

読解支援のための簡略化手法に対する一考察

— 「やさしい日本語」への変換による読解支援を目指して —

中 尾 桂 子

あらし

専門教育を受ける日本語学習者の読解支援の一つに文章の簡略化という方法がある。これは、自動、または、人手により、1文の構造を単純にし、表現を初級レベルの学習内容に合わせて書き換えるというもので、主に、教育的立場と工学的な立場から読解を補佐するために利用されている。立場の違いで実行方法に違いがあるが、いずれにせよ、言語学的な格枠組みで言う余剰成分を極力消去するという格構造の単純化がメインである。これは内容把握を第一に優先することによるが、読み手がいずれ独り立ちして読解を進めていくことも視野に入れた読解支援としては、教育的な指導といった観点を考慮し、実際の学習者の自律的学習を促進するための配慮を十分に行うことが望まれる。この場合、簡略化の指標を構文情報だけではなく、教育的示唆を含めて考えること、また、余剰部分を消してしまう簡略化ではなく、消すべきとされる余剰の語とそうではない必須の部分との価値の違いを把握できるような方法を検討すべきである。すなわち、表示の切り替えで注目すべき箇所を強調するということである。このような考えに基づき、文章を「簡略化」するために、どのような指標を用いるべきか、留学生のニーズも踏まえて簡略化に対する考え方や視点を変更すること、また、簡略化の際の指標を拡大するといった3点を提案する。

キーワード：読解支援、簡略化、word colligation、基本文型、構文解析

1. はじめに

1.1. 読解にまつわる2つの立場と社会状況の変化

従来より、日本の技術や文化に関する専門知識を求める人々の間では、専門知識を学ぶ目的で日本語の文章が読まれてきた。この場合は、通常、日本語を学んでから読むというのではなく、既に、母語である程度の基礎知識を十分に備えている研究者が、それを確認する目的で、辞書や文章構造の手がかりを元に日本語の文献を読む場合が多く、読解という観点から見れば、これは、母語での読解能力や情報収集スキルを駆使して、内容のみを読み取り、専門に生かすというものである。技術的には、主に、スキミングやスキニングのスキルを駆使するものである^[1]。

このような知識収集方法は、日本語教育という観点から見た読解とは全く異なるものであると言える。それは、日本語教育における読解の第一目的が、既習の日本語や文法知識を利用するフィールドを与えることにあり、第二の目的が、日本語の文章構造を知り、徐々に、知識を得るために文章を利用するための橋渡しの機会を与えることにあるからで、専門知識の為に日本語を読む立場とは優先順位が異なるものだからである^[2]。

専門知識のために外国語の文章を読むという立場では、ただテクニックとして情報を収集するスキルだけを身につけるという手法のみが求められる。すなわち、日本語の文章構造を知り、構造と表現といった言語上の手がかりから、必要な情報を収集する方法や、また、既習の専門知識の翻訳

専門日本語を手がかりにして内容を確認する知識のようなものである。

しかし、日本語学習者の学習ニーズや年齢層、立場、社会的身分が多様化するとともに社会情勢が異なってきた結果、共通語としての1つの言語が使えない場合や、専門教育という枠が拡大して大学への正規入学を目指す留学生や、義務教育機関での年少者も専門日本語教育の対象者として考慮に含まれるようになった。そのため、昨今では日本語教育の分野に「専門教育のための日本語教育」という分野が確立している。

1.2. 専門教育のための読解支援における課題

義務教育機関で学ぶ日本語学習者や大学以上の高等教育を受ける留学生など、日本語を学ぶという語学教育が目的ではなく、専門知識についての教育を受けることが主たる目的である場合、日本語学習者、並びに、その指導者には、できる限り早く、専門知識の学習に移行したいという希望がある。

しかしながら、昨今では、日本語と同時に、専門知識の概念も学び、さらに、論文指導や文献収集の指導とも関連する読解テクニックも併せて学んでいかなければならない場合が多々ある。このような学習者の場合、必ずしも、母語での読解力や書物からの情報収集スキルを応用できるわけではなく、文脈から事実関係を捉えることに困難を覚える場合が多い。そのため、社会的な立場を確立できずに自尊心が保てないという問題や、日本語学習の放棄という問題があり、その動機付けと自律学習継続習慣が成功を左右するとされる^[3-5]。

そこで、ごく初級レベルの日本語で専門教育を目指す学習者のための指導法を工夫する必要から、学びたいという意思を動機とし、次の学習につなげる原動力とするよう、主に、専門書の読解力向上を目指す自律的読解教育が検討されるようになってきた。特に、小中高の外国人児童・生徒や大学院の留学生には、初級の段階から専門教育につなげるための教育が、読解を中心に考慮されはじめ、多くの教材が作成されている^[6-9]。

ただし、指導者が常に存在するわけではなく、読解すべき課題の文章も個人により異なることが多いため、授業や課外における教員の指導、その他、逐次読みや内容解釈の支援として、チューター制度や読解支援システムが利用され^[3-5]、また、昨今では、自律的な支援の一環として、e-learningに関連する支援が望ましいという考えがあり、翻訳システムの利用や、辞書機能の活用、そのほかに、簡易日本語に書き換えた文章を用いて読解指導することも多く^[9]、教育的指導を備えた読解支援システムの自動化を検討する向きも増えている^[10]。

2. 従来の読解支援のための文章簡略化

読解支援の多くは、文章の構成要素である単語（形態素）レベルに対訳をつけるというものが最も多く、語彙力の補助や辞書で意味を確認する作業の軽減が利用の主な目的と想定されている。しかし、多様化する学習者のニーズとして、文章の内容を理解するのに、個々の単語の意味が分かるだけでは十分とは言えない場合が多くなってきている。

そこで、初級レベルの文法や構文知識でも、ある程度の因果関係などを論理的に把握することを目的とし、初級学習者の既習の文法構文知識の範囲で専門書などを書き直す、簡略化が行われ、簡易版テキストの使用や簡略文章を用いた指導が行われるようになってきている^[8-9]。

このような文章の簡略化による読解支援のあり方は、機械処理か否かで2つに分けられる。1つは自然言語処理技術を応用して自動化されたもので、自動翻訳、要約等を目的として大意の把握の

ために自動的に簡略化するものである。簡単に言えば、設定されたレベルに位置づけられた語彙や構文のパターンを、シソーラスや品詞情報、語彙の上位、下位概念の区別といった情報を指標に用いて、別のレベルとして設定されたパターンに差し替えるというものである^[11-13]。機械的な変換による場合、文章内容を極力変えずに簡易化するために名詞や動詞といった主要な実質語を平易な表現に差し替えたり、構文を別の構文に変換するのであるが、その際、不要な情報となる必須格成分以外の余剰成分、すなわち、修飾語を割愛する方法をとることが多いのが特徴となる。

もう1つは、読解支援者が、手作業により、読み手の読解状況を踏まえて調整していくものである。こちら読解者の内容把握が主な目的とされるが、全ての文に対して一様にパターンを差し替えるのではなく、余剰成分である修飾語を省略・割愛し、出来事の流れの把握に必須の情報のみを文中、あるいは、段落に残すという方法をとることが多い。しかし、構文自体が複雑な場合は、自動処理同様、差し替えも行われる。手作業による場合、文章構成上のまとまりごとに、複数の文をまたいでコンテキストベースの判断が行われ、文の価値の違いに着目した省略・割愛が期待できるため、さらに、読解者の能力や教育的配慮を加味して、省略や割愛を変えていくことも多くなり、支援者によって、簡略化の結果が異なることも多くなる。手作業による簡略化のための書き換えの場合、特徴として指摘できるような特に決まった方法があるわけではない。しかし、そのことが特徴とも言える。それは、書き換えに携わった個々人の文章表現力や主観に影響されることも多く、一貫性や均一性に疑問が残る場合もあることによる。また、書き換えの労力が大きいことから、分野に関わらず、希望する資料や文章の簡易版を学習者の誰もが必要とときにすぐ手に入れられる環境にあるわけではないという点も考慮すべき点である。

人手による場合は、読解支援体制下にある文章を量産できず、簡略化の箇所や簡略化の方式が個人により異なるが、内容に応じて、正しく簡略化が行える。また、機械的に行われる自動処理では、簡略結果にぶれが生じることなく、画一的に大量の多様な文章に対しての簡略化が見込まれ、読みたいものが即読めるという利点があるが、文構造において必須となる語を指標に、それ以外の余剰成分となる語を消去する方法、すなわち、1文の構造を、述語ベースの必須格に基づいて余剰成分を落とすという方法で単純な文に作り変えるため、日本語教育で利用される、いわゆる、表現文型といった、複数の語や助辞が組み合わさって形成される、一種の慣用句的な文法上の単位 (colligation) や連語 (collocation) は、考慮の対象として十分カバーされにくい。

自動処理と手作業での文章簡略化は、基本的には余剰成分を消去するという点で同じ考え方のものである。ただ、技術的な可能性により、応用範囲が異なり、同様に比較できるものではなく、どちらにも良い点と不十分な点がある。どちらにしても、読解する日本語学習者から見れば、1文中の省略・割愛という点で同じだが、差し替えの際にゆれが生じる手動よりも、一様にパターン変換が行われる画一的な変換の方がわかりやすい場合もあり、情報収集の効率化という点で個人的な好みも反映される。しかし、ごく初歩の学習者においては、ただ差し替えや省略を行うだけでは情報の選別に対する支援は得られず、コンテキストの把握には不十分である。

余剰成分消去法では、文構造のどの部分を主要情報とするかという読解の手がかりが明示されず、文章のどこに着目し、何を手がかりに内容を的確に取るかという技能的なスキルを養うことができない。簡略化文章で与えられる情報からは自律的な読解方法を養うことには至らない。根本的に、読解能力向上は期待できない点が惜しまれる。

3. 研究目的

現在の学習者の状況やニーズから考えると、読解支援を必要とする読解者にとっては、日本語教育上の読解指導という観点、専門知識を収集するための技術的な観点といった、読解そのものにおける個人それぞれの欠点を補う形での読解支援が望ましい。

そう考えれば、現在の読解支援としての簡略化という手法は、機械、人手のいずれにしても支援としてはニーズにあっていないのではないだろうか。複雑な構造の文を、単純な構造の文で記述された文章へと書き換える簡略化という支援では、教育活動として授業内で行われる読解支援の観点は考慮されないし、また、日本語教育における知識の単位を踏襲した簡略化ではなく、機械的処理の正確さを第一に優先するための自然言語処理上の簡便さを優先した簡略化単位では、文章読解時の捉え方が人間の読解支援にそぐわない。

そこで、日本語教育における表現の単位と、読解支援という教育的示唆を加えて「簡略化」するために、文法上の単位と格構造上の単位を融合し、かつ、教育的に読解支援を行う指標を再検討する必要があると考える。

ここで言う日本語教育における単位とは、一種の慣用句的な文法上の単位 (colligation) や連語 (collocation) を指し、前者は、いわゆる教育用の「文型」であり構文として扱われるものも含む。後者は、いわゆる慣用表現や寺村 (1982, 1984) の言う「表現文型」を指す^[14]。構文としての「文型」の中には、連語レベルではあるものの結びつきのたいへん強い「表現文型」が含まれる場合も多く、実際には、その差は連続的なものであるため、分けられないが、便宜上、「文型」と「表現」とする。

本稿では、まず、簡略化の指標である構文情報を見直し、さらに、日本語教育で言う「基本文型」や「表現文型」を再検討する。そして、量的処理が可能な機械処理上の簡略化の方法として、「消さない」簡略化という発想の転換を提案する事が可能か、すなわち、表示の切り替えて、注目すべき箇所を強調するという方法について考察する。これらに基づき、表示切替による教育的読解支援方法として提案したい。

4. 簡略化指標としての構文情報

4.1. 言語学的に見た「文型」と「構文」

自然言語処理の技術では、構文解析を行うことが自然言語としての日本語の工学的な利用の前段階として必ず必要であること、ならびに、機械処理上、コンピュータ言語としての観点からもその法則性が重要となることから、一文中の動詞に基づいた格枠組みによる構文情報が重視される。日本語教育では、構文ということばを言語学的に文の構造を広義に捉えた場合に利用される術語であり、実際の現場では文型ということばが利用されている。使い方によっては類義語とも考えられるが、どのように考えればよいか。

そもそも「文型」とはどのようなものか。そして、「構文」とはどのように違うのか。「文型」について言語学大辞典には「広義には文をいろいろの観点から分類した結果得られる文の種別、類型をいい、狭義には、文の構成要素の組み合わせの型をいう」とあり、意味的側面と形式面側面の2つの観点から分類されるとある^[15]。

意味的側面から分類された文型には3ジャンルあると言うが、それは、まず、平叙文、疑問文、

命令文、感嘆文の4種を言うもので、機能、性質による伝統的な分類である。次は、存在文、所有分、条件文、推定文、現象文など、形式面での対応があるもので、表現内容や機能に基づくため、「～表現」と称されることもある。最後が、結果文、目的文、時間分、場所文、譲歩文、比較文、関係文、前提文、条件分、帰結文、引用文といった、複数の節の関係を表す分類である。これらが意味的側面での分類である。

形式的側面から分類する場合は音声面や構造面の類型で分類されるが、構造面からの分類では、様々なレベルでの様々な大小の構造体の類型として、その構成要素である語の配列の文中での機能により抽象化されて分類される。言語学大辞典によると、形式面で分類された文型には4ジャンルあると言う。まず、西欧の伝統的文法観に倣う主述関係から見た文型、次は、述語の有無と種類から見た文型である。述語の有無と種類で見る文型は、国語教育で伝統的に名詞述語や形容詞述語文と分けることの多かった、助動詞「だ」「です」「である」をつけて述語とする、いわゆるコピュラ文と動詞のある動詞述語文とを比較するものである。そして、単文、複文、重文といった節と節との関係から見た場合があり、最後が、英語学習で必須の知識とされるSV、SVC、SVO、SVOCの基本5文型を示す、構成要素の配列から見た場合である。英語学では、この短文の構成要素の組み合わせタイプを指して文型と称してきた。

以上を概観すると、「文型」とは、文としてのまとまりの中の語配列の組み合わせ類型という捉え方でよいだろう。

一方、「構文」という言い方もある。たいていの場合、「構文」は文の「構造」、「組み立て」という意味で利用されが、“construction”や“structure”が多分野で多義的に利用されることからわかるように、こちらは一義的には捉えにくい。

言語学大辞典によると、構文は、一般に、言語の形式的、機能的単位の構成を指す。また、言語の構造と言え、言語を構成している要素がその言語の中でどのように配置されているかを言う。広義には体系と同義語で、狭義には談話中の統語関係における配列の仕方となる。

現在、自然言語処理では、例えば、「この形態素解析では構文情報利用している」などと使用されたりするが、これは言語学的に言う、「語が語幹と接辞からなる場合、語幹と接辞の間には一定の配列の仕方が決まっています、それが語の構造をなす」という意味である。構文解析とは、具体的には、配列がどのような規則で生成された規則にあてはまるか逆算することで、つまり、決まっている規則の複数のあり方のうち、いずれに、該当の自然言語が合致するかを判断することである。「構文」という術語を利用する立場では、文や句が語から構成され、その構成が形態指標によるか、構成要素の位置によるかで、識別、判断されるものだという考え方に立っている。

以上、類義語として漠然と同義的に使用される「文型」と「構文」について区別して整理したところ、厳密に言えば、文型と構文は何をどう見るかという点で異なっており、「構文」には文全体をパターンとして類型的に捉える観点が低いと考えられる。しかし、「文型」は文の構成要素の組み合わせの型であり、これは、実質的な意味を持つ単語と文法的な機能上の意味のみを持つ単語の組み合わせで構築される文、またはその構成全体を指し、言語の形式的、機能的単位の構成全体、すなわち、コーパス言語学で言うところの colligation というものである。

「文型」という術語は、組み合わせの型が存在するとして、その組み合わせ（パターン）に名をつけるような立場で使われる。一方の「構文」は、構成のあり方や構成要素の関係がどのようになっているかに着目する場合を指す。いずれを利用するかで立場の違いと、着目する視点、観点が、全体か個別かの違いにつながっている。同義のようで別のものである。

本稿では、「文型」を広義の意味で捉え、表現時の構成配列の類型としてセットで捉える場合に使

用し、「構文」を個別の統語的配列として捉える場合に使用する。「構文」では、特に、動詞述語文で必須の格関係を指すものとする。

4.2. 日本語教育で言う「文型」

現在の日本語教育では、学習者の誤用に対する経験的な知識に基づいて誤用を回避しながら、また、基本構造を知り、語彙を差し替えるだけで利用できるような応用力を身につけるための基礎として、単純な文法から積み上げるという方式で、場面に応じた表現形式を学ぶ。

指導における表現形式は一般に「文法」、「文型」と呼ばれているが、これらは一義ではない。現在の日本語教育教材に見られる「文型」を言語学大辞典の基準に照らせば、前節で述べたように、広義の意味と狭義の意味とが混在したものであり、出自や性質での分類から、コロケーションに準拠したものまで意味においてもレベルの差があり、また、意味面での分類だけではなく、形式面の分類まで幅広く扱われていると考えられる。

日本語教育は、植民地時代の、実践と暫定的な研究に端を発し、アメリカ言語学の流れや第二外国語教育の観点など、時代に応じて、複数の視点が混在して形成されてきた。昨今では学習者ニーズが多様化していることや、教師の背景の異なりから内容的や教授法が拡張していることもあり、その術語の捉え方も多様化しているようである。したがって、そこで取り扱われる「文型」や「文法」は抽象的な概念となっている。効率と目的を優先するという機能面優先の立場があり、その一方で、基礎と応用力の関係を優先する自律的、言語的構造把握を優先する立場とが常にせめぎあう中で、学習者ニーズに合わせて指導項目や観点も変容してきたが、今後も変容していくだろう。

現在は形式的な意味で、文法項目を組み合わせて形成される「文型」の見直しの必要性が指摘されており、最近の学習者の多様化に伴い、目的に応じた学習を進めることを第一に優先する立場がとられることが多くなった。結果、年少者教育や大学留學生意外の目的別学習者には、特に、文型規定の基準を、機能や場面中心で選別するような考え方が出てきており、また、それが文型を見直そうという動きにつながっている（野田，2005）。しかし、能力試験を大学入学や就業条件とされることに変化は見られないため、構造的な正確さという観点に根ざした文型教育が消えるほどの変化ではない。したがって、教育現場では、教師も学習者にとっても概念でしかない仮想の「文型」が、実態を持っていることには変わらない。

「文型」が実際どのような虚構であるかについて例示する。最も、使用率の高い教材として、翻訳文法書や副教材が充実する、スリーエーネットワークの『日本語の基礎』、『みんなの日本語』という初級用教材があるが^[16-17]、これは技術研修者を学習者として想定した教材と、それをベースにして、後に、技術研修生以外にも対応され得る語彙や文法項目に差し替えた教材とであり、それぞれが前者後者である。これらの教材では、学習者に、「文法事項や単語が文型や会話の中に入っている」ので、それらを繰り返し練習し、日本語の表現を一定の文の形として理解し、覚えることで、日常よく使われる日本語表現を自らも会話で使えるようになる」と教科書の仕組みと合わせて学習内容を示している。

この教材で扱われる項目は、語彙、文法事項、文法事項の利用法（組み合わせの型）である。例えば、第1課「自己紹介」では、「私はマイク・ミラーです」「サントスさんは学生じゃありません」「ミラーさんは会社員ですか」「サントスさんも会社員です」という4つのパターンが紹介されているが、テキストには、語彙として、「私、あなた、～人、～さん、先生、研修生、留学生、誰」等を紹介している。クラスの学習者に合わせて適宜語彙を追加して紹介しつつ、この課で学ぶべき文法事項で会話できることを目指す教材である。第1課の文法項目として「～は」「～も」「です」「じゃ

(では) ありません」「か」があげられている。そして、語彙や文法事項といった文の構成要素の組み合わせが抽象化されたものが、図1にあげたような形で紹介されている。図1は、語彙と文法事項を用いて利用できるようになるべき表現として、テキストの「練習A」にあげられたものであるが、これは、実は練習と言うより、この課で学んでもらいたい抽象化された文の類型的組み合わせパターンリストとなっている。

このパターンを見ながら、四角の枠内に、紹介された語彙を挿入することで、必須学習文法事項である「～は」「～も」「です」「じゃ (では) ありません」「か」の使い方がわかり、自己紹介、その後の会話、名前や出身国の確認を行う方法を習得できるよう練習が繰り返される。これは、日本語教科書の典型例であり、この1課のような学習の流れは他の教材でもおおむね同じである。

図1 『日本語の基礎』第1課練習A

つまり、教師や学習者は、上記のように文法事項を組み合わせたセットを文の形として提示し、それに語彙を当てはめるという方法で、読む、聞く、話す、書くといった運用と知識の構築学習を行っているのであるが、この、文を抽象化した文法項目のセットが、いわゆる、教育用「文型」である。

そして、教材を利用した学習の経験を持つ日本語学習者は、基本的に、文型、すなわち、文法項目の組み合わせの型に応じて新出語を覚え、挿入していくという学習スタイルに馴染んでいる。それは、この基本文型との兼ね合いで新出の文法項目を積み上げ、日本語の文法力、会話力を養っているためであるが、そのために、学習者には、読む(読解)、聞く(聴解)、書く(作文)、話す(会話)の4つの技能の運用のための取っ掛かりが、この基本文型に準拠したものとなるのである。

さらに、日本語教育の教科書の多くは、学習を始めた段階の初級レベルから、日本語能力検定試験1級合格といった上級レベルの段階まで、教科書内の抽象化された構成要素の組み合わせパターンで学習していくが、この流れの中では、主に初級レベルで指導される文法項目の組み合わせパターンが、教育用「基本文型」として認識されており、中上級の学習者の言う文型とはいわゆる表現的な項目の組み合わせである。慣用表現ほど語彙的なものではないが、結語よりは連語程度にまとまりとしての意味が問題となるものである^[18]。

5. 読解支援を中心に考えた簡略化のあり方

日本語学習者の視点から読解支援を見れば、自然言語処理技術を応用した方法と手作業による簡略化は、文の持つ情報自体の格差に対して配慮があるかどうかという点で異なる。文が持つ情報とは、内容上重要な部分を表す文や、文章の流れや展開がつかめるか否かという点から見た重要性に関するもので、それは機械的に判断が難しいものである。しかしながら、読解初心者には、支援として必要になるものである。

読解スキルが十分に身につけていない学習者は、一般に、初級～中級者といわれるレベルの学習者に多いが、一概に初級といっても、その上下の範囲には開きがあり、文章の内容が得意な分野か否かの好みによっても読解とその理解の程度が異なるため、学習者の能力に合わせて、ある程度の「教育的配慮」による調整が可能なことも重要になる。読解力が十分ではない日本語弱者にとっては、一様に、ゆれなく、語彙や構文の差し替えを正確に行ってくれることよりも、実は、情報としての価値の違いと文章構成の関係が把握できる支援の方が、自律的な読解につながりやすく、教育的効果が期待できる^[1]。

しかし、ある程度、読解力があり、ただ、語彙力だけが十分でない日本語弱者にとっては、一様にゆれなく語彙をパタンで差し替えてくれる安定した簡略化で十分である。教育的配慮を加味した読解支援としては、手作業で行われるような、コンテキストベースで、学習者のレベルに応じて簡略化されるものとなり、そこに、ある程度の自律的な読解を継続させるために、安定した変換が望めるためである。そうすると、初歩的な日本語学習者から中上級にかけての学習者に対する読解支援を念頭に置くと、画一的で一様なパタン変換が組み込まれたものが同時に必要になると考えられる。

ところが、手作業での教育的配慮を加味する場合、自動化されたパタン変換とは異なり、書き換え作業の労力が大きく、学習者全員が、希望する資料の簡易書き直し版を必要となきときやすぐ手に入れられる環境にあるわけではないという点が問題になる。また、書き換えには、特に決まった方法があるわけではなく、予想される「簡易さ」に基づいて行われるため、個々の読解支援のための書き換え担当者の表現能力や主観に影響されることも多い。

このような問題に対して、支援としての読解用文章簡略化を考え、かつ、教育的示唆にも配慮に置くと、作業量、主観的判断のぶれには自動化手法、そして、的確な要約には、キーワードの選定への配慮と文脈情報の利用ということが必要である。

作業量の問題は、自動化を検討することで解決が期待できるが、主観的判断の問題については、ごく簡易な状態は、文脈の流れを決定付ける文、すなわち、ナレーションの部分のみを必須成分のみの文にすれば、ある程度の統一感が期待できる。それは、手作業で行う簡略化では対応させられる対象数が少ないし、同じ文章を様々な視点で、多様に変化させるには、オリジナルの情報は保持したままで可逆的に変換する仕組みが必要であり、そのように何度も繰り返す簡略化変換に対応するには、紙面上ではなく、電子的な文章を用いた方が効率がよく、制約が課しやすいことによる。

ただし、そもそも、簡略化による省略という読解支援には、根本的な問題がある。簡略化した結果である文章のみを表示することは、「自律的」な読解力を養う支援としては不十分なことによる。自律的な読解力とは、母語（または第一言語）における基礎教育の中で養われるが、その方法は、通常、「流れを掴む」ことや、「何が」「どうした」という情報を掴むことを目的として、文章構成や文の構造を見極めるような指導をうけることで各自が自分なりの読解戦略を確立し、読み取った情報を統合して理解を進めていくようになるというものに裏打ちされている^[1,4-5]。そのため、教育的配慮で「自律」を促す読解支援を考えるならば、簡略化結果のみを示すよりも、文章のどこに着目するかということ把握できるように、オリジナルの文を残して、着目点のみを示す方法で簡略化を考えるべきである。

必要な文だけを掴めれば、文中での話題や内容が表している場面状況が掴め、それらを統合していくことで、文章に表される場面展開や文脈という出来事の流れが把握できる。このことから、自動化の場合に、概要が理解できるようにどの程度流れを汲んだ割愛が行われたか、その軌跡もあわせて見られるように配慮することが望ましい。

6. 表示切替の指標としての文型

5章までの考察からすると、基本的な文型を簡略化の際の指標として利用する方法を考えるには3つ問題がある。

まず、文型そのものが不明瞭な存在であるということである。教科書で扱われる文法項目はコマ切れに提示されるものであるが、指導時には組み合わせて文の骨格を作り、そこに語を当てはめるため、あたかも、実際に、日本語がその文法項目を組み合わせて作られた骨格を持っているように見せる。しかし、その骨格をまとまりとして見る知識は日本語母語話者にはないもので、学習者特有の知識である。ネイティブにも不明瞭な組み合わせの規則を指標とするには、何か辞書や教材に準拠することになる。

また、同列には扱えない情報を複数利用する必要があるという問題がある。この文法項目を組み合わせて形成される学習用「文型」は、辞書的な語だけではなく、品詞情報、構文情報、意味論的、語用論的情報といった文法的な情報を結合した colligation 情報である。品詞情報と文字列の組み合わせに意味と語用論的情報を加味して調整したものであるから、性質が異なるものが複数寄せ集められていることになる。データベースとしてデータを扱うという観点から考えれば、異なった性質のデータを一位に扱うのは問題があるため、データは可能な限りは1次元のものとして扱うことを考える必要がある。

そして、これが最も大きな問題であるが、文型としての重要性の判断基準には恣意性があるということが挙げられる。「文型」として採用されるか否かは教科書の編纂者の意図によりかなりの差がある。積み上げ式に将来を見越して文法項目を積み上げるならば、初級のごく最初に学ぶ語と組み合わせられて使われる格助詞の意味に差が出る内容は計画的に示す必要があるだろう。ところが、あって然るべきはずのものが、上級でも、表現としても、取り扱われることがない場合があり、その数は決して少なくはないのである。

例えば、森田(1989)は、語の意味分析の観点から、文型と意味との関係に着目して、文型判断に影響する語の意味の差を指摘しているが、「出す」には、「AがBにCヲ出す」形式を取るものと、「AがCヲ出す」形式のように「相手B」を必要としないものがある。それは「出す」が「提出」の意か「外に取り出す」意かにより決まるという^[19]。また、「(相手B)に」の格助詞の意味が、相手か場所か何かの目的や存在としてかといった3つの意味に応じて、語順入れ替えの可能性に影響があるというのである。似たようなものはまだまだ多いだろう。実際の教育現場では、生教材と言って、新聞や小説を用いた読解指導を行う場合に、たまたま出てきた表現を数多く取り扱うことで、教材等の穴を埋めていると考える。

以上のような3つの問題に対処するため、3つのうち、まず、最初の、何を文型とするかという問題には、「基本文型」を国際交流基金(1994)に準拠し、「文型」をジャマシイ(1998)に準拠することで対応したい^[20-21]。また、2つめの輻輳的な情報の統合、すなわち、同列化、単純化という問題には、検索条件としての文型情報を階層的に整理し、同時に、複数の性質の異なる部分に検索をかける方法をとることとする。最後の3つめの問題はここでは対処済みとして話を進める。それは、本来なら問題ではあるものの、交流基金(1998)相当でなければ、日本語能力試験に出ないためである。不十分な対応ではあるが、今回は基本文型の基準をこれに準拠するためである。

以上のように、文型情報をどのように考えるか明確にし、簡略化の指標を決定する。これを指標に、日本語の文章の簡略化を行うわけであるが、これらを残して他を消すことで文章の簡略化を実

行するのではなく、自立的読解力を要請するためには長い目で見てもよくないということで、これら指標を強調表示し、他を残したまま、焦点化する方法をとる。そのため、日本語のテキストを構造化し、指標検索のために必要な文法的情報をテキストの下位情報として付属的に階層化してデータを整理する。このようにデータ化した文章を利用する事で、自動的に WEB ブラウザ上で「簡略化」が行えるようになる。具体的には以下のような手順で行う。

(1) 簡略化指標となる「基本文型」特定のための下処理

- ① 『日本語文型辞典』の解説部分の情報を参照して、文字列と品詞情報により「文型」リストを正規表現で記述する
- ② 文字列から見れば同じ配列を含む「文型」は、1項目として集約し、統一する
- ③ パタン検索のため最長一致に配慮し、文型を長いものから順にリスト化する
- ④ 『日本語文型辞典』の見出し語を形態素解析し、形態素解析システムに合わせた組み合わせで検索できるように調整する

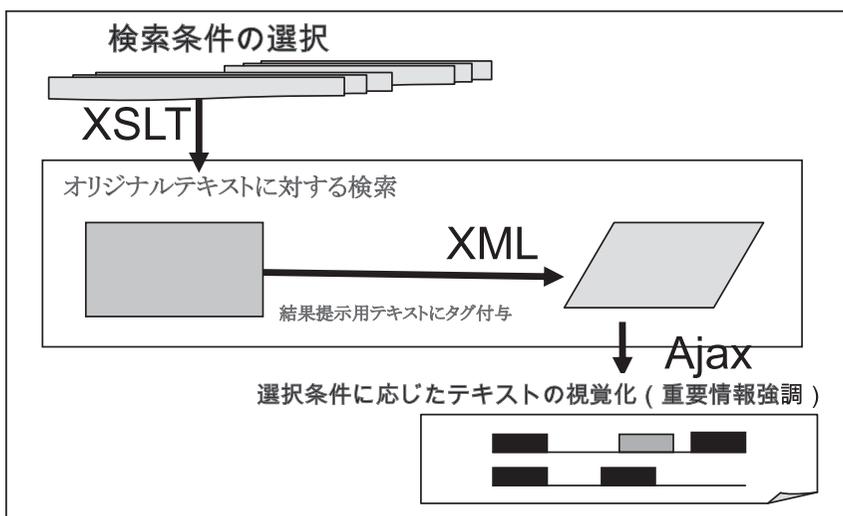


図2 補填情報の付与の流れ (xslt 検索⇒xml 変換⇒ブラウザ表示)

(2) テキスト構造の変更

まず、読みたい文章をシステムに投入すると、次に、内部で下処理を行って xml データに変換し、その後、最後に、投入時に設定した表示方法に基づいて表示返す、という3段階の処理を行う。処理の3ステップは図2に示す。表示切替実行方法は以下①～④のようになる。

- ① アップロードされたテキストに対して掛かり受け解析を行う。分かち書き習慣のない日本語の単語認定のため、形態素解析と構文解析を経て語同士の掛かり受けの解析に至るのであるが、この処理には次の7章で述べる KNP^[22]を用いる。この場合の形態素解析は JUMAN^[23]を利用する。
- ② ジャマシイ (1998) と国際交流基金 (1994) の文型をリストにし、それらにも、形態素解析を行い、「文系リスト」に合成する。
- ③ テキストに対して文型リストを元に検索し、ヒットした部分の情報を保有しておく。

- ④ XML^[24]を用いて、解析結果と文型リストとの照合結果を各単語単位の属性情報として、テキストの下位に格納しながら、テキストを任意に構造化する。

以上の①から④の手順を踏むことで、テキストは、結果として、複数の情報からなるテキスト型簡易データベースに変換されたことになる。再構造化された XML テキストを準備することで、ブラウザを介してテキストの表示を切り替えることができ、これにより、示唆的に読解上の重要箇所を示すことになる。

(3) 強調表示で示される内容

- ① 『日本語文型辞典』の見出し語は「文型」情報として下線をつける
- ② 「文型」の前の実質語など、意味上重要な部分を太字にする
- ③ 学習漢字レベルでの表示切替を選択制にする
- ④ 『日本語文型辞典』の解説部分を別ウインドウかフレーム等で別に表示させて意味を参照させられるようにする
- ⑤ 連体修飾節はマーカーとなる指標が文型として存在しないため、節のマークとして“)”を挿入する（例：「～節）名詞」）

これが普通だと思っている) 日本人の習慣からすると、その時一回限りのお礼だけでは物足りなさを感じる) のも無理はない。

図3 「文型」による強調表示例

これが普通だと思っている日本人の習慣からすると、その時一回限りのお礼だけでは物足りなさを感じるのも無理はない。

図5 KNP による係り受け情報を利用した強調表示例

「文型」による表示例を図3に示す。図3はスリーエーネットの『日本語中級 J301』という教材の7課の本文400字程度の文章に、表示切替処理をかけた結果から、一部を抜粋したものである。

7. 構文情報としての係り受け情報

「構文」判断の情報は、基本的に生成文法に根ざした語配列の格関係に基づく構造化のための規則である。述語の必須項をベースに、配列に、機能に応じた構造を与えるものである。これは、基本的に、文字列と品詞情報の2種類の性質の異なる情報を基準にしており、意味論的、語用論的情報は、将来的課題である。結局、解析結果の候補のうちのどれに該当するのかを選択する決定的な術はないが、統計的に算出された結語傾向で文脈の他の多頻度語を元に統計的に算出するなどの方法を併用することになる。格関係と意味上の連語関係を少し併用していることになる。

昨今、一般に、電子的に自然言語を扱うならば、ぶれなく機械的に処理でき、文の配列に構造を与えられる、構文情報に基づいた構造化が可能な解析結果を使うべきだという指摘もある。ただ、実際には、フリーで簡便に使える構文解析ソフトはあまりなく、係り受け解析システムがその代用となる。今回は、KNPを利用して格関係の情報に基づき、必須格を中心にみる強調表示を実行した。

図4は、図3で用いた例文と同じ箇所解析結果である。この結果に基づき、強調表示をかけたものが、前章図5である。図5は、図4のような解析結果から得られた主要情報に基づいて強調表示をかけた例である。KNPの解析結果では、節ごとに分かれて出力されているが、文章で強調する際、その分割が表示されにくいという問題があり、図5のようになってしまう。これは分かち書き等で表示の方法を工夫する必要があるが、今後の課題である。

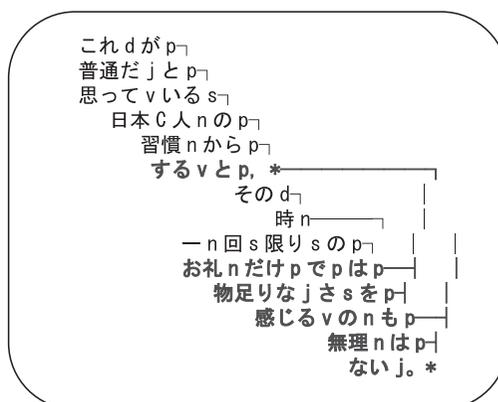


図4 KNPによる係り受け情報出力結果

8. 成人留学生に対する印象調査

1文の構造上の必須要素である構文情報のみによる表示強調と、文型情報も加味した表示強調と、どちらが、初級日本語学習者にとって、元の文章よりも読みやすくなっているか、また、内容の理解に齟齬が生じていないかについて、読解テストとアンケートを行った。

テストは、トピック、事実関係、因果関係が理解できているかについての内容把握問題とし、テスト時の文章は、中級日本語の総合教材『日本語中級 J301』の4、7、9課の本文や、日本の小学校の教科書などを用いた3問の読解問題である^[25]。

今回の検証は、中、上級段階の漢字圏の留学生10人にアンケート協力を依頼した。それは、自身の読解の問題や初級時からの課題を客観的に振り返ることができ、かつ、余裕をもってテストとアンケートに臨めるとともに、自身の意見の十分な説明をしてもらうことを期待するためである。

アンケートの手順は、まず、読解テストとして文章を読み、内容が把握できているかどうかを確認する質問に答えてもらい、その後、文章の視覚的な効果に対する印象や、自身の読解の問題について回答してもらうというものである。

文章読解テストでは、8人が問正解で、不正解の回答を含む留学生は、全て、全文表示で一部強調表示の文章の方が情報収集に便利であるとしていた。

不正解だった学生は、おそらく、漢字圏であるため、語彙情報にのみ頼り、構文情報を考慮せずに語彙力だけで読んでいたことにより、内容の因果関係が十分読み取れていないことが原因ではないかと考えられる。

テスト後、主文の述部と連体修飾節内の動詞述語とそれらの述語が必須とする語を重要度に応じて強調表示した文章を見せ、先の2つの文章と比較して検討してもらったところ、10人中8人が、消去法ではなく、表示切替による文の簡略化が望ましいと答えている。その理由としては、文法的な手がかりが示されていることと、それに応じてどこを見て読めばよいか分かるので助かるという印象を答えている学習者がいた。

簡単な印象アンケートではあるが、これらの結果から、トピック、事実関係、因果関係について、書き換え前と書き換え後では、読解の簡便さが違うという印象を持つ学習者が多くいたことが明らかになり、簡略化における表示切替法とその指標の拡大の良さが確認できたと考えられる。

表示部分への希望や希望する読解支援については、漢字のよみ、ことばの意味の他に、ある程度のイメージがつかめないものもあるため、もし、ウェブ上での表示切替を考えるのであれば、通常の読解支援にあるような単語の意味を添付する辞書機能の他に、単語の使用例、文法的な意味の説明、画像がほしいという希望があった。また、小さい文字や行間の狭い文章を1行表示するなど読みやすくする必要があると希望的に述べられていた。

今回協力を依頼した留学生は、概ね、文章の簡略化には好意的であったが、インターフェースが問題だと考える人が多く、実用化にむけての積極的な希望が多く見られた。また、アンケート対象の人数が少なく、また、漢字圏の留学生だけに頼るという偏りがあることは否めない。システムを実装し、実際の使用に耐えうるものとして公開するためには、アンケート調査を継続して、理解程度の変化を分析、検証することで、システムの有効性をさらに考察すべきであるだろう。

9. まとめ

複雑な構造の文を、単純な構造の文で記述された文章へと書き換える簡略化という支援では、通常、教育活動として授業内で行われる読解支援の観点は考慮されない。そこで、教育的示唆を加えて「簡略化」するために、文法上の単位と格構造上の単位を融合し、かつ、教育的に読解支援を行う指標を再検討した。まず、簡略化の指標である構文情報を見直し、さらに、日本語教育で言う「基本文型」や「表現文型」を再検討した結果、量的処理が可能な機械処理上の簡略化の方法として、表示の切り替えて、注目すべき箇所を強調するという方法について、留学生へのアンケートを参考に考察した。

結果、簡略化に対する考え方や視点を、消去法ではなく、軌跡を残すために表示切替で強調する方向に変更すること、また、簡略化の際の指標を構文情報だけではなく、いわゆる、基本文型に拡大するほうが、より決めの細かい支援が可能であるということが指摘できた。

よって、日本語教育で言う colligation 情報を利用し、消さずに表示切替で主要箇所を強調する「簡略化」いうの良さを主張し、今後の機械的な読解支援のあり方に対する1つの方法を提案する。

ただし、表示切替の方法についてはまだ実際には何も検討していない。文字の大小等の表示方式の確立やインターフェースを含む実際の使用上の諸問題は多い。システムを公開した上でのアンケート調査を継続し、理解程度の変化を分析、検証することが必要である。今後の課題である。

本研究は、中尾桂子・森下淳也^[25-28]の一連の読解支援システム研究における中尾の研究担当の部分から現在検討中の課題を取り出して再考し、加筆修正したものである。

- [1] 甲田直美, (2009) “文章を理解するとは——認知の仕組みから読解教育への応用まで” スリーエーネットワーク.
- [2] 日本語教育学会『日本語教育事典』(1982) 大修館書店
- [3] 村岡貴子・因京子・仁科喜久子・深尾百合子・加納千恵子 (2001) 「専門日本語教育の現状と将来の方向」専門日本語教育研究3, 15-19.
- [4] 近松暢子 (2003) 「外国語としての日本語の読み・読解研究」『第二言語習得研究への招待』くろしお出版: 67-85.
- [5] 岡崎敏雄 (2004) 「外国人年少者日本語読解研究方法論: 原理論」『文藝言語研究 言語篇』45号 筑波大学大学院人文社会科学研究科文芸・言語専攻 29-46.
- [6] 加納千恵子 (1990) 「専門書を読むための読解指導について」『筑波大学留学生教育センター 日本語教

育論集』第6号：35-64.

- [7] <http://language.tiu.ac.jp/>
- [8] 仁科喜久子・土井みつる・高野知子(2007)『初級文型で学ぶ科学技術の日本語 -An Introduction to Technical Japanese-』スリーエーネットワーク.
- [9] 光元聰江・岡本淑明(2006)『国語教科書対応——外国人児童・生徒を教えるためのリライト教材』ふくろう出版.
- [10] http://www.mext.go.jp/a_menu/shotou/clarinet/003/001/004/002.pdf
- [11] 乾健太郎, 藤田篤(2004). “言い換え技術に関する研究動向” 自然言語処理, Vol. 11, No. 5, pp. 151-198.
- [12] 呉浩東 “語彙・構文的言い換え表現の自動生成”, 情報科学研究 (25), 95-99, 2008/2.
- [13] 近藤恵子, 佐藤理史, 奥村学(2001). “格変換による単文の言い換え” 情報処理学会論文誌, Vol. 42, No. 3, pp. 465-477.
- [14] 寺村秀夫(1982-4)『日本語のシンタクス』I, II, くろしお出版.
- [15] 亀井孝・河野六郎・千野栄一(1996)『言語学大辞典』第6巻術語編 p1171r, 三省堂.
- [16] (財)海外技術者研修協会(AOTS)(1974)『日本語の基礎I』スリーエーネットワーク.
- [17] (財)海外技術者研修協会(AOTS)(1998)『みんなの日本語初級I本冊』スリーエーネットワーク.
- [18] 野田尚史(2005)『コミュニケーションのための日本語教育文法』くろしお出版.
- [19] 森田良行(1989)『基礎日本語辞典』角川学芸出版.
- [20] 国際交流基金(1994)『日本語能力試験出題基準』凡人社.
- [21] グループ・ジャマシイ(1998)『日本語文型辞典』くろしお出版.
- [22] KNP: <http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/knp.html>
- [23] JUMAN: <http://nlp.kuee.kyoto-u.ac.jp/nl-resource/juman.html>
- [24] XML: <http://www.xml.org/>
- [25] 土岐哲・関正昭・平高史也・新内康子・鶴尾能子(1995)『日本語中級J301』スリーエーネットワーク.
- [26] 中尾桂子・森下淳也(2007)「読解支援のための語彙統計量に基づくテキストの動的拡張——語彙の重要度に基づくテキストの簡略化——」情報処理学会シンポジウムシリーズ, 人文科学とコンピュータシンポジウム2007: ポスター発表.
- [27] 中尾桂子・森下淳也: 語彙の重要度に基づく文構造自動書き換えシステム——「やさしい日本語」への変換による読解支援——, 人文科学とコンピュータシンポジウム論文集, Vol. 2008, No. 15, pp133-140, 2008.
- [28] 中尾桂子・森下淳也: 基本文型と語彙の重要度を用いた強調表示による読解支援——, 人文科学とコンピュータシンポジウム論文集, Vol. 2009, No. 16, pp256-234, 2009.