおよそ残す。

v 接頭辞・接尾辞

「無収入」の「無」や、「入居者」の「者」のような接頭辞や接尾辞は、辞書で引くにくいので、分割して書き換える。

ここまでは、どのように書き換えたら日本語は読みやすくなるのかという問題を文法面と語彙面に分けて論じてきた。ただし、本作業でわかったことは、書き換え基準だけでは日本語をやさしくすることはできないという事実である。そもそも書き換え基準以前に議論すべきである2つの問題を取り上げその対応策を考察する。一つは、情報内容や伝達方法に関する問題であり、もう一つは書き換えの大前提となる語彙数をどう制限するかという問題である。それぞれを5.節と6.節に分けて論じる。

5. 公文書の伝達方法・伝達内容に関する問題

書き換え基準で書き換える以前の問題として、公文書の伝え方には様々な問題点が見られる。責任の所在を明らかにしないように敢えて断定表現を使わない以下の例は、断定表現にしてもかまわないはずである。

(16) 季節性のインフルエンザと同じ飛まつ感染と接触感染であると考えられています.

→季節性のインフルエンザと同じ飛まつ感染と接触感染です。

そもそも公文書を作って配布してまで伝えたい情報であるならば、責任の所在がわからないようなものは採用すべきではないのではないだろうか。

次の例は丁寧な文を心がけた結果こうなるのであろうと予測するが、下線部の情報が後件と完全に重複している。下線部を削除しても伝達情報に問題はない。こうやって長くなることで意味が取りにくくなることをここでは指摘したい。

(17) 「入居資格本審査に必要な書類」(4ページ参照)は、入居資格本審査日において入居候補者及び補欠順位者に提出していただく書類となりますので、申込みの段階では必要ありません。→下線部削除

次の2例は共通しており、法律の内容を過不足なく伝えることに集中しているため、(18)のように伝達内容が長くなったり、(19)のように伝達内容がわかりにくくなったりしている。

(18) 二輪自動車小売業（原動機付自転車を含む）を除く →下線部削除

(19) 入居者が60歳以上の方又は昭和31年4月1日以前に生まれた方であり、かつ、同居し又は同居しようとする親族のいずれもが60歳以上の方若しくは昭和31年4月1日以前に生まれた方又は18歳未満の方である世帯。 →53歳以上の人…

50ccのバイクを自転車であると解釈している一般市民はいるのであろうか。法律上での規定をそのまま表現すると(18)のような長い文になってしまう。(19)はもっとひどい例で、下線部は簡単に言うと、書き換え例のよう

に「53歳以上の人」となる。しかし、過渡期にある法律の旧版と新版の内容を過不足なく伝えようとするとこのような表現になるということである。これは筆者が複数で話し合っても内容が理解できなかったため市の担当者に電話をして確認したところ返ってきた答えである。なお、その担当者も最初は「私もよくわかりません」と答えていたため別の担当者に替わってやっと答えが返ってきたという経緯がある。ここで指摘したいことは、こういった伝達方法に関する問題は、もはや外国人のためではなく、日本人にも関わってくるということである。

また、今回書き換えた中には、紋切り型でほとんど内容がないものも見られる。以下の例は、地デジがどうして必要なのかという説明であるが、「やさしく便利なテレビ放送」とは一体何なのかわからない。

(20) 地デジは、今までのテレビ放送よりきれいな映像か楽しめるだけでなくあなたにやさしく便利な 21世紀のテレビ放送です。

→下線部にほとんど伝達情報がない。あいまい表現。

こういった伝達内容自体が不明の表現は、いくら「やさしい日本語」で書いてもやさしくはならない。これらを改めるには、書き手の認識を変えるような働きかけが必要である。今後の機械化に向けては、書き換え基準だけで書き換えていくのではなく、削除といった選択肢も用意して書き換え作業を継続していく必要がある。

6. 語彙数の問題と対応策

今回の作業では、4. で述べたように我々執筆陣の直観によって、やさしくなるであろうという語彙に変換した。しかし、機械化を目指す以上、客観的な語彙制限を確立していかなければならない。公的文書の伝達目的は、「書き言葉で過不足なく情報を伝達すること」と考えられる。その意味で、

語彙制限を厳しくした結果、過不足のない情報伝達が不十分ということになってしまうと本末転倒である。

そこで、今回の10文書の書き換え作業において、原文に比べてどのような語彙制限が行われたのかについて、まずは概観し、その結果から語彙数の問題について指針を述べたい⁽¹⁰⁾。

表1 10文書の書き換えにおける基本情報と累積頻度

	総文字数	総単語数 ⁽¹¹⁾	異なり語数	累積頻度80% % 超えの単語順位	累積頻度90% % 超えの単語順位
原文	11,808	6,830	1,093	258	494
書き換え	11,743	6,771	1,005	237	447

今回の書き換えた分量は、原文で10文書、11,808文字にとどまるため、異なり語数でも1,093語と少なく、公的文書における書き換への語彙制限を語るのに十分とは言えない。しかし、原文で見ると、異なり語数が1,093語のうち、約1/4の258語で累積頻度が80%を超え、90%超えでも約1/2で済むということは興味深い。書き換えについては、総文字数、総単語数、異なり語数のいずれにおいても、原文と比べて数字としての削減は感じられず、語彙に関してはもう少し大胆な語彙制限を意識して書き換える必要があるだろう。

それでは、書き換えによって頻度はどのように変化したかについても見てみる。原文でも書き換えでも頻度1位は「の」であり、これはどのようなコーパスでもほとんど同じ結果になると思われる。特徴的なのは、「する」である。原文で「する」は頻度3位で155回出現するが、書き換えでは頻度12位まで下がり、103回出現にとどまる。これは書き換えによって、漢語サ変動詞が和語動詞に置きかえられた結果である。「又は」は、原文では頻度24位で出現46回と高頻度だが、書き換えでは頻度219位で5回まで抑えられている。これは書き換え時の見落としである。下記に書

き換え後の「又は」の例を示す（全部で5回だが、一つは(23)と同表現が別の部分で重複出現）。

- (21) 受付印のあるもの又は収支明細書
- (22) 婚約者又は結婚はしていないけれども一緒に住んでいる男女
- (23) 障害の程度が恩給法の特別項症から第6項症の方又は第1款症の人
- (24) 他の家族と炊事場又は便所等を一緒に使っている

いずれも、「か」で置きかえることが可能である。つまり、「又は」のような、機能語の場合は語彙制限により、書き換えで使わないことは可能である。

また、実質語では「世帯」が原文では頻度31位で37回出現しているが、書き換えでは頻度165位で7回出現にまで抑えられている。残っている7回は「又は」と同じく書き換え時の見落としと思えるものもあるが、他の語彙で言い換えにくい文脈であったものもある（全部で7回だが、一つは(26)と同表現が別の部分で重複出現）。

- (25) 多回数落選世帯については、21ページの「10 選考方法」を見てください
- (26) 母子世帯
- (27) 心身障害者世帯
- (28) 原爆被爆者世帯
- (29) 引揚者世帯
- (30) 炭坑離籍者世帯

原文で「世帯」となっていて、書き換えた例を見てみると、「家族」がもっとも多い（例えば、「未就学児世帯」→「小学校にまだ入らない子どもがいる家族」）。これに従えば、(25)は「何回も抽選に落ちた家族」、(26)は

「母と子だけの家族」などのように言い換えられる。「家族」への言い換えを採用すれば、「世帯」を完全に語彙から外してしまうこともあり得るが、本当にそれで過不足のない情報伝達と言えるかは検討が必要であり、「世帯」のような公的文書に特有な専門語彙とでもいう語をどう扱うかが今後の課題である。このように具体的に語彙制限について課題を持ち、検討する必要がある。

さて、それでは本稿の書き換え作業を通して、語彙数にどのような目安を立てるべきであろうか。一つには、書き換え前の原文において、異なり語数が1,093語のうち、約1/4の258語で累積頻度が80%を超え、90%超えでも約1/2で済むということ、すなわち一定の語彙数における被覆率に注目したい。林(1971:9)では国立国語研究所の現代雑誌九十種調査をもとに被覆率を産出しており、5,000語で81.7%、10,000語で91.7%という結果を示している。同様に、松田ほか(2010)においても、Wikipedia、Yahoo!知恵袋、日本経済新聞、livedoor Blogというタイプの違う4つのコーパスを組み合わせ、上位5,000語の被覆率の調査を行った。この上位5,000語による各コーパスの被覆率は、Wikipediaが82.26%、Yahoo!知恵袋が87.33%、日本経済新聞が89.68%、livedoor Blogが87.98%であった。

これらの結果から考えると、公的文書は、日本語能力試験3級の1,500語や、簡約日本語や国立国語研究所(1984)の2,000語のレベルでは、公的文書の書き換えにあたっては、被覆率80%の達成は難しい。せめて、6,000語(国立国語研究所(1984)の基本語六千、日本語能力試験2級)ないし10,000語(専門教育出版(1998)、日本語能力試験1級)のレベルで書き換えを行うべきである。

しかし、公的文書の書き換えに6,000語~10,000語の語彙を用いると、「やさしい日本語」と言えないのではないかという批判があろう。そこで、この6,000語~10,000語の語彙を収録した外国人向けの辞書の作成を考えている。この辞書の特徴は、

1. 語彙の意味記述のメタ言語を約 2,000 語レベルに制限する
2. 語彙と文法が同じように 50 音順で引ける

の大きく二点である。まず、1 についてであるが、公的文書をいきなり 2,000 語レベルに書き換えると、過不足のない情報伝達に支障を来すと考えられ、一方で 6,000 語～10,000 語レベルだと難しすぎるというジレンマがある。そこで、書き換えについては 6,000 語～10,000 語を用い、その書き換えを 2,000 語で記述された辞書で調べながら読むというイメージである。次に 2 については、例えば「した方がよい」を辞書で引くとき、普通は辞書で「し」の項目を引いても載っていない。そこで、語彙と文法を同時に引ける辞書が外国人向けの辞書としては適当であると考えられるからである。

そして、この外国人向け辞書は、機械翻訳のための基礎語彙データとして作成し、並行して紙ベースの辞書としても作成したい。前者の機械翻訳のための基礎語彙データは、過去に作成された公的文書を書き換えるために用いる。具体的なイメージとしては、10,000 語レベルで書き換えた文書が html などで表示され、カーソルを合わせるとポップアップで、この辞書での解説が表示されるといった仕組みを考えている。後者の紙ベースの辞書は外国人が引くだけでなく、公的文書を作成する際の指針にも使用できるように配慮したい。

本稿の「やさしい日本語」における公的文書の書き換えにおける語彙のシステムは、以下のようにまとめられる。

- a. 『にほんごこれだけ！』レベル 800～1,000 語
- b. 外国人向け辞書記述レベル 1,500～2,000 語
- c. 公的文書書き換えレベル 6,000～10,000 語

7. まとめと今後の展望

本稿では公文書の書き換え作業を通して見えて来た書き換え基準を4節で提示し、書き換え以前に問題となる伝達内容・方法についての問題点を5節で、語彙数に関する見通しを6節で考察した。本稿の主張は、書き換え基準を文法、語彙別に増やしていくだけでは「やさしい日本語」にはならないというものである。今後も客観的な記述としての書き換え基準は増やしていかなければならないが、同時に5節で指摘した伝達方法や伝達内容についても書き手が内省しなければならないであろう。受け身などでごまかしたりする表現は避ける、無くても困らないような情報はカットするなど、公文書の在り方自体を議論しなければならない。またその結果は、行政側に助言していくべきであろう。語彙数に関しても、今回は目安を示したに過ぎない。本格的な議論は今後継続して行っていく。ここでもう一度確認しておきたいことは、このような公文書の書き換えは外国人だけのためではないということである。5節で見たような例は、日本人の高齢者には理解が難しいものであろうと想定できるし、中には筆者達にも理解不能なものがあった。

今後の課題を述べると、まず、本稿で言う「やさしい日本語」が、本当にやさしいのかという検証が必要である。その際には、対象者の母語についても併せて考慮する必要があるだろう。次に、日本語教材『にほんごこれだけ!』との接続についても課題である。語彙については、3つのレベル設定と共にその接続についての構想を示した(6節参照)。文法については、話し言葉に特化した『にほんごこれだけ!』の「STEP1・2」に対して、書き言葉に対応できるSTEP Rの提示などが考えられる(4-2節参照)。さらに、1節で述べた「HKプロジェクト」の構想に基づいて、今後は工学系など他分野との共同研究により具体的なプロジェクトとして多様な問題を解決していく必要がある。

参考文献

- 福生市 (2009) 『外国人のための生活便利帳 やさしい日本語版』『進む内なる国際化～「やさしい日本語」を活かした地域共同の実現～』グループ編
- 林四郎 (1971) 「語彙調査と基本語彙」『国立国語研究所報告 39 電子計算機による国語研究Ⅲ』 pp. 1-35, 秀英出版
- 庵功雄 (2009) 「地域日本語教育と日本語教育文法——「やさしい日本語」という観点から——」『人文・自然』3 pp. 126-141 一橋大学
- 岩田一成 (印刷中) 「言語サービスにおける英語志向——「生活のための日本語：全国調査」結果と広島的事例から——」『社会言語科学』13-1, pp. 81-94, 社会言語科学会
- 小林ミナ (2009) 「——過去から現在へ——文法研究と文法教育」小林ミナ・日比谷潤子編『日本語教育の過去・現在・未来 5 文法』 pp. 18-37 凡人社
- 国立国語研究所 (1984) 『日本語教育のための基本語彙調査』秀英出版
- 国立国語研究所 (2004) 『行政情報を分かりやすく伝える工夫に関する意識調査 (自治体調査)』国立国語研究所 研究プロジェクト「日本語の現在」意識調査グループ
- 国際交流基金・日本国際教育支援協会 (2002) 『日本語能力試験出題基準 (改訂版)』凡人社
- 松田真希子・児玉茂昭・竹元勇太・石坂達也・森篤嗣・川村 よし子・山本和英 (2010) 「コーパスの異なりと単語親密度を活用した日本語共通基礎語彙の抽出」『言語処理学会第 16 回年次大会予稿集』 pp. 579-582, 言語処理学会
- 松田陽子 (1999) 「外国人のための災害時の日本語」『月刊言語』28-8 pp. 42-51
- 松田陽子・前田理佳子・佐藤和之「災害時の外国人に対する情報提供のための日本語表現とその有効性に関する試論」『日本語科学』7 pp. 145-159
- 御園生保子・前田理佳子 (2007) 「やさしい日本語の仕組み」『「やさしい日本語」が外国人の命を救う』 pp. 31-37, 「やさしい日本語」研究会
- 水野義道 (2006) 「災害時のための外国人向け「やさしい日本語」」『月刊言語』35-7 pp. 54-59
- 森由紀 (2006) 「ユニバーサル日本語の観点をふまえた公用文の考察」『三重大学国際交流センター紀要』8 pp. 1-14
- 野元菊雄 (1990) 「簡約日本語」『文林』26, pp. 1-36, 松蔭女子学院大学
- 野元菊雄・川又瑠璃子・義本真帆 (1991) 「簡約日本語の創成」『日本語学』10-4, pp. 94-105

- 入管協会 (2009) 『在留外国人統計』 入管協会
- 埼玉県総合政策部国際課 (2006) 『外国人にやさしい日本語表現の手引』 埼玉県総合政策部
- 佐藤和之 (2004) 「災害時の言語表現を考える」『日本語学』 23-8, pp. 34-45
- 佐藤和之 (2007) 「被災地の72時間——外国人への災害情報を「やさしい日本語」で伝える理由——」『「やさしい日本語」が外国人の命を救う』 pp. 9-27, 「やさしい日本語」研究会
- 専門教育出版 (1998) 『品詞別・A～D レベル別 1万語彙分類集』 専門教育出版
- 田中ゆかり (2009) 「首都圏の多言語表示——“標準化”の観点から」『日本語学』 28-6 pp. 10-23
- 寺村秀夫 (1975) 「連体修飾のシンタクスと意味——その1——」『寺村秀夫論文集 I (1992)』 所収 pp. 157-207 くろしお出版
- 山田泉 (2002) 「第8章 地域社会と日本語教育」細川英雄編『ことばと文化を結ぶ日本語教育』 pp. 118-135 凡人社

付記

本稿は日本学術振興会科学研究費補助金による基盤研究 (A) 「やさしい日本語を用いたユニバーサ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, 研究代表者: 庵 功雄) の成果の一部である。

注

- (1) 例えば、『日本語教育』138号(2008.7)で「多文化共生社会と日本語教育」が特集されている。
- (2) 庵 (2009) の Step1 レベルに対応. 続けて『にほんごこれだけ! 2』を2011年に出版予定。
- (3) EJ は Easy Japanese の略である。
- (4) 減災 EJ においても文法項目の選抜などのマニュアルの改訂が考えられているようである (御園生・前田 2007: 36-37)。
- (5) ここで言う「公文書」には学校で担任から保護者へ配布される「お知らせ」なども含まれるが, 本稿では自治体が発行した文書のみを対象とした。
- (6) 国立国語研究所 (2004) による広報誌編集担当者への調査によると, 「広報誌の編集で高齢の読者に配慮していること」で「特に配慮はしていな

い」は3.9%だが、「広報誌の編集で外国人の読者に配慮していること」で「特に配慮はしていない」は78.9%である。

- (7) 例えば、地域日本語教育向け教材として開発中の『にはんごこれだけ!』の語彙は Step1, Step2 を合わせても、800~1,000 語程度になるとみている。
- (8) 埼玉県(2006)、松田(1999)他にも同様の指摘があるが、本稿では内の関係については修飾を認める。公文書において、すべての連体修飾を解体することは実質上不可能である。
- (9) 本節に関しては、我々執筆者が書き換えたデータを基本としているが、筒井千絵氏にいただいたコメントも参考にしている。
- (10) 書き換え作業は共著者3人がそれぞれ行ったが、ここでは森篤嗣が書き換えたデータのみを用いる。
- (11) ChaSen-2.4.2 及び、IPAdic-2.7.0 を用いて、形態素解析をした結果である。なお、全角数字、漢数字、丸数字は半角数字に置換した。また、未知語、記号、固有名詞は除いた。