

日韓文法の類似性を利用した韓国語作文演習
知的CAIシステムの構築に関する研究

2S-7

李 圭建* 小西 達裕* 高木 朗** 小原 啓義*

*早稲田大学 理工学部

**株式会社CSK

1. はじめに

本稿では日韓両国語の文法の類似性を用いた韓国語作文演習知的CAIシステムの構築について論じる。まず日韓文法の類似性と相違点を明らかにする。また学生が犯す誤りをタイプ分けし、それに適した誤り解析と教示手法を提案する。さらに試作システムを構築し、有効性を評価する。

2. 日韓文法の特徴と韓国語作文教育への利用

日韓の両国語は高い文法の共通性をもつ言語である⁽¹⁾。また両国語の教育においては文法の共通性を学生に認識させ、文法の相違点を強調する教育手法が一般に用いられる。本研究ではこのような教育手法を用いて日本人を対象とする韓国語作文演習知的CAIシステムを実現する。そのために日韓文法が具体的にどの程度共通性を持つかを調査し、次のように整理できた。文法の類似性は①文節内の順序には強い制約がある。②文節同士の順序は係り受け関係が交差しないことと、述部が文末に位置する緩い制限しか存在しない。③述部は述語以外に助動詞などの付属語が付けられることが多い。④漢字による名詞はほとんど同じ概念である。また相違点は①助動詞、助動詞句の適用法や意味に相違がある。②活用用法の違いがある。③韓国語は時制の表現が豊富である。このように日韓文法はほぼ同じ語順を持ち、文法の相違は制限されたところで生じる。従って初等文においては日本語文の単語を韓国語に置き換えるだけで作文が可能になる。この範囲では初学者にも簡単に作文させることができる。またこの作文法を用いた場合、学生の誤りは文法の相違から生じることが多いため、学生の誤り原因同定の際には文法の相違点に着目することで処理を効率化・単純化できる。これに対し日英等の文構造が全く異なる言語同士の教育においては、本稿で用いる手法とは異なる方法を用いる必要があると考えられる。

3. 学生の誤りタイプ及び誤り検出手法

3.1 学生の誤りタイプ

韓国語作文における学生の誤りを明らかにするために、韓国語の知識がない学生を対象に韓国語作文演習の予備実験を行った。実験の際には学生に文法の簡単な説明を行った後、単語辞書を与えて作文させた。その結果得られた誤りを学生の作文過程を基にタイプ分けした。一般に日本語の単語を韓国語で

置き換えて作文する場合に学生は次のような過程で作文すると考えられる。A-辞書を引いて韓国語の単語を選択する。B-単語を語順に従って並べる。C-韓国語の文法を適用させ、文字列の生成や変形を行う。この各過程における誤りを表1のA~Cに示す。

表1. 学生の作文過程による誤りのタイプ

誤りのタイプとその内容	誤り例(日本語:正解/誤り)
A. 単語選択の誤り	
①類似単語の混同誤り	家 : "JIB" / "JIP"
②日本語の選択誤り	私です : "IBNIDA" / "SEUBNIDA"
③日本語意味の選択誤り	見られる : "BOSIDA" / "BOGE DOEDA"
④韓国語意味の選択誤り	見せる : "BOGE HADA" / "BOSIKIDA"
⑤韓国語の選択誤り	見せる : "BOIDA" / "BOHIDA"
B. 単語並びの誤り	
①欠落の誤り	家へ : "JIBEURO" / "JIB__"
②過剰の誤り	家に : "JIBE" / "JIBEDO"
③連体形句の誤り	行く人 : "GANEUN" / "GADA"
④述部の順序の誤り	ました : "SSBNIDA" / "BNISSDA"
⑤時制の誤り	行った : "GASSDA" / "GAN"
C. 表層文字列生成の誤り	
①綴り誤り	座る : "ANJDA" / "ANZDA"
②助詞音韻適用誤り	私は : "NAEUN" / "NAEUN"
③分かち書き誤り	私は : "NAEUN" / "NA_NEUN"
④用言活用の誤り	行けば : "GAMYEON" / "GAEUMYEON"
⑤品詞転成の誤り	読む : "IRGGI" / "IRGDA"
⑥現在形表現の誤り	行く : "GANDA" / "GADA"

3.2 誤り検出手法

このような原因による誤りの検出手法についてシステムが行う処理手順に従って考察する。

(1) 辞書情報による誤り解析: 表1のCの誤りは、単語の屈折や変形に関する文法適用の間違い等の原因で、誤った表層文字列を生成する誤りである。この誤りの検出には単語辞書と、語の前後の音韻上の関係、付属語の意味上の関係等を表す活用テーブル等を用意する必要がある。誤り解析は韓国語入力文の形態素解析を行う際に、韓国語の文字列に上述の辞書情報などを適用しつつ行う。例えば同一意味の助詞でも表層が異なる助詞を適用した誤りの解析の場合、その助詞適用知識(先行語の最後の文字の属性により後行助詞が決まる)により入力文の助詞を調べて入力文の正誤が判定できる。また用言活用の誤りは形態素に活用形、語基などの属性を用意し、その形態素を調べることにより検出できる。

(2) バグ知識を用いる誤り解析: 表1のBの誤りは学生が日本語の語順に従って単語を並べる際に日韓の文法的相違により生じる。この誤りは日韓の文法知識の相違を調べることで検出できる。このことから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。一般に知的CAIにおいてバグ知識は、そのバグを十分に用意することが難しい。しかし日韓の文法の類似性が

Applying Japanese-Korean grammar similarities to an ICAI system for Korean composition training.

Kyu Keon LEE. *Tatsuhiko KONISHI.

**Akira TAKAGI. *Hiroyoshi OHARA.

*Waseda Univ. **CSK Corp.

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バグの数はそれほど多数必要とされない。この題材では構文構造を知らなくても語の並びを調べるだけで誤り検出ができる。よって構文解析を行わずに、形態素列レベルのチェックでも誤りを十分に検出できる。バグ知識は条件-結論型の一定の構造を持つ。バグ知識は表1のBの各々の誤りタイプに対して数種類ずつ用意する。誤り解析手法はまず入力文と正解との比較により誤り位置を決定し、次にバグ知識を適用して誤りを検出する。例えば[行く人: GANE UN SARAM]は連体形動詞が名詞を修飾する文構造であるが、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出するバグ知識」によって誤りが検出され、誤り仮説が立てられる。このバグ知識の条件部は正解文の構造が連体形句で、入力文の形態素の属性が用言、終止形語尾、体言の構造である場合に真になる。

(3) 日韓意味対応関係を用いる誤り解析: 表1のAの誤りは単語の記憶間違い、単語の意味関係の相違等の原因により起こる。また学生が単語を選択する際に、どの段階で誤ったかにより誤り原因は異なる。一般に学生が日本語から韓国語の単語を選択する過程では、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と考えられる。この際、日韓の単語の意味は必ずしも1対1に対応しない。そこで単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1のようなネットワークで表すことにする。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、日本語の単語からの韓国語単語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1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1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-正解文の単語のパス[a1-d1]と入力-問題文の単語のパス[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。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: ~することができる等)を含む助動詞句の述部で多く見られる。そこでこのシステムでは日本語の助詞と連語を含む助動詞句についてこの方法を用いる。

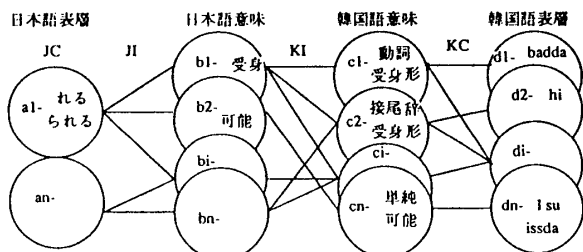


図1 日韓意味対応関係

4. アドバイス生成

学生に対するアドバイスでは誤り原因と韓国語の基礎的文法を詳しく教示する必要がある。従って本研究では正解だけではなく、誤りの原因、ヒント、詳細な文法知識、例文などを提示する。アドバイスは誤文解析の結果を基にアドバイス生成知識を用いてテンプレートに埋め込むことにより生成する。

5. システム構成と処理例

システムの構成図を図2に示す。システムは日本語の問題集から問題を出題し、学生の韓国語ローマ字入力文を3.2の誤り検出の手順に従って誤りを解析する。また誤文解析の結果を用いて学生に文法的アドバイスをを行う。その処理例を図3に示す。

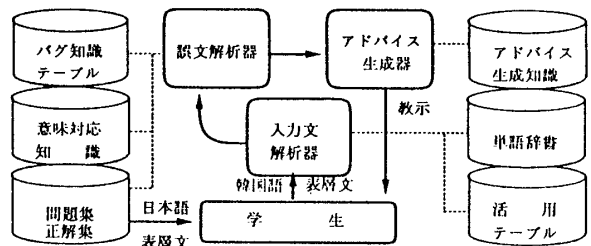


図2 韓国語作文演習知的C A Iシステム構成図

- >>教師: [問題1]=努力をする人はほめられる
- >>学生: "noryeogreul hada saramneun chingchanhal su issda"
- >>教師: あなたの作文には「3」つの誤りがあります。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] - あなたの作文では((1) [名詞, noryeog]に接続された[助詞, reul] (2) [名詞, saram]に接続された[助詞, neun])の助詞音韻を誤っています。
- [B] - [ヒント] - 助詞を適用する時は先行語の音韻を考えましょう。
- [C] - [文法知識説明] = [助詞音韻適用]について説明します。
先行語の最後の字が子音の場合は助詞の頭文字が母音、母音の場合は子音です。
- (中略) - 例えば[人は-saramteun, 私は-natneun]のようになります。
- [2] = [連体形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[する人]のところで韓国語の[終止形動詞, hada]を用いて[名詞, saram]を修飾しています。この場合の韓国語の表現は連体形活用語尾を用います。
- [B] - [ヒント] - 日韓の連体形構造について考えましょう。
- [C] - [文法知識説明] = [連体形用法]について説明します。
韓国語は連体形動詞、形容詞から連体形活用語尾を用いて体言を修飾します。
連体形活用語尾は[n, neun, 1, deon]があります。 --- (以下省略) ---
- [3] = [問題文と異なる意味の作文をした誤り]
- [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[1 su issda]と作文しています。
- [B] - [ヒント] - 日本語に対する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, 尊敬, 可能, 自発]などの複数の意味を持ちます。
また問題文の意味によって韓国語の表現も異なります。 -- (以下省略) --

図3 処理例

6. まとめ

学生の誤文を解析して教示を行う手法について論じ、実際に試作システムを構築した。また今後はより複雑な文体について取り扱うために構文解析を導入する等、誤り解析について検討を進める。

【参考文献】

- [1] 朝鮮語を学ぼう: 菅野裕臣監修, 日本朝鮮語学研究会編著, 三修社(87)