

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

A Web Dictionary Compilation for Japanese Language Learners

Word selection and order of word definitions

金庭久美子* KANENIWA, Kumiko, 川村よし子** KAWAMURA, Yoshiko
横浜国立大学* 東京国際大学**
〒240-8501 神奈川県横浜市保土ヶ谷区常盤台 79-1
Yokohama National University : byh04643@nifty.com

Abstract: This study reports our ongoing international project on developing a web-based multilingual dictionary for Japanese language learners (<http://marmot.chuta.jp/>). After reporting the history of basic word selection, we discuss how word entries should be selected and in what order word definitions should be presented. Then we discuss the editing system and the editorial standard for our web-based dictionary based on the above discussion.

キーワード：基本語、語の出現頻度、単語親密度、意味の提示順序

1. はじめに

現在、国際共同編集により日本語学習のための多言語 web 辞書の開発を進めている。

辞書の編集においてまず考えなくてはならないのは、語彙数と語の種類、つまり、見出し語として何をどれだけ取り上げるかである。しかしながら、日本語学習者の視点から辞書を考えると、語彙数や語の種類もさることながら、語の意味の説明の仕方、意味がたくさんある場合の提示順序、用例などへの配慮が不可欠である。

本研究では基本語選定の変遷をたどり、辞書の見出し語選定のあり方を考えるとともに、日本語学習者に必要な意味の提示順序を考え、それに基づいた web 辞書の編集システムと編集基準を述べる。

2. 基本語選定の変遷

1) 必要な語を優先して選定した時代

日本語教育の分野における基本語に関する考え方は、まず日常生活の言語活動に必要な語彙かどうかということだった。『日本語基本語彙』(土居光知 1933)は 1000 語、それに続く「日本語教育における基礎学習語」(加藤彰彦 1963,64)は 6 種の資料に基づき 1393 語を選定した。さらに、専門家による判定方式で『日本語教育のための基本語彙調査』(国立国語研究所 1984)では「基本語二千」、「基本語六千」を選んでいる。辞書では、日本語学習書やその後の語彙調査などから『外国人のための基本語用例辞典』(文化庁 1971)では 4600 語、『基礎日本語学習辞典』(国際交流基金 1986)では 2873 語が選ばれた。

2) 文献中の出現頻度によって語を選定した時代

電子計算機の登場によって、数量的に語の出現率が計れるようになった。『現代雑誌九十種の用語用字』(国立国語研究所 1962)で得られたデータは、

その後の『分類語彙表』(国立国語研究所 1964)で活用されている。また、『高校教科書の語彙調査』(国立国語研究所 1983,84)や『中学校教科書の語彙調査』(国立国語研究所 1986,87)なども行われ、数量的な語の出現率が明らかになった。『日本語能力試験出題基準』(国際交流基金・日本国際教育協会 1994, 2002 年改定)において特に 1, 2 級に選定された語は、基本的には上記の資料等に基づいたものである。本研究の辞書に登録済みの語もこの『日本語能力試験出題基準』を基に選定した。

3) 現代：単語親密度によって語を選定する時代

しかしながら、文献中の出現頻度が低くても実生活ではよく目や耳にする語がある。近年、基本語の選定において新たな選定方法が可能になった。それは単語親密度に基づく選定である。『日本語の語彙特性(第 1 期)』(天野成昭他 2000)では、ある語に対して馴染みがあるかどうか、18 才以上の被験者 40 名が 7 段階評定を行った結果の平均値を求めている。単語親密度の高いものは基本的な語彙と考えられる。本研究の辞書においてもこの資料に基づいて追加登録を行う予定である。徳弘康代(2005)も『日本語の語彙特性』に着目し、使用頻度と親密度のデータから日本語学習者に必要な漢字語彙を選び、それを利用して日本語教材の開発を進めている。今後も、このような方法で、学習者に必要な教材が整備されていくと思われる。

3. 意味の提示方法

日本語学習のための辞書を考えてとき次に問題となるのが、辞書の内容、つまり、各々の語の辞書情報のありかたである。語の意味の説明の仕方、意味がたくさんある場合の提示順序、用例、文法説明等多岐にわたるが、中でも語の意味の提示順序は重要である。

伝統的な国語辞典の場合、基本的な意味が上位に、派生的なものが下位に並べられる。しかしながら、日本語学習者にとってはその意味が派生的な意味かどうかはあまり問題ではなく、探したい意味がすぐに見つかるかどうかの方が重要となる。

北村・川村(2000)ではある語にいくつかの意味が並ぶ場合、EDR コーパス(注1)の意味の頻度情報に基づいて、よく使われる順番に並べ替えるという試みを行った。それによって学習者に必要な語の意味が上位に提示されるようになった。しかしコーパスの大きさが十分でない、コーパスで使用されていない意味の頻度は不明、近い意味のものが近くに提示されない等の問題があった。

こうした点を踏まえ、本研究の多言語版 web 辞書では意味の提示順序を再検討した。

4. 本研究における意味の提示順序

1) 辞書構造

従来の辞書においては意味の提示順序を決定すると、その順序を変えることは容易ではなかった。図1は本 web 辞書の構造である。見出し語単位ではなく意味毎に別ファイルを作成し、各ファイルが見出し語に格納されるようになっている。そのため意味の提示順序は容易に変更できる。

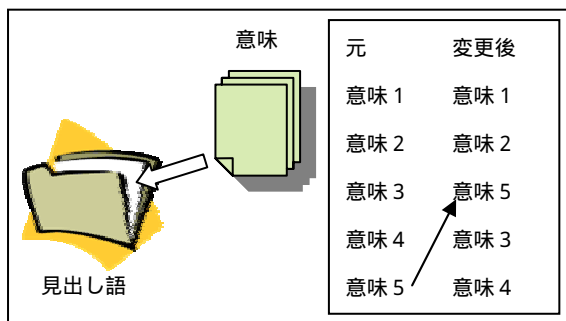


図1 web 辞書の構造

2) 意味の提示順序

各語の意味情報を入力する際に、日本語学習者にとって必要な意味かどうかを判定し、意味の提示順序についても再検討することにした。

その基準は、次の通りである。

- その語のもつ本来の基本的な意味
- 学習者に必要だと思われる意味
- 学習者に必要だと思われる意味に近いもの
- あまり使われないと思われる意味
- 将来的に辞書から削除してもよい意味

この提示順序が学習者にとって有益かどうかの調

査は今後の課題である。しかしながら、構造上意味の提示順序が自由に換えられること、国際共同作業により各国の複数の編集者が提示順序についても継続的に議論可能なこと、今後研究が進み意味について出現頻度や親密度が明らかになればそれに基づく変更が可能なことなど、新しい形の編集システムの持つ意味は大きいと考えられる。



図2 web 辞書の画面

5. まとめ

本研究の多言語版 web 辞書の試作版は <http://marmot.chuta.jp/> において利用可能である。今後もよりよい辞書を目指して開発を進めたい。

注1 日本電子化辞書研究所(1996)『EDR 電子化辞書』

参考文献

- 天野成昭他(2000)『NTT データベースシリーズ日本語の語彙特性 (第1期)』三省堂
- 加藤彰彦(1963,64)「日本語教育における基礎学習語」『日本語教育』2, 4, 5号 日本語教育学会
- 北村達也・川村よし子(2000)「読解学習支援システムにおける辞書提示方法の改善」『日本語教育国際シンポジウム』韓国国際交流基金(1986)『基礎日本語学習辞典』凡人社
- 国際交流基金・日本国際教育協会(1994,2002)『日本語能力試験 出題基準』凡人社
- 国立国語研究所編(1962)『現代雑誌九十種の用語用字』
- 国立国語研究所編(1964)『分類語彙表』
- 国立国語研究所編(1983,84)『高校教科書の語彙調査』
- 国立国語研究所編(1984)『日本語教育のための基本語彙調査』
- 国立国語研究所編(1986,87)『中学校教科書の語彙調査』
- (国立国語研究所編は全て秀英出版)
- 土居光知(1933)『日本語基本語彙』六星館
- 徳弘康代(2005)「中上級学習者のための漢字語彙の選択とその提示法の研究 - 学習指標値の設定と概念地図作成の試み」『日本語教育』127号 pp41-50 日本語教育学会
- 文化庁(1971)『外国人のための基本語用例辞典』大蔵省印刷局

