

表示切替による読解支援システムでの「基本文型」の有効性

中尾 桂子

大妻女子大学短期大学部

森下淳也

神戸大学大学院国際文化学研究科

日本語学習者を対象とした読解支援では、読解時に必要なスキルを明示した教育的な示唆のあるものが望ましく、同時に、様々なテキストに対応できることが望まれる。これまで我々の目的に合うシステムの開発に取り組んできたが、これまでのシステムでの表示では、日本人利用者からも強調表示の意味が理解しにくく、それは「構文」情報を利用しないことによると指摘されることが多かった。そこで、「文型」による検索を厳密化するために、品詞情報、構文情報、意味論的、語用論的情報を統合的に利用する教育用「文型」を詳細に規定して文中の重要語句を強調表示する方法、ならびに、構文解析（係り受け解析）でテキストを処理した結果で表示語を特定する方法を比較した。結果、留学生にとっては文型に基づく強調表示の方が学習者に馴染みのある方法であり、表示上の手がかりとして有益であることから、「文型」情報の利用の妥当性が判断された。

Effectiveness of Educational "Sentence pattern" in comprehension support system by display switch

Keiko Nakao

Department of Junior College
Otsuma women's University

Jyun-ya Morishita

Graduate School of Intercultural Studies
Kobe University

We suggest a Reading Comprehension Support System. This is an interface system emphasizing the indication of the text for an index at the pattern of a basic sentence and importance of the vocabulary. First, system analyzes the input text data and second, it converts that into structured text data by XML. At the time of conversion, it is searched detailed condition setting by incorporating analysis in the low rank of the text as a text attribute. The 10 learners said that it was able to expect some supports as a result of effectiveness interviewing comparing two display methods by this system. We can say that this system has utility in the beginner's comprehension support.

1. はじめに

日常生活でのやりとりでは問題がなくても、長文やまとまった文章から内容を正確に読解するには少々問題があるという日本語弱者を対象とした読解支援においては、読解時に必要なスキルを明示した教育的な示唆が可能な読解支援が望ましく、同時に、特定のテキストだけではなく、希望する様々なテキストに対応できることを目的としたシステムであることが望まれる。

そこで、中尾・森下は、1999以降、読解支援システムの開発に取り組んできたが[1][2]、我々の検討してきた読解支援システム”Text Viewer”では、必須情報の部分を強調表示にする方法で、読解スキル向上のための手がかりを残すという、ある種の「簡略化」を提案している。ここで言う簡略化とは、いわゆる、自然言語処理の分野で行われる研究成果を駆使したものとは異なり、簡略化を短文化とは考えず、どの部分が副次的でどの部分が必須

の情報か学習者に明らかにされるべきであるとの観点に立つ。そのため、述語と格関係のある必須要素を残して副次的な部分のテキストを消すという方法ではなく、文中のどこをどのような理由で見るとすべきかという文章中の着目点を指示することによる読解支援を実行しようとしている。

我々のシステムでは、文の構造明示の方法として、語の組み合わせパターンを元にした強調表示を考えている。いわゆる、教育用基本文型のみを利用するもので、その理由は、読解等の手がかりとして、留学生や日本人に心理的に馴染みがあることによる。

しかし、これまでのパターン検索法では、強調表示による強調の意図が一般日本人には理解しにくく、また、協調された情報が見にくいという問題が指摘されていた。現在、日本語教育で、主に中級以上で利用される教育用基本文型は、品詞情報、構文情報、意味論的、語用論的情報を統合した colligation 的な情報である。そこで、パターン検索をより厳密に

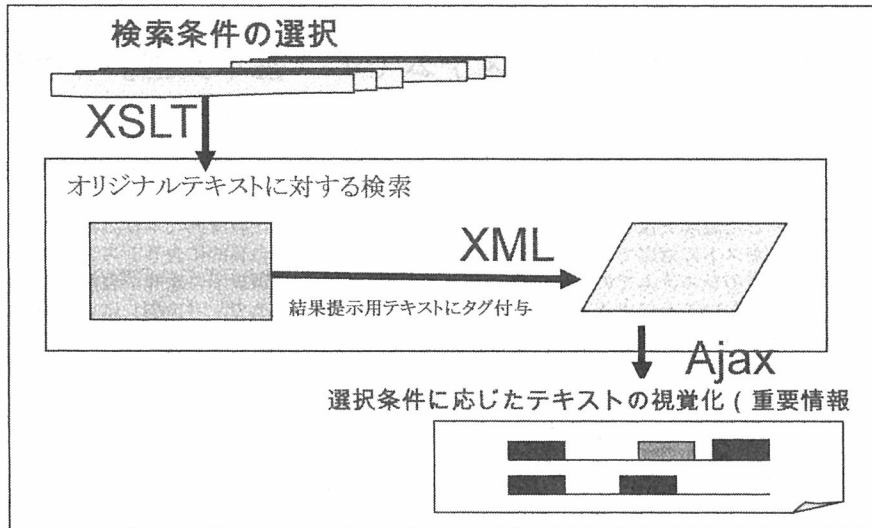


図1: 補填情報の付与の流れ(xslt 検索⇒xml 変換⇒ブラウザ表示)

するために、基本文型の正規化の際、基本文型を構成する colligation 的な統合情報をより厳密に検索に利用する方法を検討する必要があるが、これまでは、文型を深く考察せず、一列に扱っていたのみである。指摘された問題点に対処すべく、本稿では、読解時の手がかりとして文型情報をどのように指定するかを検討する。以下、まず、これまでのシステムとその問題点を挙げ、次に、「文型」と「構文」の違いを踏まえた検討方法について、さらに、変更による表示例を比較、検討し、最後に、今回の手法の有効性、課題について、留学生へのアンケートに基づき、考察した内容を報告する。

2. 読解支援システム"Text Viewer"

教育効果を念頭に置いた読解支援システムを実現するために、テキストの不要な箇所を消してしまうのではなく残したまま、必要な箇所だけを強調表示するというテキストの表示切替による「簡略化」方法での読解支援システムを検討してきた。

可逆的な表示切替には、XML を利用してテキストに構造を与え、テキストにはない品詞、漢字の読みといった深層情報をテキスト部分の属性として付与した。「文型」パターンリストに基づき、XSLT でテキストを検索した結果は、別の XML に吐き出し、それに対して HTML でマーカーやライン、字色の表示を切り替えて表示している。概略を図1に示す。

また、強調表示の他に、イメージ画像と辞書、KWIC 検索の機能を補填させている。これは各単語単位で、その文字列をキーワードとしてネット上のイメージ画像や別途リンクさせている電子辞書を参照情報として呼び出すもの、読解テキスト中の他の使用方法を KWIC 検索するというものである。図2に呼び出した情報が個別の資料となった状態の概念図を図2に示す。

このようなある種の「簡略化」を実行するためには、同じ文章を様々な視点で、多様に变化させるということになるが、これは、すなわち、オリジナルの情報を保持したままで可逆的に数種の表示切替を可能にするという仕組みであり、必要に応じて、つまり、利用者のレベルに合わせて補填する情報の有無を取捨選択できるということである。

3. 読解支援システム"Text Viewer"の課題

本システムの問題点としてこれまで指摘されてきたことには、意味を理解する上で必要

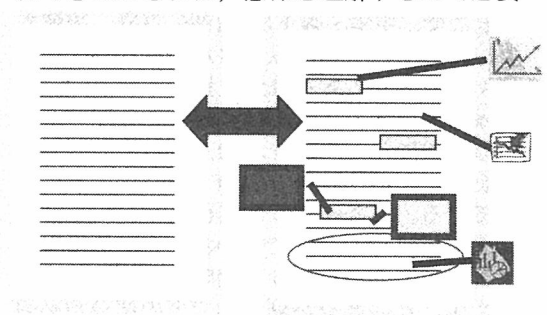


図2: 元データの独自利用イメージ

ハンゴウスイザンをするとき、薪がうまく燃えず、煙だけが出ることもある。しかし、薪の並べ方を工夫すると、よく燃えるようになる。

火を燃え続けさせるには、どのようにすればいいのだろう。

図 3: 語彙中心の表示マーカー採用時

な部分がどこかわかりにくいというものがあった。インターフェースの問題もあるが、これはより根本的な問題で、マークする情報の選択条件の規定の甘さによると考えられる。

これまで、表示の対象となる構文上の重要語の特定には、述語、その必須格と主題に着目して選んでいた。述語は、基本文型のうち、文末形式を指標にして、述語部分を示す指標「である」「だ」「です」「ます」の現在、過去、否定、過去否定、推量の形、「。」の前に最初に現れる動詞普通体と、その直前の名詞、動詞、形容詞 2 種と副詞を含めて述語と一括指定した。

また、他に、日本語教育の一般的な指導で用いられる文法要素を組み合わせた文型情報も利用した。例えば、「～は～です」といった「基本文型」で、述語のペアとなる要素は、ボタンとしての開始部分であるが、これは、文中の構文的関係ではなく、あくまでもいわゆる「基本文型」に即して、単純に、述語の前の最初に現れる助詞「が」「は」「を」とその直前の名詞としていたが、これは確かに乱暴な話である。傾向としては組み合わせて利用される文法項目で示される縦横情報である場合が多いが、そうでないものも含まれる可能性も割りにあることによる。

また、マークするということが重要な語は、文法項目という機能的なもので、読解時の内容には直接結びつかないため、実際の強調表示は、これらの重要文法機能語の前の実質的な意味を持つ名詞や動詞を対象としていた。図 3 にその表示例を示すが、このような「基本文型」の指定でカバーされる文型は、実は出現頻度が高いことから、ほぼ全ての文に含まれる情報である場合も多くなる。それが、最も特別視すべき情報の強調だけの表示切替につながりにくいというのが問題の原因の一つだったと言える。

また、基本文型で自然言語を判断するのは不十分であるという指摘も受けたことも考慮すると、正確に重要なものを選び、教育的にも内容読解の上でも役に立つ表示切替方法を検討するには、本当に、教育用基本「文型」

では難しいのか、また、自然言語処理の分野で利用される構文情報が有益なのか、指標となる「文型」を見直し、「重要」の意味を検討し直すことが必要であるだろう。「構文」情報と「文型」からの情報は、性質が異なる情報であるかどうか、特に、日本語教育を受けた留学生には、構文情報よりも「文型」の知識の方が意味の把握には有益であると考えられるかを考察し、システムの検索手法を考え直すために、本稿では、構文解析の結果も利用し、両方の検索結果を比較してみる。

4. 基本文型の正確な検索に向けて

4.1. 言語学的に見た「文型」と「構文」

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。

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。

一方の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るが、「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」にはパターンという文全体を典型的に捉える観点が低いと考えられる。すなわち、

「文型」は文の構成要素の組み合わせの型であり、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ということである。「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、学習者の誤用に對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で形成される「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とがとられることが多くなった。結果、年少者教育や大学留学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りない。

「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、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、それらを繰り返して練習し、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4にあげたような形で紹介されている。図4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。

私は [] です。
あなたは [] ですか。
私は [] ではありません。
あなたは [] ですか, [] ですか。
[] も [] です。

図4:『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」というものである。

そして、教材を利用した学習の経験を持ち日本語学習者は、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[8]。

5. 表示切替の指標としての文型

表示切替の指標として基本的な文型を利用する方法を考えるには3つ問題がある。

まず、文型そのものの存在が不明瞭なことである。教科書で扱われる文法項目はコマ切り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で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。この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の複数利用という問題がある。この文法項目を組み合わせで形成される学習用「文型」は、文法的な語だけではなく、品詞情報、構文情報、意味論的、語用論的情報を結合した colligation 的な情報でもある。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、データは1次元のものとして扱いたい。

そして、これが最も問題であるが、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ものがあり、その数は決して少なくはない。

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 ガ B ニ C ヲ出す」形式を取るものと、「A ガ 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ふ。また、「(相手 B)ニ」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ふのである。似たようなものはまだまだ多いだろう。

以上に対処するためには、3つのうち、まず、最初の1点目に配慮して「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。また、同列化、単純化のためには、検索条件としての文型情報を階層的に整理し、同時に、原義と文字列から問題がなければ、派生を文字列で統一するという変更を加える。残りの1点は、日本語能力試験に出るか出ないかと言えば、交流基金(1998)相当でなければ、出ないため、本来なら問題ではあるものの、ここでは対処済みとして話を進める。

(1) 検索のための「基本文型」の下処理

- ①『日本語文型辞典』の解説部分の情報参照して、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パター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

(2) テキスト構造の変更

- ①テキストは形態素解析を行い、解析結果を属性に格納するが、同時に別途係り受け解析を行い、その結果も別テキストに属性として盛り込み、再構造化しておく

(3) 強調表示の変更

- 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

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ずる) のも無理はない。

図 5: 「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ずるのも無理はない。

図 7: KNP による係り受け情報を利用した強調表示例

- 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」)

「文型」による表示例を図 5 に示す。図 5 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をかけた結果から一部を抜粋したものである。

6. 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。基本的に、文字列と品詞情報を基準にしており、意味論的、語用論の情報は、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[10]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 6 は、図 5 で用いた例文と同じ箇所の解析結果である。この結果に基づき、強調表示をかけたものが、図 7 である。図 7 は、図 6 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 7 のようになってしまった。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。

7. 留学生へのインタビュー

「文型」情報と、「構文」情報(係り受け解析結果)との表示箇所の違いを利用者の視点で比較するために、中、上級の留学生に、「読解時の手がかりを得ると考えれば、どちらの表示が好ましいか」質問した。

今回、協力いただいたのは、H 県の R 大学の学部留学生の中・上級以上のレベル 10 人で、漢字圏の学生が中心である。PC リテラシーには問題がないが、時間の関係で、出力を印刷したものを利用して印象に対する聞き取り調査(インタビュー)を行った。

10 人中 8 人は、文型に着目して、重要な語を見る「文型」情報に基づく図 5 のような文章の強調表示の方が馴染みやすく、また、学習にも便利であるという意見であった。残りの 2 人はどちらとも言えないということで、積極的には「文型」を選ばなかった。そして、文型の意味を追加する際、「説明だけではなく、例文もほしい」とか、「文型説明の表示が別ウインドウで開いて煩雑」だろうという

これ d が p₁
 普通だ j と p₁
 思っ て v いる s₁
 日本 C 人 n の p₁
 習慣 n から p₁
 する v と p、*
 その d₁
 時 n
 — n 回 s 限り s の p₁
 お礼 n だけ p で p は p
 物足りな j さ s を p₁
 感ずる v の n も p
 無理 n は p₁
 ない j。*

図 6: KNP による係り受け情報出力結果

意見を述べていた。インターフェースの問題に不足感を感じていたということだろう。

しかし、おおむね、「文型」による表示強調の方がわかりやすいという意見であった。

8. 考察

「文型」情報を検索に利用するには、構文解析結果の情報が必要だという意見を参考に、本稿では下処理とテキスト構造に変更を加えた。そして、表示結果として、「文型」ベース(図5に例示)と、「構文」ベース(図6に例示)を比較し、かつ、留学生に印象をインタビューした。

結果、「文型」を利用した方が、形態素解析に基づく「構文」結果よりも、語の組み合わせをまとめて取り扱えることから、留学生には文型の意味の把握や内容読解には有益だと受け取られることが確認できた。

中上級の漢字圏の学習者には、ある程度の助詞の機能から述部と格の関係を知ることができるため、どちらかといえば、語の組み合わせに意味がある「文型」を手がかりに、読解を進める方がよかったということだろう。

係り受け解析では、ベースが形態素解析であること、さらに、格関係の検出に強いことから、結果が美しいが、いわゆる「文型」の単位ではなく、あくまでも形態素単位での係り受け表示になり、語彙の拡張としてでしかキーワードを捉えられないことも、不十分な手がかりという印象を与えたようである。

ただし、今回、格関係の分別を上手く表示できなかったせいで、「構文」の表示に対する反応がよくなかったとも考えられる。述語に直接係る部分のみに強調表示をかける工夫をすると、動詞が問題となる文など、文章によっては、より有益な手がかりが与えられると考えられる。テキストの種類やジャンルを変えて、「構文」情報が望ましい場合と「文型」の得意な場合をさらに検証していく必要がある。

9. まとめ

日本語弱者を対象とした読解支援においては、読解時に必要なスキルを明示した教育的な示唆が可能な支援が望ましく、また、様々なテキストに対応できることを目的とした支援が望まれるため、これまで我々は読解支援システムの開発に取り組んできた。従来型のシステムでは、強調表示による強調の意図が

一般日本人には理解しにくいという問題があったため、パタン検索の厳密化のために、品詞情報、構文情報、意味論的、語用論的情報を統合的に利用する文法項目の組み合わせのパタンで文中の重要語句を強調表示した。今回はジャマシイ(1998)を利用した。結果、学習者のニーズに合う形で、表示をきめ細かく的確につけられるようになり、ある程度の質的向上が実現できた。ただの読解支援ではなく、自律的な読解を念頭に置いた教育的読解支援として向上が見られたと考える。

ただし、今後、さらに、文型自体の collocation 情報を検証することや、さらに、colligation 情報から文の構造に対する分類を行い、基本文型が基本的な文構造のパタンを示すものかどのようなものかについて再確認の必要性がある。これは、いわゆる、基本文型の見直しともなり、興味深い観点が得られることが期待できる。

参考文献

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] (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
- [6] (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
- [7] 森田良行(1989)『基礎日本語辞典』角川学芸出版.
- [8] 国際交流基金(1994)『日本語能力試験出題基準』凡人社.
- [9] グループ・ジャマシイ(1998)『日本語文型辞典』くろしお出版.
- [10] KNP:<http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/knp.html>