

外国語学習に係る速訳用課題提示システムについて

— 状況的課題選択 —

北 垣 郁 雄

(2008年10月2日受理)

On a Computerized Foreign Language Learning System
Displaying the Material for Prompt Translation
— Situational selection of a material —

Ikuo Kitagaki

Abstract: I report a situational prompt translation practice system for language learning. The system first builds a database holding many of short sentences. After that, it selects a sentence from the sentences through a quasi-random process and then displays it on the screen. The learner promptly and verbally translates it into the foreign language. The system employs such a method that the learner verifies how precisely the sentence has been translated by comparing the learner's translation with the display answer and the voice answer indicated by the system. The system aims at the learner who finds difficulty in promptly speaking what he/she wants to say in the foreign language. Because there is not necessarily one correct answer, the learner must have the ability to judge how precise the his/her answer is by visually and verbally checking the answer indicated by the system, which would be one of the correct answers. The adjective "situational" derives from the fact that the system has adopted such an algorithm that the occurrence probability of a short sentence which is suitable under a certain situation will be increased and thus the occurrence probability of an unsuitable sentence will be decreased, by calculating the compatibility between the property of the short sentence and the property of the user, even though the short sentence displayed on the screen has basically been selected through a quasi-random process. In the calculation above, fuzzy theory is implemented when comparing two properties mentioned above.

Key words: CALL, prompt translation practice, situational selection of a material

キーワード：コンピュータ支援語学学習，速訳練習，状況的課題選択

1. はじめに

情報通信技術の発達と共に，その教育への応用に関心が高まっている。コンピュータの支援による語学学習（CALL: Computer Assisted Language Learning）もその一つである。

本研究では，短文を当該外国語に速訳するという一つの外国語学習法に関し，これを支援するための速訳用課題提示システムを報告する。そこでは，数多くの

日本語(元言語)による難度・音声・画像付き短文データベースの中から，準ランダムに短文を課題として画面・音声提示し，その当該言語（速訳言語）への速訳を学習者に促す。回答ボタンを押すと，速訳言語による一つの回答が文字と音声で提示される。正解は一つとは限らないが，自分なりの回答をその正解に照合することができる。新たな短文提示を求めると，準ランダムにそれが決定され提示される。この手法を順次続けていく。したがって，学習者には，提示された正解

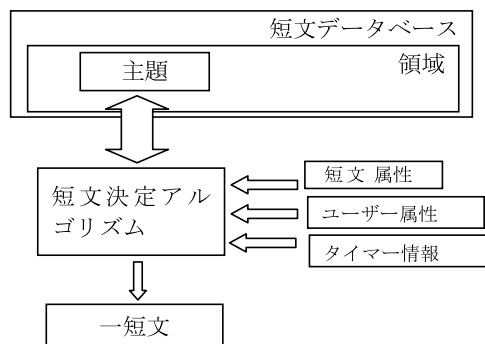


図1 システム構成

に対して自分の回答を評価できるなど、ある程度の学力的条件が必要となる。換言すれば、正解を見聞きすれば当該短文に対する的確な一訳文であることを理解できるものの、条件反射的に自らそれを文章構成するには難がある、というレベルの学力者を対象とする。もしも所与の短文がそのレベルに該当しないのであれば、難度に関しての不適切な課題と思われるから、提示要求する際の難度を変更すればよい。したがって、本システムは、当該外国語学習に関する、初心者以外の学習者を対象とするものと言って良い。

本システムでは、全短文に共通する属性を持たせ、教材開発者が属性ごとにその短文が適合する割合を数値で割り当てるようにしている。そして、ユーザーにも同じ属性を割り当て、各属性に数値を設定させる。そして、両者の対応を図ることにより、当該ユーザーに整合する度合いが高い短文の選択確率が高くなるようなアルゴリズムを作成している。

次に、先のユーザー属性と短文属性の設定では、いずれもファジィ論的扱いを可能にしている。具体的には、各属性をいくつかの区分に分け、区分ごとに適合度を割り当てることにより、メンバーシップ関数を設定するようにしている。

また、本システムは、語学学習に資するオーサリングシステムとしても活用を考慮することができ、応用範囲が広い。さらに、モバイル PC 等携帯端末を通して使用するといったシステムの拡張性をも有する。

筆者は、以上のような特徴をもつ CALL を on-going プロジェクトとして開発している。本報では、中間報告として、設計の骨格をまとめている。

2. システムの概要

本システムの中核部分である短文課題の決定方法は、図1に示すとおりである。

短文データベースは、全体>領域>主題の順に下位に構造化されている。そして、所与の主題に属する短文の中から、1つの短文が所定のアルゴリズムにしたがって決定されるようになっている。数多くの短文を想定しているため、互いに異なる領域・主題の短文を継続提示するよりも、一つの主題の範疇である方が、学習者の興味を継続させると思われるからである。

短文決定アルゴリズムでは、当該の主題の短文群の中から、準ランダムに一つの短文を選択する。そこでは、選択の対象となるすべての短文の選択が等確率ではなく、短文やユーザーの属性等を加味しつつ、ランダムに選択する。より具体的には、一つの属性に関し、「短文」と「ユーザー」を照合し、その整合度が高いほど選択確率が高くなるよう設定する。したがって、短文決定アルゴリズムでは、乱数を使用しているものの、一つの短文群の中で、短文の選択されやすさには、ある程度のむらがある。「準」ランダムとは、そのような意味である。

属性は、非時間属性と時間属性の2種類がある。

非時間属性には、「地域」、「性別」、「年代」、「職業」などがある。さらに、各属性に対していくつかの区分を設け、区分ごとに[0,1]の数値を与えることができる。例えば、「性別」という属性は、{男, 女}といふ2つの区分がある。そして、一人のユーザーが「男」「女」のいずれに対しても0.0～1.0の数値を与えることができる。もし、「男」に1.0、「女」に0.6を割り当てたとすると、男性固有の内容の短文が選択される確率に対して、女性固有の内容の短文が選択される確率が、6割程度と考えればよい。そのようなファジィ論的処理方法については、次節で詳述する。

短文には「難度」を[0,1]の数値で与えている。しかし、操作上の煩雑さを避けるため、区分別の処理は行っていない。「難度」も非時間属性の一種と言えるが、先の「地域」等とは別の処理となっているので、便宜上、非時間属性のカテゴリーから外すことにしたい。

時間属性は、「日」と「年」がある。「日」は、1日を0時から24時まで1時間毎に24の区分に分けている。また、「年」は、1月を上・中・下旬に分け、全部で36の区分に分けている。そして、ユーザーのアクセスタイムと時間属性（および区分）を照合し、その整合度が高いほど当該短文の選択確率が高くなるようになっている。したがって、「朝食」の短文と「夕食」の短文が存在すると仮定して、当該ユーザーが朝アクセスすれば、前者の短文が出やすく後者が出にくくなるように属性設定することができる。一方、夕方アクセスすると、その逆になるように設定できる。「桜」の短文と「紅葉」の短文が存在するときも、ほぼ同様

の考え方で適切な属性設定が可能になる。

このように、本システムは、ユーザーの属性やアクセスタイムが短文課題の選択に反映されている。いわば「ユーザーが置かれた状況」を考慮したシステムであり、状況的課題選択と特徴づけることができる。

3. 短文・ユーザー属性のファジィ論的情報処理

第*i*短文 ($i=1, \dots, I$) において、第*j*非時間属性 ($j=1, \dots, J$) の第*k*区分 ($k=1, \dots, K$) の値を $s(i, j, k)$ と表記する。また、あるユーザーに関し、第*j*非時間属性の第*k*区分の値を $u(j, k)$ と表記する。

$$\begin{aligned} 0 \leq s(i, j, k) \leq 1 \\ 0 \leq u(j, k) \leq 1 \end{aligned} \quad (1)$$

同じように、あるユーザーで第*j*時間属性の第*k*区分の値を $u_i(j, k)$ と表記する。このとき、ユーザーのアクセスタイムが第*k*₀区分に該当するときは、次式のようなになる。

$$\begin{aligned} u_i(j, k) = 1 \text{ for } k = k_0 \\ u_i(j, k) = 0 \text{ for } k \neq k_0 \end{aligned} \quad (2)$$

なお、第*i*短文で第*j*時間属性 ($j=1, 2$) の第*k*区分の値を $s_i(i, j, k)$ と表記しておく。

$$0 \leq s_i(i, j, k) \leq 1 \quad (3)$$

難度は、第*i*短文の元言語と速記言語の双方に設定できるようにしている。それぞれの値を $d_u(i)$, $d_s(i)$ と表記する。一方、ユーザーの設定した難度を d_u で表す。

このとき、第*i*短文に対する状況的整合度 μ_i を次式で定義する。

α_i : 非時間的属性の整合度

β_i : 時間的属性の整合度

γ_i : 難度的整合度

$$\mu \equiv \alpha \wedge \beta \wedge \gamma \quad (4)$$

$$\alpha = \bigwedge_{j=1}^J \left(\bigvee_{k=1}^K (s(i, j, k) \wedge u(j, k)) \right) \quad (5)$$

$$\beta = \bigwedge_{j=1}^J \left(\bigvee_{k=1}^K (s_i(i, j, k) \wedge u_i(j, k)) \right) \quad (6)$$

$$\gamma = 1 - |(d_u(i) \wedge d_s(i)) - d_u| \quad (7)$$

その後、(4)式を用いて、(8)式の規格化状況的整合度を求める。そして、(8)式より(9)式の確率分布関数Fを求める。

$$\mu_i^0 = \frac{\mu_i}{\sum_{i=1}^I \mu_i} \quad (8)$$

$$F(\mu_i^0 | i \leq i_0) = \sum_{j=1}^{i_0} \mu_j^0 \quad (9)$$

ここに、

$$F(\mu_i^0 | i \leq I) = 1 \quad (10)$$

そして、乱数の値 r_0 ($r_0 \in [0, 1]$) に対し、

$$F(\mu_i^0 | i \leq i_0 - 1) < r_0 \leq F(\mu_i^0 | i \leq i_0) \quad (11)$$

となるような第 i_0 短文を選択する。

以上の方法により、ユーザーと短文の属性的整合性を加味しつつ、準ランダムな短文選択が可能になる。

4. 課題提示システムの全体的構成

当システムの全体的流れ、データベースの構造などを説明する。

4.1 短文データベース

一つの言語による短文は、表1の構造を持っている。したがって、日英のシステムであればこの構造が2つとなり、さらに日中を加えるのであれば3つとなる。

2節で、本システムではユーザーの置かれた状況を短文選択方法に反映させる旨を述べたが、短文提示における背後的な「状況」の説明も重視している。例えば、提示課題として「その赤い荷物を取り上げていただけますか?」と提示するよりも、「空港で、荷物を受け取るベルトコンベヤーの脇でポーターに向かって『その赤い荷物を取り上げていただけますか?』」と表示したほうが、如何なる状況での会話であるかがイメージでき、それを想定した回答に絞られるはずである。そこで、本システムにおける教材開発では、状況説明文と速記を求める課題文の双方を含めた文、および課題文だけの2種類を用意している。前者と後者が、

表1 一言語による短文の情報

データの種類	データの属性
状況説明付きの短文	文字列
状況説明付きの短文音声	パス情報
画像情報	パス情報
状況説明なしの短文	文字列
状況説明なしの短文音声	パス情報
発音記号	文字列
ヒント	文字列
難度	数値[0, 1]

表2 短文属性

(a) 非時間属性

地域

短文キー	日本	アメリカ	中国	...
0001	1.0	1.0	1.0	...
0002	1.0	0.0	0.0	...

年代

短文キー	...	20代	30代	...
0001	...	1.0	1.0	...
0002	...	1.0	0.7	...

性別

短文キー	男	女
0001	1.0	0.0
0002	1.0	0.6

職業

短文キー	...	大学生	勤労者	...
0001	...	1.0	0.8	...
0002	...	0.4	1.0	...

表現

短文キー	基礎・基本...	慣用的表現	...
0001	0.8	1.0	...
0002	1.0	0.4	...

(b) 時間属性

日

短文キー	...	7時	8時	...	18時	19時	...
0001	...	1.0	1.0	...	0.0	0.0	...
0002	...	0.0	0.0	...	1.0	1.0	...

年

短文キー	...	3月下旬	4月上旬	4月中旬	...
0001	...	1.0	1.0	0.8	...
0002	...	0.0	0.0	0.0	...

それぞれ表1の「状況説明付きの短文」、状況説明なしの短文」に対応する。またそれぞれの音声も用意している。

例えば、日本語から英語への速訳の場合は、日本語による「状況説明付きの短文（音声）」と英語による「状況説明なしの短文（音声）」を提示する。一方、英語から日本語への速訳の場合は、英語による「状況説明付きの短文（音声）」と日本語による「状況説明なしの短文（音声）」を提示することになる。

速訳において「画像情報」を使用することもあるのですが、データベースには各言語による画像情報を登録することができる。

データベースには、「発音記号」を登録することもできる。特に、中国語ではピンインを表示するためにこのデータを用いる。

「ヒント」は、構文や難しい単語などを表示するために利用する。

「難度」は、言語ごとに設けている。

4.2 全体の構成と流れ

本システムは、大きく4つの部分に分かれている。

1. 短文管理機能— 管理者は短文データベースを閲覧したり、「領域」や「主題」を設定することができる。また、一つの「主題」に対して、特定の短文を登録・削除・修正することができる。短文の登録では、表1に示した教材を準備しておく必要がある。一つの短文は、いくつかの言語にまたがって共通する内容であるので、現実には、それらはほぼ同時期に登録されるはずである。

短文の登録では、併せて属性も登録しなければならない。各属性には任意の個数の区分を設けることができる。本研究で設定した属性と区分の例を表2に示す。どの属性も、2つの短文事例で示している。表(a)の非時間属性のうち、「地域」に対しては、特定の地域を想定しない場合は、短文キー0001の設定のようにすべてに1.0を与える。一方、日本以外の地域を想定しないときは、短文キー0002のように設定すればよい。

同じような考え方で、「年代」で、年代に依存しない場合は、短文キー0001のようになり、20代が最も相応しい内容の場合は、短文キー0002に準じた設定となる。

「性別」は、ふつうは短文キー0001のように、一方が1.0でもう一方が0.0となるように設定するが、短文キー0002のように設定することもできる。この場合は、男性向けの内容の短文が現れる割合に対して女性向けの割合が6割程度であると解釈すればよい。したがって、男女不問の短文に対しては、男・女いずれの区分に対しても、1.0と設定することになる。

「職業」、「表現」についても、ほぼ同様である。

「表現」は、日本語の文法（庵功雄，2001）をまとめ直し、つぎの5つに分けた。

- ・基礎・基本、待遇表現
- ・慣用的表現
- ・数量、比較、図表説明
- ・不確実、仮定、疑問、否定
- ・複文、重文、装飾

次に、表(b)の時間属性の「日」に関し、短文キー

0001と短文キー 0002は、それぞれ朝食、夕食を話題にした短文属性の設定例である。

同じく「年」に関し、短文キー 0001と短文キー 0002は、それぞれ桜、紅葉を話題にした短文属性の設定例である。

すべての属性に関することであるが、ユーザーは、自分に該当しない区分に対しては、0.0を設定することができる。しかし、そのような区分が多いユーザーに対しては、ときには、実際に選定される短文の範囲が狭くなって、同一の短文がしばしば現れるという不都合が起り得る。そのため、本システムの実運用では、0.0という設定に対しては、0.1～0.2程度の割合を与えた上で前節の計算を行うようにしている。

2. 短文提示機能— 本システムの中核部分である。短文の提示過程には、3種類がある。

第一は、準ランダム系列による提示過程である。これは、選択の対象となる短文の中から、前節で述べた短文決定アルゴリズムにしたがって一つの短文を選択し、画面表示する方法である。そこでは、状況説明文と課題文を表示する。メディア制御に関するコマンドボタンとして、「状況説明付き短文音声」、「文字回答」、「音声回答」が用意されている。どの短文画面でも、「難度」の設定変更が可能であり、それを変更すると、(7)式が変更されるので、(11)式が再計算される。

第二は、再現系列による提示過程である。準ランダム系列によって選択された一連の短文は、その提示順序を一つのレコードとして登録するようにしている。学習の途中に他の「主題」や「領域」に移動して準ランダム系列にて画面提示すると、移動後の系列は新たなレコードとして登録するようになっていく。一方、ユーザーがすでに登録されたレコードの中から特定のレコードを指定すると、その系列を再現することができる。したがって、準ランダム系列にて学習した短文に関し、それと同じ系列にて復習することができる。

第三は、固定系列による提示過程である。これを用いると、短文データベースに登録された短文を順次表示することができる。難度など特定のキーに該当する短文を表示させるのに役立つ。

再現系列と固定系列の実施では、メディア制御の方法は準ランダム系列のときと同様である。その際、難度の設定変更は、無効となる。

3. ユーザー情報管理機能— ユーザー ID の登録、抹消、Login/Logout 管理、ユーザー属性の管理を行う。

4. 実行情報管理機能— ユーザーの選んだ速訳言語、難度、準ランダム系列における短文順序の記録や管理を行う。

5. 本システムの活用方法と特徴

本システム構成に関する特徴は、つぎのようにまとめることができる。

第一に、短文を登録対象としたデータベースを用いている。コンピュータを教材提示・回答分析手段として用いる語学教育では、Vocabulary Learning (Ma, 2006; 澤田, 1996) や単語の発音比較 (牧田, 1996) など、単語レベルでの使用を対象とすることが多い。これに対し、本システムでは短文を対象にしている。

第二に、短文を多数掲載した印刷教材では、その提示順序は当然ながらいつも同じである。しかし、本システムでは、短文の提示系列に乱数を利用しているので、ある提示短文の次に提示される短文を予測することが困難である。ここに、ゲーム感覚が生じることが予想される。

第三に、自己学習を可能にし、速訳練習を繰り返し行うことができる。自己学習と繰り返し学習の重要性は、語学教育研究者がしばしば協調するところである (伊藤, 2003; 浅沼, 1996)。

第四に、先の乱数を活用した短文選択では、当該ユーザーが想定する「状況」を反映させている。これにより、ユーザーの置かれた立場を考慮した短文選択を可能にしている。

第五に、前記「状況」の設定では、短文とユーザーにいくつかの属性をもたせ、さらに各属性に区分を設けることにより、ファジィ論的処理を行っている。つまり、各属性をメンバーシップ関数にて特徴付けており、ユーザーの設定した諸状況に対する整合性の高さを短文の選択確率に反映させている。これに関連して、子供用の辞書における単語検索で、ファジィ推論を応用した例がある (St-Jacques et al., 2005)。

第六に、どの短文提示画面でも、その提示中に難度の属性を変更することができる。「難度」は、上記と同様、ユーザーとの整合性が高いと短文の選択確率が高くなる傾向がある。その際、ユーザーは難度が大幅に異なる短文に対しては、回答ボタンを押すことなくパスすることができる。その場合、パスした短文は再現系列としては記録されないようになっていく。したがって、再現系列には、難度等からみてそのユーザーにより整合する短文だけが残されていることになる。

次に、本システムでは、短文データベースに登録した短文を用いて速訳用短文提示を行うことができるが、そのほかに、オーサリングシステムという形式にて活用することができる。つまり、語学教育者が各自必要とする短文等を開発し、短文データベースに登録する。また、非時間属性や区分も任意に設定して使用

するという方法である。いわば、カスタマイズ方式である。これにより、本システムの活用範囲が広まることが予想される。

また本システムは、モバイル PC や携帯電話を情報端末の対象とするという拡張性を持っている。特に、海外では、ローミング機能により、その場に整合する主題やそこでの状況に即した属性や主題による速訳練習が可能になる。

6. おわりに

本報では、語学学習のための速訳用短文提示システムについて述べた。そこでは、ユーザーの置かれた状況を加味した短文選択を行うために、短文データベースから短文選択する際に、ファジィ論的アルゴリズムを用いていることを述べた。今後は、実際に教材開発を行い、実践を通してシステムの効果を検証したい。

【参考文献】

浅沼大海 (1996) : Hyper Card ドイツ語教材の活用, コンピュータ&エデュケーション, Vol.1, pp.80-84.

庵功雄 (2001) : 新しい日本語学入門 ; ことばのしくみを考える, スリーエーネットワーク, 東京.

伊藤健一 (1996) : 外国語コミュニケーション教育における学習・教育観—使える英語力の獲得を目指して—, コンピュータ&エデュケーション, Vol.15, pp.19-23.

Ma Qing and Kelly Peter (2006): Computer Assisted Vocabulary Learning: Design and evaluation, Computer Assisted Language learning, Vol.19, No.1, pp.15-45.

牧田英二, 楊立明, 楊達, 平林宣和, 遠藤雅裕, 舩谷鋭 (1996) : 中国語学習支援ソフト開発への試み, コンピュータ&エデュケーション, Vol.1, pp.100-104.

澤田肇 (1996) : 自主制作ソフトウェアの教育への利用, コンピュータ&エデュケーション, Vol.1, pp.95-99.

St-Jacques C. and Barriere C. (2005): Search by Fuzzy Inference in a Children's Dictionary, Computer Assisted Language Learning, Vol.18, No.3, pp.193-215.