

日韓作文演習用知的CAIにおける文法の類似性を用いた誤文解析及び指導戦略

李 圭建+ 小西 達裕++ 高木 朗+++ 白井 克彦+ 小原 啓義+

+早稲田大学 理工学部 ++静岡大学 工学部 +++株式会社CSK

本稿では、韓国語作文演習用知的CAIの構築における日韓の文法的類似性を用いた誤文解析及び指導戦略について述べる。一般に日本語と韓国語は文法の類似性が高いため、基本的な指導法として日本語の単語を韓国語の単語で直接置き換えさせることによって作文させることにした。またこの指導法による学生の作文過程と文法の相違点に着目し、誤り原因の同定に必要な知識を次のように整理した。①表層文字列変形知識、②バグ知識、③日韓意味対応関係知識。さらにこれらの知識を用いる誤文解析手法、誤りに対する指導戦略等を提案した。また試作システムを構築し、その有効性を評価した。

Applying Grammer Similarities to the Japanese-Korean composition training ICAI
System for the Method of Composition Error Analysis and Advice Strategy.

Kyu Keon LEE+, Tatsuhiko KONISHI++, Akira TAKAGI+++,
Katsuhiko SHIRAI+, Hiroyoshi OHARA+

+Waseda Univ. ++Shizuoka Univ. +++CSK Corp.

In this paper, we propose the method of composition error analysis and advice strategy between Japanese and Korean by means of the applying grammer similarities to the Japanese-Korean composition training ICAI system. In general, Japanese and Korean language are grammatically very similar. So we choose a basical teaching method that a Japanese student is exchanged Japanese words for Korean words directly. And we summarise the necessary knowledge about the analysis of composition errors reason by means of composition process and grammer difference. ①knowledge of charactor string searching. ②knowledge of bug. ③knowledge of meaning relation between Japanese and Korean. And we propose the method of composition error analysis and advice strategy between Japanese and Korean using these knowledges.

1. まえがき

本研究は、学習者の母国語と教育対象である言語の間の文法的類似性が高い場合の教育支援を行う語学用知的C A Iの構築を目的とする。教材としては韓国語を選び、日本語を母国語とする初学者が韓国語を学習する際に学生が作文した入力文を解析し、評価する能力に重点をおいて述べる。一般に、日本語と韓国語は単語的、文法的に類似性が高く、文構造と語順の一致性が見られる^[1-3]。そこで本研究では、基本的な指導法として日本語の単語を韓国語の単語で直接置き換えさせることによって作文させることにした。これには作文において複雑な文法説明を行わなくても、直ちに学生が能動的に学習を行うことができるという利点がある。またこの指導法による学生の作文過程と文法の相違点に着目し、誤り原因の同定に必要な知識を以下のように整理した。

①表層文字列変形知識

②バグ知識

③日韓意味対応関係知識

さらにこれらの知識を用いる誤文解析手法及び誤りに対する指導戦略等を提案した。また試作システムを構築し、その有効性を評価した。

2. 基礎的考察

2. 1 日本語と韓国語の文法の類似性と、その教育システムへの利用

語学教育C A Iでは学習者の作文した文の誤りを発見し、それに応じた個別指導を行うことが望ましい。その誤りは学習者の母国語と教育対象の言語の類似性・相違点が原因となることが多いため、誤り発見における両国語の文法的類似性と相違点に着目することが有効であると考えられる。そこで対象とする日本語と韓国語がどの程度文法が類似し、また韓国語を学習する際にどれほど有効にその情報を使えるかを調べる必要がある。このことにより両国語の文構造と単語の意味対応関係から文法的類似性と相違点を整理した。その結果を表1に示す。また日本語に対する韓国語の作文の具体例を下記に示す

例①日：花子は 学生である

韓：花子NEUN 学生IDA

例②日：太郎は 毎日 電車で 行きますか

韓：太郎EUN 毎日 電車RO GABNIGGA

例③日：店で 赤い 帽子を 買った

韓：GAGEESEO BURGEUN 帽子REUL SASSDA

例④日：Aが 食べられました

韓：AGA MEOGEUSIEOSSSEUBNIDA

例⑤日：歩いて きたが、 遅くは なかった

韓：GEOLEO WASSJIMAN, NEUJJINEUN ANHASSDA

例①-⑤では漢字による名詞概念、係り受け関係、文構造、語の順序関係等で文全体について類似性が高い。このことから初等文（後述する取り扱う入力文の範囲）に関しては日韓の文法的類似性を用いて韓国語を作文する際に、日本語の単語を、対応する韓国語の単語で直接置き換えさせる方法でも作文が可能になる。この場合、初学者でも韓国語の作文が可能になり、作文において複雑な文法説明を行わなくても、直ちに学生が能動的に学習を行えるという点で優れていると考えられる。但し、一部（述部、助詞、助動詞等）において部分的に文法的な相違がある。例えば例①②の格助詞「は」は韓国語の音韻によって表層が異なる助詞である。例④の述部においての付属語の接続関係、例②③の活用規則、⑤の「歩いて」の変格活用規則等の表層文字列の変形の文法的相違と、例③の「赤い帽子」の連体修飾構造等の文構造的相違がある。さらに例②③の格助詞「で」、例④の助動詞「られる」等で語の意味対応関係が多対多である。このように学生が作文する際、単語を置き換えさせる指導法だけでは正確な文にならない場合もあり、学生にとって困難な部分となる。それらについては上述のような日韓の文法の相違点に限定できるので、それらの部分にきめ細かい教育を行うようなシステム設計を考えるべきである。文法の相違点に着目することには、①一般の語学用知的C A Iに比べて学生の誤り原因同定が容易になる。②アドバイスが提示しやすくなる。③システムが簡単に構築できる。等の利点がある。

表1. 日韓文法の類似性と相違点

類似性	相違点
自立語だけで文節が生成できる 自立語に付属語が接続される 文節内の順序には強い制約がある 文節同士の順序は係り受け関係が 交差しない 文構造が類似している 述部が文末に位置する 述部は述語以外に付属語が付く 修飾語は被修飾語の前に置かれる 格助詞により主語と述語の関係が 決定される 用言活用が激しい 述部が発達している 尊称語が発達している 漢字による名詞はほぼ同概念である	連体修飾構造が異なる 分かち書き法が異なる 助詞、助動詞の適用法、意味 対応関係が異なる 述部の付属語順序関係が異なる 修飾構造が異なる副詞がある 付属語は音韻属性に依存する 付属語の接続構造が異なる 尊称語表現が異なる 語の短縮及び省略形が異なる 用言活用が異なる 日本語は助動詞が豊富である 韓国語は時制表現が豊富である

2. 2 作文過程による誤り原因同定

入力文を解析して誤り原因を診断し、指導するためには学生の作文過程、即ち学生が持つ文法知識、用いた作文手法などを把握する必要がある。学生の作文過程を調べるために、初学者に韓国語文法の簡単な説明と単語辞書を与えて作文の実験を行った。その際の学生の作文行為より作文過程

と誤り原因を推定した。その関係を図1に示す。一般に、日本語の単語を韓国語で置き換える作文行為は次のようなプロセスを持つと考えられる。

- ①単語を選択する (見る=BODA, た=SSDA)
 - a-日本語の単語表層を選択する
 - b-日本語の意味を選択する
 - c-韓国語の意味を選択する
 - d-韓国語の単語表層を選択する
- ②単語を並べる (見た=BODASSDA)
- ③表層文字列を生成する (見た=BOASSDA)

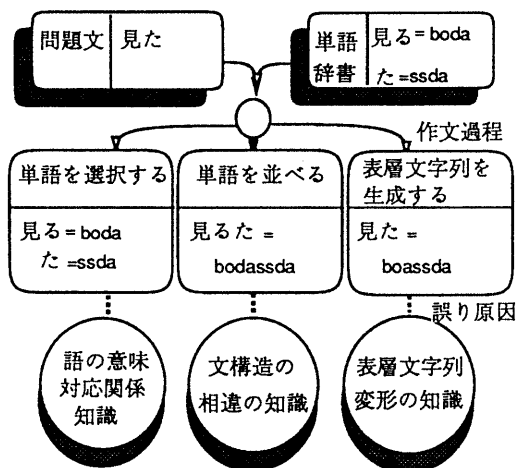


図1. 作文過程と誤り原因の相関関係

さらに作文実験から得た誤りと一般に韓国語作文において起こりやすいと考えられる誤りを推定し、作文過程別に分類した誤りタイプを表2に表す。学生の誤りは、各作文過程で誤った行為をしたり、間違った知識を用いた場合に起こり、その原因は次のように推定できる。

- ①単語を選択する際には辞書の引き間違い、記憶間違い、日韓の意味的対応関係知識による間違いなどがある(表2のA)。一般に両国語は単語の表層上、意味上の対応関係が多対多であり、助詞、助動詞句(述部)に学生が最も多く間違える原因が見られる。よってここでは単語の意味上の複雑な対応関係より間違った単語を選択する場合を、主な誤り原因とする。
- ②単語を並べる際には学生が母国語の文法知識を適用したため、連体修飾構造、分ち書き規則、係り受け構造等の文構造的相違点の知識を誤る場合が多い(表2のB)。
- ③表層文字列の生成や変形の際には語尾変形、活用規則、音韻適用規則等の韓国語の表層文字列を変形する知識を誤る場合が多い(表2のC)。

表2. 作文過程と誤り原因による誤りのタイプ

誤りのタイプとその内容	(日本語)	韓国語正解	韓国語誤り)例
(A) 単語選択に関する誤り			
①日本語表層に関する選択誤り			
・丁寧指定詞の選択誤り	私です	"IBNIDA"	"SEUBNIDA"
・異なる意味表層の選択誤り	見せる	"BOGE RADA"	"BOGE DOEDA"
②日本語意味に関する選択誤り			
・複数意味の選択誤り	見られる	"BOL SU ISSDA"	"BOGE DOEDA"
③韓国語意味に関する選択誤り			
・複数意味の選択誤り	見せる	"BOGE RADA"	"BOSIKIDA"
④韓国語表層に関する選択誤り			
・助詞「の」の特殊用法誤り	本の前	"CHAEG AP"	"CHAEGEUI AP"
・複数表層の選択誤り	～られる	"~BADDA"	"~DANGRADA"
⑤類似語に関する選択誤り	家	"JIB"	"JIP"
(B) 単語並びに関する誤り			
①語の欠落に関する誤り	家で	"JIBEURO"	"JIB"
②語の過剰に関する誤り	家に	"JIBE"	"JIBEDO"
③連体修飾構造に関する誤り			
・連体形を終止形とする誤り	行く人	"GANEUN SARAM"	"GADA SARAM"
・連体形活用語尾の適用誤り	見た人	"BON SARAM"	"BODAN SARAM"
・連体修飾の時制適用誤り	行く人	"GANEUN SARAM"	"GAN SARAM"
④分ち書きに関する誤り	行った人	"GAN SARAM"	"GANSARAM"
⑤述部順序に関する誤り			
・否定表現の述部の順序誤り	ません	"JI ANHSEUBNIDA"	"BNJJI ANHDA"
・丁寧表現の述部の順序誤り	ました	"SSBNIDA"	"BNISSDA"
・時制を含む述部の順序誤り	なかった	"ANHASSDA"	"ANHEOSSASSDA"
⑥語順に関する誤り			
・副詞修飾の順序誤り	もう一杯	"HAN JAN DEO"	"DEO HAN JAN"
⑦時制に関する誤り	行った	"GASSDA"	"GAN"
(C) 文字列生成に関する誤り			
①綴りに関する誤り			
・韓国語のマ字表記法間違い	座る	"ANJDA"	"ANZDA"
・韓国語外来語表記法間違い	バス	"BEOSEU"	"BBEOSU"
・韓国語文字構成規則間違い	知らせる	"ALRIDA"	"ALLIDA"
②付属語音韻適用に関する誤り			
・助詞音韻適用に関する誤り	私は	"NANEUN"	"NAEUN"
③用言活用に関する誤り			
・格基適用誤り	行って	"GASEO"	"GABOSEO"
・変格用言活用誤り	助けて	"DOWASEO"	"DOUSEO"
④語尾変形に関する誤り	して	"HAYBOSEO"	"HAESSEO"
④品詞転成に関する誤り			
・動詞による転成名詞誤り	読み	"IRGGL"	"IRGDA"
・形容詞による転成名詞誤り	赤さ	"BURGEUM"	"BURGDA"
⑤動作状態表現に関する誤り	私は行く	"GANDA"	"GADA"

2. 3 誤文解析に必要な知識

ここでは、誤り解析のためにシステムが持つべき知識を、その解析順に考察する。

2. 3. 1 文字列生成の誤り解析に必要な知識

表層文字列の生成や変形の際には、語尾変形、活用規則、音韻適用規則等の表層文字列を変形する知識の適用を間違ったために誤りが生じることが多い。この種の誤りは、前後の単語の文法的、意味的、音韻的關係から単語を変形して表層文字列を生成する知識を用いれば誤り検出が可能になる。このタイプの知識を表層文字列変形の知識と呼ぶことにし、表3に表す。誤り解析は次のように行うことができる。まず形態素解析を行って文中で用いられている単語を同定し、これに表3の知識を適用して正しい変形が行われているかを調べる。韓国語の形態素解析には単語辞書と活用テーブル(活用語尾、接辞、語尾等の用言に接続される付属語)が必要となり、その結果は文字列変形の知識を検索できる情報を持つ必要があるため、韓国語の音素、音節、格、語の音韻的、文法的、意味的属性を用意すべきである。例えば[座る=A NJDA]を[ANZDA]と綴り誤りを犯した場合、単語

辞書と韓国語ローマ字表記規則、文字構成規則等の知識が必要となる。また同じ意味の助詞でも表層が異なる助詞を適用した誤りは、単語辞書、助詞音韻適用則（先行語の最後の文字の属性により後行助詞が決まる規則）が必要である。さらに用言活用誤りは単語辞書と活用形、語基などの属性を用意した活用テーブル、用言活用規則の知識等が必要である。この手法を表層文字列変形知識を用いる誤り解析手法と呼ぶことにする。

表3. 韓国語の表層文字列の変形に関する知識

区 分	用いる表層文字列を変形する知識
①音韻構造	音素：半母音，単母音，重母音音素の構成 子音音素の構成 音節：単母音，重母音等の音節構成 子音+母音の開音節の構成 子音+母音+子音の開音節の構成 韻律：音の強度よりは音の高低による音度 接頭辞，接尾辞，複合語，転成名詞の語彙 漢字語の表記，意味，発音の規則 仮名混用・仮名表記のハングル表記規則
②語彙	基本構造：主格，補格，提示格が用いられる 拡大構造：連体格，接統格，処所格，器具格 資格格，比較格，与同格が用いられる
③格助詞	形態論的 意味論的
④活用語尾	助詞間の結合度が高い 多義性 正格活用 変格活用：D, L, B, S, REU, EU, U, YEO, JEO, H, GEORA, NEORAの変格活用
適用規則	母音調和規則が適用される 終止形，転位形，接統形，装統形の活用則 動詞は態，時相，意志，試図を持つ 語基が適用される
⑤指定詞	役割，音韻，接統法語形，転位法語尾
⑥時制 時相 態：使役 受身	過去，現在，未来，完了，進行の時制 継続相，完了相，結果相等の動作の表現 使役補助語幹，使役助動詞，使役接尾辞 受身補助語幹，受身助動詞，受身接尾辞
⑦照応	人称，指示表現
⑧待遇表現	終結語尾の形態によって異なる 平叙文，疑問文，否定文，命令文，請誘文
⑨慣用表現	感謝，紹介，依頼，状況等の表現 成句：否定，対立，その他の表現

2. 3. 2 単語並びの誤り解析に必要な知識

学生が日韓の文法の類似性を用いて単語を並べる際には日韓の連体修飾構造，係り受け構造等の文構造的相違点で誤りが起こるため，日韓の係り受け関係の相違を調べることで誤りが検出できる。よって入力文の形態素列と，文構造の相違を判断できる知識が必要となる。この種の誤りは構文解析，意味解析を行って誤りを検出する場合が多い。しかしこの題材では文構造の類似性が強く，構文構造を解析しなくても語の並びを調べるだけで誤り検出ができる場合が多い。また誤りを含む文の構文解析^[4]は語の係り受けを判断する知識が膨大になるため極めて困難であり，構文解析を行ってもその精度を高めることは非常に難しい。よって

ここでは文構造の相違を調べる知識を文法の相違点で限定できるために構文解析を行わずに，バグ知識を用いて誤りを検出する。

一般に，知的CAIにおけるバグ知識は，そのバグを十分に用意することが難しい。しかし文法の類似性が高く，句構造がほぼ同じであるため，初等文において単語を直接置き換えさせる作文を題材とした場合のバグの数はそれほど多数必要とされない（我々が構築したシステムでは100程度のバグを想定しており，これで初等文についてはほぼサポートできると考えられる）。よって文法の相違点についてのバグ知識をシステムに用意することが有効であると考えられる。バグ知識は条件-結論型の一定の構造を持つ。また表4のようなバグ知識を数種類ずつ用意し，誤り検出に用いる。例えば日韓の連体修飾構造が異なるために起こる誤りは，連体修飾構造の混同を検出するバグ知識が必要となる。この手法をバグ知識を用いる誤り解析手法と呼ぶことにする。

表4. 文構造的相違によるバグ知識

- | | |
|---------------------|---|
| ①語の欠落を検索するバグ知識 | ・必須語，文節の欠落を検索する |
| ②語の過剰を検索するバグ知識 | ・語，文節の過剰を検索する |
| ③連体修飾構造の混同を検索するバグ知識 | ・連体修飾と終止形の混同を検索する
・連体形活用語尾適用の混同を検索する
・時制を含む連体修飾の混同を検索する |
| ④分ち書き誤りを検索するバグ知識 | |
| ⑤述部の係り受け誤りを検索するバグ知識 | ・否定表現の係り受けの混同を検索する
・補助語幹の係り受けの混同を検索する |
| ⑥語順の係り受け誤りを検索するバグ知識 | ・副詞修飾の係り受けの混同を検索する |
| ⑦時制の混同を検索するバグ知識 | ・叙述形と連体形の時制の混同を検索する
・未実現と実現の時制の混同を検索する
・現在時制と未来時制の混同を検索する
・過去時制と完了時制の混同を検索する |

2. 3. 3 単語選択の誤り解析に必要な知識

日韓の単語は多対多の意味対応関係を持つために単語を選択する際に誤りが多く生じる。この種の誤りは，学生が単語選択の過程でどの段階で誤ったかによりその原因が異なる。このため誤り検出には日韓の語の意味対応関係を探索できる知識が必要となる。2. 2で述べたように，単語を選択する過程はまず日本語の単語からその意味を抽出し，それに対応する意味を持つ韓国語を選択する。この際，語の意味的対応関係を，日韓の表層語と意味をノードとし，そのノード間の関係をアークとする図2のようなネットワークで表し，誤り検出に用いる。このとき学生の単語選択の過程は，このネットワークにおける日本語の単語から

韓国語の単語への探索としてモデル化できる。学生がこの探索のどこで誤ったかを知るためには、問題文の単語、正解文の単語、入力文の単語が図2のネットワーク上でどのようなパスで接続されるかを調べ、そのパスの分岐位置によって判定すればよい。例えば図2で問題文の単語がa1、正解文の単語がd1、入力文の単語がdnの場合、問題-正解文のパス[a1-d1]と入力-問題文のパス[dn-a1]を辿ると分岐する位置[J1]が誤り（問題文と異なる意味の作文をした誤り）であると推定できる。そこでこのネットワークを検索するために日韓の表層語と意味をフレーム構造で表す意味対応知識を用意する必要がある。この種の誤りは韓国語作文において頻繁に起こる誤りである。特に日韓における意味の相違が著しい助詞と連語（例：～することが出来る等）を含む助動詞句の述部で多く見られる。そこで本システムでは日本語の助詞と助動詞句についてこの方法を用いる。

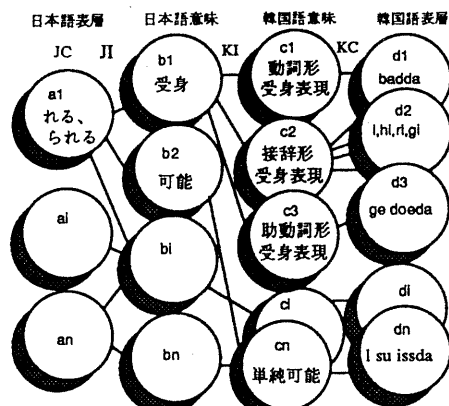


図2. 日韓意味対応関係

2. 4 指導に必要な知識

教師は学生の理解状態の探査、新しい知識の提示、誤りを修正するための説明など、種々の教育行動を行う。また教え方としては学生の個性、誤り原因によって直接その誤りを指摘したり、ヒントや例文を用いる文法的説明等、学生自ら解けるように誘導しつつ、理解しやすい形で教える。例えば綴り誤りが辞書引き間違いによる場合、教師は文法的知識よりは直接その誤りを指摘する。反面、用言活用規則誤りはその活用知識を学生に教えないといけない。よって誤り原因毎に次のような教示生成知識を必要とする。

- ①韓国語の表層文字列変形知識に関する教示、
- ②日韓の文法の相違点に関する教示、
- ③日韓意味対応関係に関する教示、
- ④その他の辞書引き法、正しい入力方法等に関する教示の知識

またこれらの知識を誤り種類、誤り原因を表す内容、ヒント、文法知識説明等のテンプレートで構成し、誤文解析や教示生成に必要な情報を参照可能な形で作成しなければならない。さらに指導戦略として日本語の単語を韓国語で直接置き換えさせる指導法を用いるため、学生が作文を行う際には常にその指導法を強調し、混同しないように指導する必要がある。

3. システムの構成

上述のような誤文解析と指導能力を持つ知的C A Iシステムの実現方法について述べる。

3. 1 システム概要

本稿で報告するシステムの構成を図3に示す。システムは出題部、入力文解析部、誤文解析部、教示生成部から成り、演習問題・正解集、単語辞書、活用テーブル、文字列変形知識、バグ知識、意味対応知識、教示生成知識等に関するデータベースを持つ。出題部は、演習問題を学生に提示する。問題の選択は学生が主導する。入力文解析部は学生が韓国語ローマ字表記で作文した入力文を、単語辞書と活用テーブル等を用いて形態素解析を行う。誤文解析部は入力文解析部の結果と2. 3で述べた知識を用いて誤りを解析する。教示生成部は誤文解析の結果と2. 4で述べた知識を基にアドバイスを生じ、教示する。システムへの入力はキーボードより行う。

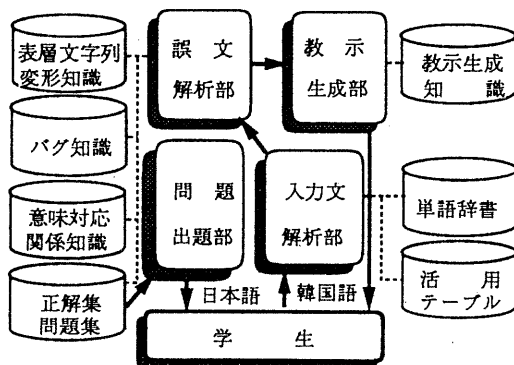


図3. システムの構成図

3. 2 入力文解析部

入力文解析部では、韓国語入力文を解析し、誤り解析に必要な文法的、音韻的、意味的性質を持つ形態素列をリスト構造で生成する。韓国語は文節毎に分かち書きされ、自立語に付属語（助動詞、接辞、語尾、助詞、指定詞）が接続する性質がある^[2]。また活用の際には正格用言活用と変格用言活用等があり、不規則に用言が活用する性質を持つ^[5]。よって韓国語の入力文解析は、単語辞書と活用テーブルを用いて最長一致法での形態素解析

を行い、次のような段階を経て誤り解析に必要な情報を持つ形で形態素列を生成する。

- ①入力文の文字列を各形態素単位で分離する
- ②用言の場合、分離された形態素の原形を作る
- ③各形態素間の接続関係を検査する

またシステム処理は体言部と用言部で、異なる処理方法で行う。例えば2.1の例④は体言部で名詞[A]、格助詞[GA]、用言部で動詞[MEOGDA]、補助語幹[EU SI EOSS]、語尾[SEUBNIDA]の形態素列がリスト構造で生成される。

3.3 誤文解析部

誤文解析部は、誤り解析器、誤り訂正器、誤り探索器で構成される。処理手順は次のようである。

- ①表層文字列変形知識を用いて誤りを検出する
- ②誤った語を正しい語に訂正する
- ③訂正された入力文の形態素列と正解文の形態素列を検索し、誤りの位置を決定する
- ④バグ知識を用いて誤りを検出する
- ⑤意味対応関係知識を用いて誤りを検出する

これらの手順を繰り返すことによって文中にある誤りを検出する。以下、システムの処理手順に従って説明する。

3.3.1 文字列変形知識による誤文解析

ここでは、韓国語入力文の形態素解析結果に、語の前後の意味的、文法的、音韻的性質の関係を検索する文字列変形知識を適用することで誤りを検出する。処理手順は次のようである。

- ①入力文形態素列に文字列変形知識を適用する
- ②形態素列の前後の関係を探索する
- ③誤った語を正しい語に訂正する

これらを繰り返し、文全体について検索する。例えば[彼は]を[GEUEUN]と作文した場合、形態素列に助詞音韻則を用いて誤り検出を行うと共に、[GEUNEUN]と正しい文で訂正する。また[食べれば]を[MEOGMYEN]と作文した場合、形態素に用言活用語の語基適用則(第2語基)を用いて語幹と語尾を調べることで誤り検出ができる。こうして[MEOEUMYEON]と訂正され、次の誤り解析処理に送られる。

3.3.2 バグ知識を用いる誤文解析

ここでは、バグ知識、正解文の形態素列、入力文の形態素列、単語辞書、活用テーブル等を用いて誤り解析を行う。処理手順は次のようである。

- ①正解文の形態素列と入力文の形態素列を検索し、文節、または文節内の単語の誤り位置を決定する
- ②未処理のバグ知識を検索し、条件部が真の場合、結論部で誤り仮説を生成する

バグ知識の条件部処理は各バグ知識が持つ入力文の誤りを判定する関数を実行することで誤りを検出する。結論部処理は、誤り内容、正解文と入力文の検索内容などをリスト構造で生成する。例えば[行く人]は連体修飾構造であるが、学生が[行く]を終止形を用いて[GADA]と作文した場合、[連体形と終止形の混同を検索するバグ知識]で条件部が真になるため、ここで誤りが検出される。これらのバグ知識の代表例を表5に示す。

表5. バグ知識の代表例

(A) 連体修飾構造の混同を検索するバグ知識

- ①正解文が連体修飾構造で、
 - ・入力文が用言+終止形活用語尾+体言の場合、連体修飾と終止形活用語尾の混同の誤り仮説を生成する
 - ・入力文が用言+転位形活用語尾+体言の場合、連体修飾と転位形活用語尾の混同の誤り仮説を生成する
 - ・入力文が用言+接続形活用語尾+体言の場合、連体修飾と接続形活用語尾の混同の誤り仮説を生成する
- ②正解文が動詞+(EU)N+体言で、
 - ・入力文が動詞+NEUN+体言の場合、動詞過去形の連体形時制の混同の誤り仮説を生成する
- ③正解文が動詞+NEUN+体言で、
 - ・入力文が動詞+(EU)N+体言の場合、動詞現在形の連体形時制の混同の誤り仮説を生成する
- ④正解文が形容詞+(EU)N+体言で、
 - ・入力文が形容詞+NEUN+体言の場合、形容詞現在形の連体形時制混同の誤り仮説を生成する
- ⑤正解文が用言+(EU)L+体言で、
 - ・入力文が動詞+(NEUN, (EU)N)+体言の場合、未実現形の連体形時制の混同の誤り仮説を生成する
- ⑥正解文が用言+(EU)L+形式名詞で、
 - ・入力文が用言+(NEUN, (EU)N)+形式名詞の場合、推測形の連体形時制の混同の誤り仮説を生成する
- ⑦正解文が用言+DEON+体言で、入力文が用言+(NEUN, (EU)N, (EU)L, SSDEON)+体言の場合、回想形の連体形時制の混同の誤り仮説を生成する
- ⑧正解文が用言+SSDEON+体言で、入力文が用言+(NEUN, (EU)N, (EU)L, DEON)+体言の場合、完了回想形の連体形時制の混同の誤り仮説を生成する

(B) 時制の混同を検索するバグ知識

- ①正解文が用言+時制補助語幹+終結語尾で、
 - ・入力文が用言+連体形活用語尾の場合、叙述形時制と連体形時制の混同の誤り仮説を生成する
- ②正解文が用言+GESS+終結語尾で、
 - ・入力文が用言+(N, NEUN)+終結語尾の場合、未来形時制の混同の誤り仮説を生成する
- ③正解文が用言+(N, NEUN)+終結語尾で、
 - ・入力文が用言+(SS, SS EOSS, GESS)+終結語尾の場合、現在形時制の混同の誤り仮説を生成する
- ④正解文が用言+補助語幹①+終結語尾で、
 - ・入力文が用言+補助語幹②+終結語尾の場合、時制補助語幹の順序の混同の誤り仮説を生成する

【用語定義】

- a) 連体修飾構造：用言+連体形活用語尾+体言
- b) 連体形活用語尾：(EU)N, NEUN, (EU)L, DEON, SSDEON,
- c) 時制補助語幹：過去完了形-SS+EOSS, 過去形-SS, 現在形-N, NEUN, 未来形-GESS (未実現, 推測)
- d) 補助語幹①：SS+EOSS+GESSの順
補助語幹②：EOSS+SS+GESSの順, EOSS+GESS+SSの順, GESS+SS+EOSSの順, GESS+EOSS+SSの順

3.3.3 日韓意味対応関係ネットワークを用いる誤文解析

ここでは日韓の両国語における意味的相違が著しい助詞と連語を含む助動詞句の述部で頻繁に起こる誤りを解析するため、図2のネットワーク上に約100種類の日本語の助詞、連語を含む助動詞句^[6]と、それに対応する韓国語単語のデータを用意し、誤りを検出する。誤り推定は問題文の単語、正解文の単語、入力文の単語で、日韓意味対応知識のフレームの上下位ノードスロットを探索すると、問題文-正解文のパス、入力文-問題文のパスが導かれる。それらのパスを辿ると、その分岐点が誤り位置である。通常、問題に対する正解は必ずしも一つではなく、問題によって複数の正解を持つ場合が多いため、ここで生成されるパスは複数である。また入力文の単語がネットワーク上にない場合、学生が問題文を理解できなかったか、勘違いして作文したと推定できる。例えば「TVが見られる」(可能)を「TVEUL BOGE D OEDA」(受身)と作文した場合、図2のネットワークを辿ると、「JI」で日本語単語が持つ複数の意味を、学生が混同し、間違った意味を選択したと推定できるため、それに応じた教示も容易となる。

3.4 教示生成部

教示生成部では、誤文解析結果と教示生成知識を基にアドバイスを生成し、学生に教示する。教示生成器は誤文解析器と対応づけられ、各々構成される。教示生成知識は誤り毎に誤り種類、誤り内容、ヒント、文法知識説明及び例文等のテンプレートで構成されているため、誤文解析の結果を、教示生成知識のテンプレートに直接埋め込む形で次のように生成する。

- ①誤り種類は誤り原因と対応づけ、生成する
- ②誤り内容は誤文解析結果を用いてテンプレートにデータを書き込む形で生成する
- ③ヒントは学生が自ら誤りを訂正できるような知識を生成する
- ④文法知識の説明は誤り内容と対応づけられ、生成される。このとき、補足説明として例文を用いる場合がある。

また指導戦略として学生が単語を直接置き換えて作文するように、随時その知識を注意しつつ、次のような教示手法を用いる。

- ①表層文字列の生成や変形の際に、間違った文字列変形の知識を適用し、文字列に異常がある場合は日韓の表層文字列を変形する知識を明らかにさせて説明する。
- ②係り受け、文構造等の相違による誤りは日韓の文構造、係り受けの類似性と相違点を比較

- しつつ、その関係を明らかにさせて教示する。
- ③日韓の意味対応関係知識による誤りは、単語を選択する際に単語選択上の注意点、両国語の意味上の対応関係などを詳しく説明する。
- ④その他、誤り原因別に指導手法を用意する。

4. 処理例

本稿で取り扱う入力文は、平叙文、疑問文、否定文等について、日本語の単語を韓国語の単語で直接置き換えられる文を中心とした。また学生の入力文として主節、従属節の反転、関係節構造の変形等の意味表現の変形規則^[7]のような文は取り扱わないことにした。そこで韓国語の基本的文構造を表6に示す。

表6. 韓国語の基本的文構造

基本的文構造	日本語の例文	韓国語の例文
①文の基本構造 (運用成分) +++ (運用成分) + 動詞・助動詞 (注1) + 助詞	私は山を歩いて行ったが	"NANEUN SANEUL GEOLEOSEO GASSJIMAN"
②名詞句 ・ (連体修飾表現) +++ (連体修飾表現) + 名詞 ・ 数詞 ・ 代名詞	美しい花 一つの これは本です	"AREUMDAUN GGOJ" "HANA EU" "IGEOSEUN CHAEG IBNIDA"
③運用成分 ・ 副詞 ・ 形容詞運用形(形容動詞含む) ・ 動詞の運用形 ・ 連体修飾節 ・ 名詞句 + (指定詞) + 助詞	一番大きい 美しかった 歩いて行く 歩く人がいる 美しい花が	"JEIL KEUDA" "ARUMDAWEOSSDA" "GEOLEOSEO GADA" "GEODNEUN SARAMI ISSDA" "AREUMDAUN GGOJ"
④連体修飾表現 ・ 関係節(主格関係節, 目的格関係節, 所有格関係節) ・ 形容詞連体形(形容動詞含む) ・ 動詞の連体形	私は彼を愛した人を知っている 美しい花 歩く人	"NANEUN GEUREUL SARANGHA N SARAMEUL ALGO ISSDA" "AREUMDAUN GGOJ" "GEODNEUN SARAM"
⑤助動詞句 ・ 助動詞形活用語尾 + 動詞 ・ 動詞語幹 + (補助語幹) + ... + (補助語幹) + 語尾	歩いて見る 歩いた	"GEOLEO BODA" "GEOLEOSSDA"
⑥節: (運用成分) +++ (運用成分) + 動詞・助動詞 (注1)	山を歩いて見た	"SANEUL GEOLEO BOASSDA"

(注1: 助動詞, または補助語幹 + ... + 補助語幹 + 語尾)

以上の検討を基に構築したシステムによる誤文解析と教示例を図4に示す。システムはまず、学生に問題を出題し、作文させる。学生が韓国語のローマ字で作文すると、システムは前述の方法で誤りを解析し、アドバイスを提示する。学生はそのアドバイスを参照し、誤った文を訂正する。図4の[1]は助詞音韻規則を誤って文字列に異常があるため、文字列検索知識を用いる誤り解析より検出される。[2]は連体修飾構造の相違による誤りであるため、連体形と終止形の混同を検索するバグ知識を用いる誤り解析により検出される。[3]は問題文では助動詞「られる」が受身の意味で用いられているが可能な意味で作文した誤りであるため、学生が日本語の「られる」が持つ複数の意味と、対応する韓国語の意味関係を混同したと推定できる。従って日韓意味対応関係ネットワークを用いる誤り解析より検出される。次に、アドバイスは

それらの結果を元に誤り内容別にテンプレートに書き込んで生成し、学生に提示する。なお、システムはLispを用いて書かれており、ワークステーションSUN4上のKCLにより構築されている。プログラムはLispソースコードで約300KB、データは約200KBである。

>>教師：[問題1]=車道を歩く人は嫌われる

>>学生：“chadoueul geodda saramneun miumeul su issda”

>>> 誤り解析中です <<> 少々お待ち下さい <<<

>>教師：あなたの作文には「3」の誤りがあります。

それでは順に説明します。

[1]=[助詞音韻適用誤り]

[A]-[誤り内容]=あなたは(1)-[名詞, chado]に接続された[助詞, eul], (2)-[名詞, saram]に接続された[助詞, neun]について[助詞音韻適用]を誤っています。

[B]-[ヒント]=助詞適用の際には先行語の音韻により同じ助詞でも表層が異なる助詞があることを考えましょう。

[C]-[文法知識説明]=[助詞音韻適用]について説明します。先行語の末字が子音の場合は助詞の頭文字が母音で、母音の場合は子音です。但し並立助詞は先行語の末字が子音の場合は子音が、母音の場合は頭文字が母音の助詞が付きまゝ。例)[人は:saramneun], [私は:naneun]。

[2]=[連体修飾にすべきところを終止形にする誤り]

[A]-[誤り内容]=あなたは[歩く人]のところで、[終止形動詞, geodda]を用いて[名詞, saram]を修飾しています。この場合の韓国語表現は連体形活用語尾を用います。

[B]-[ヒント]=日韓の連体修飾構造について考えましょう。

[C]-[文法知識説明]=[連体修飾構造]について説明します。韓国語の連体修飾構造は連体形動詞、形容詞から連体形活用語尾を用いて体言を修飾します。

また連体形活用語尾は[n, neun, l, deon, deo]があり、時制属性を持っています。

(1)-[動詞]+[N]+[体言]の連体修飾は過去を表します。

[形容詞]+[N]+[体言]の連体修飾は現在を表します。

例)[見た人=bon saram]は過去を、

[美しい人=areumdaun saram]は現在。

(2)-[動詞]+[NEUN]+[体言]の連体修飾は現在を表します。

例)[食べる人=meogneun saram]

(3)-[動詞]+[L]+[体言]の連体修飾は未実現を表し、

[形容詞]+[L]+[形式名詞]は推測を表します。

例)[見る人=bol saram], [大きいもの=keun geos]

(4)-[動詞, 形容詞]+[DEON]+[体言]は過去回想を表します。

例)[食べたものだ=meogdeon geos]

[3]=[日本語表層文と意味の取り違い誤り]

[A]-[誤り内容]=あなたは日本語の[受身]の意味を持つ

[助動詞, られる]を[可能]の意味である韓国語表現の

[l su issda]と作文しています。

[B]-[ヒント]=韓国語表現の勘違いがないか確認しましょう。

[C]-[文法知識説明]=[助動詞, られる]について説明します。

[助動詞, られる]は[受身, 可能, 尊敬, 自発]等の意味を持ちます。またその意味によって韓国語表現も異なります。

>>教師：以上の誤りを“-”の形で直して下さい!!

>>教師：[問題1]=車道を歩く人は嫌われる

>>学生：“chadoreul geodneun saramneun miumeul badneunda”

>>> 誤り解析中です <<> 少々お待ち下さい <<<

>>教師：正解です。次の問題を出しましょうか?(YES/NO)=no

>>教師：終了です。

図4. 処理例

用できることを明らかにし、またその方法を示した。ここで述べた手法は、文法の類似性が高い言語の間の教育支援を行う知的CAIに応用することが可能であると思われる。今後の課題としては、ここで述べた誤文解析と指導戦略について、より一般性を高めるべく改善を進めると共に、複雑な文体を取り扱うための誤文解析について検討を進める。

【参考文献】

- [1] 菅野裕臣監修, 日本朝鮮語学研究会編著: “朝鮮語を学ぼう”, 三修社(1987)。
- [2] 黄燦ホ, 李季順, 張ソク鎮, 李吉鹿, ソウル大学語学研究所: “韓日語対照分析”, 明志出版社(1988)。
- [3] 早川嘉春著: “日本語から学ぶ韓国語会話”, 創拓社(1991)。
- [4] 今田敬, 小西達裕, 高木朗, 小原啓義: “機械翻訳技術を用いた英語教育用知的CAIシステム”, 人工知能学会, 第5回全国大会, (1991)。
- [5] Seung Shik Kang, Yung Taek Kim: “A Computational Analysis Model of Irregular Verbs in Korean Morphological Analyzer”, Korea computer science, Vol. 19, No. MAR-CH, pp151-164(1992)。
- [6] 日本国立国語研究所編: “現代日本語の助詞と助動詞—用法と実例—”, 秀英出版(1987)。
- [7] 高木朗, 伊東幸宏共著: “自然言語の処理”, 丸善(1987)。

5. むすび

本稿では、韓国語を題材とする語学教育用知的CAIシステムにおいて、誤文解析と指導を行うために日韓の両国語の文法の類似性と相違点が利