

9月15日(土) シンポジウム2 第2室(新館-752)

<基本語改訂委員会企画>

シンポジウム**JACET基本語(改訂4版)について**

JACET WORD LIST (FOURTH EDITION)

JACET基本語改訂委員会

○ 報告者：村田 年(委員長)

：石川慎一郎(委員)

：上村 俊彦(委員)

：投野由紀夫(委員)

：望月 正道(委員)

JACETでは、かねて懸案であった在来の「JACET基本語彙4000」の改訂に着手し、公募に応じた会員有志によって2000年春以来、その作業が進められてきた。一切の連絡や討議は、小委員会別や全体のメーリングリスト、専用のホームページを駆使して行われたことも新しい経験であった。「21世紀の英語教育—新たなる挑戦」の一環として、進行の概要を報告する。

1. 企画の経緯と先行研究

中学高校で学習すべき語彙数は学習指導要領の改訂のたびに減少し、その中で必修語の指定は事実上無いに等しいまでの数になった。効率よい学習のために、その目安になる語彙表の持つ意義は大きい。近年、プロシード5000、北大語彙表約7000、アルクSVL12000などが相次いでわが国の学習者用に発表され、片や海外ではBritish National Corpusなど、億を超える規模のコーパスを基にした各種の用語資料も出現するに至った。原資料から用語を切り出し、計測し、整理する技術も向上して、辞書編纂にもその成果が広く活用されている。今回の改訂に当たっては、それらの事実を十分に踏まえ、現状に即した対応が取られた。

2. 改訂の要点

前3版との根本的相違点は、全面的コンピュータ処理化にともなう変化である。機械処理の大原則である一語一意の制約から、同綴り語に番号を付して区別することや、英米語の綴りの差を部分的に括弧を挿入して表示するなどの便法が消えた。また、大学での2年間の課程修了までの語彙とされていたものに、さらにその後の方向性を与えるために採録語数が大きく増えることとなった。これは、近年とみに大規模コーパスが充実し、それらを参照して広範囲の統計的根拠が得やすくなったことにもよっている。書き言葉にとどまらず、話し言葉の特徴的傾向も反映したものとなった。

3. 資料の選択と収集

改訂作業における語彙出現頻度の意義と、今回準備した参照資料の特徴について概観する。資料の元データとなったのは、(1) 日常、見たり聞いたりする頻度の高い一般英語(映画、新聞、雑誌) (2) 英語の学習教材や試験問題(英語教科書、大学教養レベルの教材、入試等) (3) 専門分野の語彙(百科事典、科学雑誌の大項目や記事) (4) 口語英語データ(番組スクリプト、英語のスピーチやディベート他)等であった。著作権に配慮しつつ、これらのさまざまな分野の資料収集の経緯と、コンピュータ解析のための予備作業についても言及する。

4. 資料の吟味とデータ処理

今回の語彙リスト作成ではコンピュータによる自然言語処理技術を可能な限り導入して、人手によるリストの改訂作業をサポートする基礎資料を十分に作ることを目指している。具体的には、British National Corpus の見出し語化リストをベースに、日本の英語教育環境を考慮して選択作成された種々の言語コーパスに品詞タグ付与、見出し語化などの言語的情報を付加したのから語彙リストを抽出、それを用いてBNC list を補正する方向で進めた。

5. 語彙の選択と配列

既存語彙表の検討作業では、新基本語選定にあたり、KilgariffによるBNCレマ・リストと比較検討作業を行う既存の語彙表の選択およびその電子化作業を行った。既存語彙表の選択は、英英辞典定義語彙、英英辞典頻度情報、英和辞典重要語情報、EFL学習者用リスト、アカデミック語彙リスト、アメリカ教育語彙の6つの分野からそれぞれ2~4のリストを選び、計15を既存語彙表とした。次の作業として、15の既存語彙表をコンピュータに入力して電子化し、その点検作業を行った。これらから得た頻度情報を機軸として、わが国の現状に即した最適化が図られ、1000語刻みの配列となった。

6. 語彙表の吟味と展開

学習の目安となるべき語彙表は、本来、学習者個々の目標や現在の到達度によって異なる性質のものである。また、好んで用いられる語や、消えてゆく語の様相も、近年特に変化が激しい。全てをカバーしようとするれば、全てから遠くならざるを得ない矛盾の中で、この表の限界と有効性、これを基盤にした各種の有望な展開について考察する。