

第3回「情報システムのあり方と人間活動」研究会 開催概要記録

開催日時 平成21年9月12日(土) 午後1時30分～4時
場所 慶應義塾大学理工学部創想館2階 ディスカッションルーム7

=====

第1部 午後1時30分～2時30分 質疑20分

題目 「複数の観点で分類した自然言語処理用シソーラス」
講演者 (株)言語工学研究所 代表取締役 国分 芳宏氏

【講演概要】

自然言語処理のための「シソーラス」について国文氏より言語研究所での、約20年間にわたる研究成果である同社のシソーラスについて説明があった。

目的として、他社シソーラスもあるが日本語処理一般に使うことを狙いとしました。

2009年7月現在、収容語数は42万語で最多と考えている。

ボトムアップでの作成により名詞のみでなく動詞、形容詞、形容動詞、副詞、代名詞、擬態語さらに慣用句までを収容している。時事的な用語の積極採択、地名を除いて固有名詞は含めない編集方針としている。言葉を探すのを目的とする人間の感覚に沿った分類とした。色の分類時に、「はでな色」、「暖かい色」とする例が挙げられる。言葉の意味空間は、1次元ではなく、例えば、「料理」は、材料・地域・調理法の3次元でありこの様に複数の観点で分類している。「現在」に対する、「明日」、「翌日」、「過去」に対しては、「翌日」が意味を持つ様に分類作業における揺れを吸収している。

補助的な記述内容として「分類の観点を表示」、「ある概念を詳細に分けるとときに、分けた分野に対応する適当な用語がなく恣意的な用語になるのを防ぐ」、「多義語を区別する」があげられる。同義語、反義語、狭義語、広義語、関連語、共起語(係り受け関係を構成する用語の組等)等の用語同士の関係についても整理している。同義語の例として、大和言葉「しお」は、漢語(複合語)「食塩」、カタカナ語「ゾルト」、英字表記「NaCl」がある。人の立場によって「税金」・「公的資金」・「血税」の様に入語が異なる例もあり分類の難しさを示している。同義語の一種である表記の揺れ(犬、イヌ、いぬ)については、代表用語を推奨語として決めている。検索漏れを防止するため用語の標準化を行った。(例:米、米国、U.S.A.をアメリカに標準化)

自然言語処理で広義語に適用した規則は、狭義語にも適用できるように属性が同じものだけを広義語、狭義語関係にした。反義語は、意味的に近い関係にある言葉で、片方を否定すると相手になる(例:善 ⇔ 悪)等が特徴で、反義語の対立の種類について言うならば、家族内の「兄」に年齢で対立する語は「弟」、性別では「姉」の様な関係となる。

日本語の特徴として多義語の多さがあげられる。例として、大和言葉「うめる」で「穴をうめる」、「時間をうめる」、外来語「ライト」で「光」、「照明」、「右」等があるが多義語も複数観点で弁別している。関連語とは、ある程度の意味的な関連性を持つ用語の関係を言い、同じカテゴリーの用語(例:食材 ⇒ 肉、野菜)と異なるカテゴリー(食材 ⇒ 肉と料理 ⇒ 肉料理と異なる)が、肉・肉料理は関連語に分類している。

自然言語処理で使うために用語(動詞、形容詞)、副詞も必要なので、用語は語幹と活用形で登録してあり構文解析に利便をはかっている。日本語では、慣用句が大きな位置をしめているので収容している。語の関係は、時と共に変わるのでこの点にも注力している。

構文解析上、係り受けであいまい性のある係り先を決定したり、良し悪しを決定したりするときのために共起語を持っている。係り側の格助詞を含めて7万組を管理している。

例として、「寿命」、「延びる」、「短い」等は、組み合わせられたときに、「良し悪し」の性質が出てくる。例文：寿命が延びる（良い）、寿命が短い（悪い）用語の意味的な距離について、「意味的な距離」の定義を行うことにより確定させる方法を採用しているが、距離を測るためには多義語を区別することがポイントになると考える。

パッケージソフト機能として、辞書に未登録用語の形態素分析が行える（語を分解して検索可能、例：シソーラス研究会は未知語で、シソーラス、研究会で検索可能）、語末一致用語を検索（例：アカトンボ、イトトンボ）、古い用語を休眠させる機能（例：用語 写真機は、カメラ ⇒ デジカメと移り変わりしているので用語 写真機を休眠させる）。

又、専門用語、外国語を辞書にCSVファイルで併合も可能であり、標準用語でない場合に画面上にピンク表示、差別語・誤った表記を赤字で表示する等、機能を工夫している。

インターネット・電子化された辞書との接続、他のプログラムからも文章推敲中に、シソーラスでより適した用語を探すことできるとの説明がありました。

=====

第2部 午後3時～4時 質疑 20分

題目 「高等学校における情報教育の現状と展望」

講演者 兵庫県立社高等学校 教諭 山上 通恵氏

【講演概要】

「高等学校における情報教育」について、他教科と異なり新規教科であり共通認識がないことが課題と考える。本論に入る前に自己紹介として、1992年4月 兵庫県立神戸甲北高等学校（数学）に着任、97年兵庫教育大学大学院学校教育研究科へ内地留学し、「高等学校における情報教育のカリキュラム」開発。98年4月神戸甲北高校へ帰任（情報）し、2001年8月に現職教員等免許講習会で「情報科」免許取得、2003年度には現行学習指導要領が実施された後、2009年4月 兵庫県立社高等学校に転勤（数学）し現在に至っている旨の紹介があった。歴史的には、情報科については2003年度より開始されたが、教員養成は2000年度から3年間で9000人を養成する計画で数学科・理科・家庭科・工業科・商業科の教員を対象に夏休みの3週間で急造された。

従って、「情報教育の範疇」については異なる認識が生じている。

コンピュータが使えるれば良い、ワープロ検定に合格すれば良い、プログラミング技術の習得ではとの意見があるが、情報モラルの育成、メディアリテラシーの範疇もあると考えられる。

情報教育に関しては、「情報化の進展に対応した初等中等教育における情報教育の推進等に関する調査研究協力者会議」最終報告で、1 情報活用の実践力 2 情報の科学的な理解 3 情報社会に参画する態度の3観点が生じていて、現行科目の情報 A、情報 B、情報 Cは対応している。ただし、情報 Aは観点1に、情報 Bは観点2に、情報 Cは観点3に各々若干ウェイトをおいている。ところが、情報 A、情報 B、情報 Cが、3観点の各々に対応していると誤解されていて教育現場で先導する教師の悩みとなっている。

科目選定の実態から言えば、学習指導要領に、「・・・生徒の経験や興味・関心の多様性を考慮し、3科目を用意し1科目を選択的に履修できるようにした」との記載があるが、2003年度、2009年度共に教科「情報 A」を採択するケースが圧倒的に多い。この事由は、情報 Bを教えることのできる教員数、能力が限定されている。情報 Cは内容が高度であることが大きな理由で、消去法的に情報Aが選択されている。これは、教科書発行所別統計からも明らかである。過去、未履修問題が、進学校を中心に必履修科目「世界史」に生じたが、状況は、必履修科目「情報」についても同様に、情報処理学会などが、教科「情報」の必要性を声明として発表したことがある。今回、学習指導要領の改訂の時期を控え、各方面から議論がされた。その中でも、全国校長会、全国PTA連合会、一部の教職員団体、一部の「情報科」担当者のグループから、「情報」を必履修から外し、選択にせよとの強力な意見があったが、今回の学習指導要領改訂にあたっては必履修が堅持された。今後も必履修より選択にとの意見は生ずる可能性があるもので重要性を高める必要がある。

次期科目構成は、1 社会と情報 2 情報科学から構成され、
1については、情報 C+アルファ+情報 A
2については、情報 B+アルファ+情報 A
になると予想されるが、情報 Aについては、中学校で大半を学習済みと考えられるので、発展的に解消するものと考えている。

今回の学習指導要領に関する改訂ポイントは、「社会と情報」と「情報の科学」に共通した情報モラルの学習活動を重視する点、「社会と情報」では、「情報とメディアの特徴」、「コミュニケーション手段」などの項目が増え、情報の表現や効果的なコミュニケーションを行うために情報機器や情報通信ネットワークを適切に活用する学習活動を重視する点、「情報の科学」については、問題解決が項目立てされ、問題解決を行うために、情報と情報技術を効果的に活用する学習活動やそのために必要となる科学的な考え方を身に付ける学習活動を重視する点である。

今後の科目「社会と情報」、「情報の科学」の採択率について、2009年度に、情報Aは、60%、Bは、10%、Cは、30%であるが、2013年度には、「社会と情報」は、90%、「情報の科学」を10%と予想している。

今後、採択率に偏りが生じないように対策を提言したい。

中学校の技術過程で、機器の制御や簡単なプログラミングを経験した生徒が、高等学校でより高いレベルの学習を求めるのは当然であり、教師の力量に依存して、当たり前のように「情報 A」が開講され、生徒は中学校で学習したのと変わらない内容を学習する実態を変革するために是非、教師による採択から生徒による選択へと方針を転換することをお願いしたい。

以上