

国際共同編集による日本語学習者のための多言語版 web 辞書の開発

川村 よし子(東京国際大学)・金庭 久美子(横浜国立大学)

< 共同研究者 > 植木正裕・保原麗・川村ヒサオ・根津誠

Hans Coppen, Joyce Maeda, Jonathan Bunt,

Kristina.S.Hmeljak, Saleh Adel

1. 研究の目的

本研究の目的は、世界各国の日本語教育者・研究者との共同作業によって、日本語学習者が必要としている各言語版の対訳日本語辞書を作りあげ、多言語版 web 辞書として公開することにある。従来の辞書は各言語が別個に対訳日本語辞書を作成していたため、各国の日本語学習者が得る情報は言語によって異なっていた。本研究の多言語版 web 辞書は、意味情報等がすべて日本語で説明された日本語辞書(以下、「日日辞書」と記す)を基にして、それに各言語の対訳をつけていくという方法で編集を進めている。そのため、日本語学習に必要な辞書情報を世界各国の学習者に共通して与えることが可能になる。また、他の言語による対訳情報についても容易に参照できるというメリットもある。

2. 先行研究

本研究の基盤となっているのは日本語読解学習支援システム「リーディング・チュウ太」(<http://language.tiu.ac.jp>)の辞書ツールである。「リーディング・チュウ太」は寺ほか(1996)による読解支援システム「DL」を改良し、学習履歴管理機能(北村ほか 1999)を加えるとともに、辞書の意味情報を充実させるため日本電子化辞書研究所の EDR を用いる(川村ほか 2000)などして発展してきたものである。1999 年に公開以来、学習者が自分の読みたいものを自由に教材化して学習できる開かれたシステムとして、インターネットを活用した読解支援の分野で先駆的な役割を果たしてきた。

web 上で使える辞書の開発としては、モナシュ大学の J.Breen による edict(Breen 1993)の果たした役割は大きい。前述の「DL」も辞書情報として edict を利用している。また edict の辞書情報は、すべてダウンロード可能な形で提供されているので、これをもとにフランス語、スペイン語等の日本語辞書が作成されている。さらに、この edict を発展させた www.jdic(Breen 2000)は、検索エンジンとも連動させることにより入力された単語の辞書情報のみならず、サイト検索、画像検索も可能な仕組みとなっている。一方、英語以外の対訳辞書としては、大阪大学の Ulrich Apel(2001)による『和独辞典』があり、この辞書情報をもとに「リーディング・チュウ太」の辞書ツールにドイツ語版

を取り入れた(川村 2002)。

既存の市販辞書の web 上での公開も始まり、『大辞林』『EXCEED』『ニューセンチュリー』(以上、三省堂)や『大辞泉』『プログレッシブ』(以上、小学館)等の国語辞典、英和辞典、和英辞典が利用可能になっている。また、『英辞郎 on the Web』(アルク)は単語のみならず、句、文レベルの対訳情報が含まれ、英和辞典としても和英辞典としても使える形になっているため、読解ばかりでなく作文支援としても有用である。だが、いずれも日本語母語話者を対象に作られたものであり、初級・中級の日本語学習者にとっては利用が難しい。

こうした中であって、日日辞書、日英辞書、日独辞書を搭載している「リーディング・チュウ太」の利用者も多く、さらに、各国の学習者からは、学習者の母語で書かれた対訳辞書がほしいという切実な要望が依然として強い(金庭ほか 2005)。英語以外の言語で書かれた対訳日本語辞書は少なく、あっても例文などのない単語集レベルのごく簡略なものであったり、学習者が入手しにくい状況であったりする場合が多い。こうした状況を受けて、2003 年に「辞書ツール多言語化プロジェクト」を立ち上げ、多言語版 web 辞書を開発することにした。

本研究の多言語版 web 辞書は、日日辞書の情報を基にして、それに各言語の対訳をつけていくという方法で編集を行う。そのため、第一段階として、2003 年度はインターネット上で日日辞書の編集を自由に行える編集システムの開発を行った(川村ほか 2003)。2004 年度には第二段階として、基になる日本語辞書の語の選定を行い、概念の提示順序を決め(金庭ほか 2006)、完成した日日辞書編集システムを利用して、辞書の編集を開始した。この編集作業では語の意味を平易な日本語で書き表すとともに、例文や用法等を入力した。第三段階として 2005 年度には、多言語版日本語編集システム 版を開発し、これを用いて編集済みの日本語情報に対訳を入力する各言語版の編集がスタートした。

3. 多言語版日本語辞書編集システム

多言語版日本語辞書編集システムは、日日辞書をもとに世界各国の編集者が各言語版の対訳辞書を編集するシステムである(川村ほか 2005)。日日辞書の情報が対訳されると、その情報は自動的に SQL サーバーに送られ、他の言語の辞書情報とリンクされる。この編集システムでは、入力された辞書情報を自動的に XML 化し多言語情報を一元的に管理する形をとっている。そのため、このシステムを用いれば、世界のどこからでもインターネットを介して日本語辞書編集作業に容易に参加することが可能である。

図 1 は多言語版辞書編集システムのフランス語編集画面である。編集者は画面上のボックスに対訳情報を入力する。入力項目は、Sense(意味概念の説明)、Same Word(概念説明に該当する代表的な訳語)、Translation(概念説明に該当する Same Word 以外の訳語)、

Example(例文)、Note(用法等の説明)である。



図1 多言語版辞書編集システムの編集画面

これらの情報が入力されると、直ちに辞書として使用が可能になり、他の言語の情報が入力されていれば、比較が可能になる。

この多言語版辞書編集システムを用いた運用実験は、2005年8月から9月にかけて、英語、オランダ語、スロヴェニア語版の対訳編集者の協力を得て行った。その結果、以下の問題点が明らかになった。1) ネット環境の違いによる編集画面の見え方の相違、2) 文法項目に関する説明の必要性、3) 意味概念ごとの例文情報の必要性、4) 個々の言語に固有な意味情報、文法情報、文化情報に関する補足説明の必要性、5) 編集者相互の情報交換システムの確立である。

そこで、運用実験の結果判明した個々の問題に対して対応方法を検討するとともに、編集作業において発生する可能性のある諸問題を未然に防ぐための対応策を検討した。1) については技術的な側面に対応すべき部分であり、システムの改善作業を行った。2) については、接続する活用形や意味の使い分け等を入れることにした。3) については、意味が区別できる例文を概念毎に作成することにした。また、その例文は学習者が利用可能な例文にし、書き言葉だけでなく会話表現も盛り込んだ。さらに4) については、日本語にあって当該言語にない意味、文化等の情報を解説するためのスペースを確保した。

こうして、文法項目や例文情報等の新たな入力項目を加えるとともに、言語ごとに管理可能な入力の際に留意すべき点の洗い出し作業を行った。特に例文については、可能な限りやさしい日本語(能力試験2級程度)にする、各意味概念が推測しやすいものにする、日本語として自然なものにする、等を基本方針としたが、翻訳の際の便宜を考え、主述を明確にした短文にする、終止形の場合には句として訳されることを前提に作成する、等の配慮も行うことにした。

一方、こうした一連の対応の中で、5)の各言語の編集者との情報交換が不可欠であることが判明した。そこで、個々の単語に対するコメント等をすべての言語の編集者が編集画面上で共有できるようにした。問題があれば作業画面上の「伝言板」に書き込み、日本語編集者に伝えるしくみである。またそれによって日本語辞書内容が変更されれば、当該の言語担当者だけでなく他の言語の担当者にも連絡される。さらに、情報交換が必要な場合には、MSN Messenger、Skype等のインターネット電話を活用することにした。さらに、2005年10月よりトルコ語、ブルガリア語、フランス語版の編集を開始した。

4. 日本語学習者のための多言語版 web 辞書

多言語版日本語辞書編集システムを利用して編集された各言語版の日本語辞書を、そのままインターネット上で公開したものが日本語学習者のための多言語版 web 辞書である(<http://marmot.chuta.jp/>)。

この辞書は、1)日日辞書、多言語対訳辞書として利用可能、2)インターネット上で利用可能、3)日本語能力試験出題基準に準拠した約8600語の見出し語、4)すべての意味に学習者向けの例文付き、5)用法説明あり、6)学習者に対するフィードバック機能あり、7)語彙の追加編集が随時可能、8)辞書内容の情報の追加、修正、削除が随時可能、という特徴を持っている。特に7)と8)は従来の辞書にない大きな特徴で、「進化する日本語学習者のための web 辞書」「時代やニーズにあった最新情報の辞書」と言える。

各言語版の編集済みの単語はまだ多いとはいえないが、各単語の編集作業が済むと自動的に辞書が更新されるシステムであり、一日も早く母語による対訳辞書を利用したいと考えている学習者にとって最新版の辞書を提供できるという利点がある。また辞書編集者も自らの編集作業の成果をその場で確かめることが可能である。現在、辞書としては、日日辞書に加えて、英語、トルコ語、ブルガリア語の辞書が閲覧可能である。



図2 多言語版 web 辞書検索画面

図2はインターネット上で公開されている多言語版 web 辞書の検索画面である。入力ボックスに検索したい見出し語を入力し、中央のボックスから言語を選択し、「Search」ボタンを押すと選択した言語版の辞書情報が閲覧可能になる。また、見出し語をひらがなで入力すると「会う」「逢う」「合う」「遭う」のような同音異義語がすべて表示される。

図3は「会う」のトルコ語版辞書の画面である。見出し語右肩の星印は日本語能力試験の出題基準に準拠した級を示している。意味概念ごとに、概念説明、訳語、例文とその対訳が示されている。さらに語の用法等の説明を要する場合には、Noteの形で注がつけられている。



図3 トルコ語版辞書の画面

画面下のように単語ごとに辞書利用者からの意見・要望に対応するためのコメント欄が設けてある。書き込まれたコメントは、単語および閲覧言語情報が自動的に付加された形で当該言語担当者に届き、適宜対応できる仕組みである。

この多言語版 web 辞書については、現在、日日辞書、対訳辞書（トルコ語・英語）の編集チームで運用実験を行っている。その結果、概念の説明方法、例文の内容、辞書のレイアウトなどに関しては、基本的には問題がなかったが、各言語版と英語版を同時併記した提示方法、例文検索、例文への振り仮名付加、関連語句へのリンク、用法を対訳する際の用語の統一等の要望があった。

今後も、運用実験を継続するとともに、利用者からのフィードバックも参考にしながら、学習者にとってより使いやすい辞書をめざして改良を進めていく予定である。

5. 今後の課題

多言語版日本語辞書編集システムによる成果物は、将来的には「リーディング・チュウ太」の辞書ツールに組み込むことを目指している。各言語版辞書の編集作業を継続するとともに、欧州やアジア各地の日本語教育関係者に呼びかけ編集協力者を募っていく予定である。現在、ハンガリー語版の編集も開始され、アラビア語、フィンランド語、

中国語、韓国語版の計画も進行中である。

本研究の辞書編集システムは母語話者と非母語話者の双方の視点からの意見を取り入れた国際共同編集を可能にしている。本研究においても、例文の作成方法等について、相互に情報交換することによって、日本語学習者が必要としている辞書情報が明確になった。これは今後の言語教育用教材作成のあり方にも示唆を与えるものと言えよう。

辞書編集システムにはまだ改良すべき点も多い。操作手順の簡便化、修正箇所の明示等、個々の問題に対応しながら、より編集しやすいシステムに改良を加えていくと同時に、コメント等の情報管理システムを完成させる必要がある。一方、このシステムは、多言語情報を一元的に管理することができるシステムであり、単に辞書の編集にとどまらず、新しい形の教材共同開発システムとして活用可能なものである。教材開発等をすすめている諸機関との連携を視野に入れた活動を進めていきたい。

謝辞 本研究は電気通信普及財団の平成 17 年度・平成 18 年度研究助成および東京国際大学の平成 17 年度特別研究助成による支援を受けた。ここに記して感謝の意を表したい。

参考文献

- Apel, Ulrich(2001) WaDokuJT, <http://133.1.179.17:591/WadokuJT/default.htm>
- Breen,J.W.(1993) A Japanese Electronic Dictionary Project (Part 1: The Dictionary Files), Technical Report, Monash University.
- Breen, J. W.(2000) A WWW Japanese Dictionary, *Japanese Studies*, 20-3, 313-317.
- 金庭久美子, 川村よし子, 前田ジョイス(2005)「日本語学習者のための電子辞書編纂の基礎調査」第 10 回ヨーロッパ日本語教育シンポジウム.
- 金庭久美子, 川村よし子(2006)「日本語学習者のための電子辞書編纂 語の選定と意味の提示順序」『日本語教育方法研究会誌』13(1), 26-27
- 川村よし子, 北村達也, 保原麗(2000)「EDR 電子化辞書を活用した日本語教育用辞書ツールの開発」『日本語教育工学会論文誌』24 (Suppl.), 7-12.
- 川村よし子(2002)「日独辞書ツールの開発とその評価」『Papers for the 14th International Conference on Japanese Language Teaching』59-63.
- 川村よし子ほか(2003)「辞書ツール多言語化プロジェクトの基本構想」第 8 回ヨーロッパ日本語教育シンポジウム
- 川村よし子, 金庭久美子ほか(2005)「多言語版日本語辞書編集システムの開発と運用実験」第 10 回ヨーロッパ日本語教育シンポジウム.
- 北村達也, 川村よし子ほか(1999)「学習履歴管理機能を持つ日本語読解支援システムの開発とその評価」『日本教育工学会論文誌』23(3), 127-133.
- 寺朱美, 北村達也ほか(1996)「WWW ブラウザを利用した日本語読解支援システム」『日本語教育方法研究会誌』3(1), 10-11.