様式 C-19

# 科学研究費補助金研究成果報告書

平成 23 年 6 月 20 日現在

機関番号：82610
研究種目：若手研究（A）
研究期間：2008 ～ 2010
課題番号：20689037
研究課題名（和文） 国際医療用語集を日本語で利用するための知識ベースの開発

研究課題名（英文） Development the international health terminology knowledge base for Japanese

研究代表者
柏木公一（KASHIWAGI KIMIKAZU）
独立行政法人国立国際医療研究センター・研究員
研究者番号：20334378

研究成果の概要（和文）：
この研究の目的は、国際医療用語集を日本語で利用するための知識ベースを開発することである。知識ベースとは、単に日英の対訳データベースではなく、そこから訳語に関する知識を抽出するツールである。登録データとしてはライセンスを取得した看護診断、看護行為、看護介入の用語集、並びに無償で研究利用が可能な医療用語集を用いた。これらの用語集から対訳データを作成し、その後、単語の出現頻度などから訳語のバリエーションを調査した。さらに、この知識ベースから、翻訳作業を支援するツール、用語を意味によって整理するオントロジー構築の２つを試作した。

研究成果の概要（英文）：
The purpose of this study was to develop the knowledge base of international health terminologies which can be used in Japanese language. This knowledge base is not only the database of English-Japanese translations but the knowledge source which identifies the diversity of translation and the gap between concept and terms among other languages. The data source included NANDA, NIC, NOC, ICNP, and other health terminologies such as ICF and MedDRA. The derived products of this knowledge base are tools for selection of translation terms and the helper of developing ontology.

交付決定額

（金額単位：円）

| | 直接経費 | 間接経費 | 合　計 |
|---|---|---|---|
| 20 年度 | 3,700,000 | 1,110,000 | 4,810,000 |
| 21 年度 | 1,200,000 | 360,000 | 1,560,000 |
| 22 年度 | 1,100,000 | 330,000 | 1,430,000 |
| 総　計 | 6,000,000 | 1,800,000 | 7,800,000 |

研究分野：基礎看護学
科研費の分科・細目：
キーワード： 標準化、看護用語、対訳

## １．研究開始当初の背景

海外では医療用語の標準化が進められており、2007 年 3 月に国際医療用語集標準化団体 IHTSDO（International Health Terminology Standard Development Organization）が 9 ヶ国（アメリカ、イギリス、カナダ、オーストラリア、デンマーク、スウェーデン、オランダ、ニュージーランド、リトアニア）によって設立され、米語、英語、スペイン語、ドイツ語に翻訳された約 30 万概念を含む SNOMED-CT という国際用語集が採択された。
一方、日本では、電子カルテ普及のために、病名、検査、薬剤、看護などの用語集が整備されているが、それらを統合した標準用語集

は存在していない。特に、診療報酬請求に関連が低い症状・所見や看護に関する用語は、普及が遅れている。
このような状況の中、国際的な医療用語の標準化に対応するために、様々な議論が行われている。例えば、国際医療用語集と日本語の概念にはどのような相違があるのか、そして、それは診療記録の電子データを分析する際にどの程度の影響があるのか、また、日本語にしか存在しない概念がどの程度あり、それを国際医療用語集の中でどのように取り扱うべきかなどである。これらの議論を行うためには、実際に日本語と国際医療用語集との対応づけを行い、問題点を事前に明らかにする必要がある。
看護の分野においては、様々な海外の標準用語集が日本語に翻訳されている。しかし、その訳語が翻訳者によって決定されており、翻訳者によって相違がみられるなど、学術用語としての日英対訳集が整備されていない。

## ２．研究の目的

本研究では、すでに存在している看護の用語集について日英の対訳を作成し、他の医療用語集の日英対訳とあわせたデータベースを作成し、日本語と国際医療用語集との対応付けを行う知識ベースを作成する。

## ３．研究の方法

### １）対象用語集

① NANDA
NANDAインターナショナルによる看護診断用語集。日本看護診断学会監修により医学書院から日本語版が出版。
② NIC
アイオワ大学による看護介入分類。南江堂により日本語版が出版されている
③ NOC
アイオワ大学による看護成果分類。医学書院より日本語版が出版
④ ICNP
国際看護師協会による看護現象と看護介入分類。第１版が日本看護協会出版会より刊行
⑤ MedDRA
ICHによる副作用など薬剤関係の用語集。一般財団法人 医薬品医療機器レギュラトリーサイエンス財団により日本語訳。
⑥ ICF
国際保健機関（WHO）による国政生活機能分類。障害者福祉研究会により中央法規出版から日本語版が出版。
⑦ UMLS
米国医学図書館が作成した多言語医学用語集。日本語としては医学中央雑誌シソーラスなどが収録。

ライセンス
データの電子保存利用に関するライセンス規定がある用語集については、著作権を保持している出版社と本研究目的によるライセンス契約を締結。一部用語集については電子データを購入した。

### ２）日英対訳データベースの作成

① データスクリーニング
日本語と英語の出版物には、対応付けのミス、翻訳忘れなどが存在し、電子データや個別IDが付与されていないものがあった。これらについては人手によりデータの対応付けを行った。また、日本語翻訳が欠落している用語については、研究者により翻訳を補った。

②データベースの設計
基本的に日本語、英語の対訳が１対１で対応するデータとして登録を行った。登録単位としては、診断ラベルの場合はその診断ラベル。定義など文章の場合には１文単位。関連因子のような句については句単位とした。

コンテキスト情報
用語の意味はそれが収録されている用語集内での位置づけにより異なるため、それぞれにコンテキスト情報を付与した。
コンテキスト情報とは、ICNPのように階層構造を持つ用語集の場合には、上位・下位関係を持つ情報、NANDA用語集のように内部に定義文、関連因子といった内容を持つものについては、属性として登録を行った。これらのコンテキスト情報は、訳語検索の際に、元の構造をたどれるような構造として登録した。

単語レベルのデータベース
訳語を決定する際には、各英単語がどのような訳語として使用されているかの頻度データが必要である。そのために、対訳データベースに登録された英単語を単語別に分解し、それぞれの単語がどの日本語訳として利用されているかを登録するデータベースを作成した。英単語については活用形などを機械的に正規化した。
一方、日本語の訳語がどのような英単語から生成されているかも同様に検索できるためには、日本語についても単語レベルのデータベースが必要である。日本語については、英単語ような単語の区分けが困難であり、本研究ではChasenという形態素ツールに一部の医療用語を追加した辞書を用いて形態素分割を行った。

Nグラムデータベース
連続する単語によるデータをあらかじめ登録するデータベースである。後述する

"unintended pregnancy"をひとかたまりとして登録するデータベースを用いた。構文情報を持っていないため、このデータは単純検索として用いることとした。

３）知識ベースの作成

①設計

英単語がどのように日本語訳されているか、また日本語訳に用いられた用語がどのような英単語に基づいているか、この両者バリエーションを自動生成することを目標として設計を行った。

具体的には、英単語、日本語の形態素それぞれについて対訳データから共起情報を取得し、特異的な関係を持つ関係を訳語候補とした。

②機能のリスト

知識ベースの機能としては以下の物を実装した。

・日英の基本的な検索機能
・コンテキスト情報の表示・追跡
・単語の出現頻度とデータソース
・日英それぞれの単語の共起情報

４．研究成果

１）日英対訳データベース

日英対訳データベースからは、各英単語についてどのような日本語訳語が生成されているかのバリエーションが判明した。

①　状況によって訳語が変化する場合

意味としては同義であるが状況によって訳し分けられている単語。例えば、Ambulationという言葉は、状況に応じて「移動」「歩行」「離床」のように訳し分けられていた。同様に bathing は乳幼児と成人の違いによって「入浴」「沐浴」と訳し分けられていた。これは日本語に同一の意味を示す言葉が存在せず、場面によって使い分けられるときにみられた。

一方、comfort のように意味が広く適切な日本語が存在していない場合には、「安楽」「違和感がない」「心地よい気分」「身体的に楽」「心理的に楽になった」「気にならない」「安心感」「不快なし」「快適さ」のように多様な日本語訳が存在していた。このような語句の場合は、日本語訳バリエーションの中からそれぞれの文脈に適した用語を採用することが望ましいと考える

②訳語が定まっていない場合

日本語として存在していない言葉を訳す場合に、訳語が不定となっている場合があった。Caregiver は、「ケア提供者」「介護者」「世話をしてくれる人」「ケアの受け手」「家族介護者」のように訳されていた。同様に unintended pregnancy は、「意図しない妊娠」「非計画的妊娠」「予定外の妊娠」と訳されたいた。①の状況による訳語変化とは異なり、状況との相関がないまま訳語が複数存在していると考えられた。このような訳語に関しては、訳語の統一か、優先語(Preferred Term)の指定が有効であると考えられる。優先語は、どの訳語も否定することなく代表的な訳語として１つを指定する方法である。

③解釈が異なる場合

一つの英単語の解釈が異なり、その結果日本語にも差異が見られている場合があった。Crawling は「ゆっくり歩くこと」「這うこと」と訳されていたが、前後の文脈からはどちらが適切か判断できなかった。同様に somnolence は「倦怠感」「傾眠」と訳されていた。多言語で意味の統一を行う場合、ネイティブは気がついていない多義解釈が問題となる場合がある。英語そのものに複数の解釈が可能な場合は、定義文などの解説が必要である。

④日本語訳として不適切な場合

日本語訳として不適切と思われる訳語も存在した。

２）翻訳作業を支援するツール

翻訳作業において訳語を決定する際には、過去の翻訳事例から訳語候補を検索する必要がある。本研究の知識ベースからは、英単語、日本語両者からどのような翻訳が行われているかを探すことができ、さらにその訳語が用いられたコンテキストを調べることができる。

そこで本知識ベースを用いて、訳語候補とその管理ができるツールを作成した。このツールによって ICNP のようなリファレンスターミノロジーの翻訳作業において、複数訳語とその使い分けを提示することが可能となる。

支援ツールでは、単語レベルの訳語候補の提示、複数語の自動抽出といった知識ベースのインターフェースを利用した機能と、訳語候補とその履歴管理の機能を実装した。

３）オントロジー構築

ICNP のような日英の単語レベルの翻訳を行う場合、英語の動詞の訳し分けが困難である。本知識ベースにおいては、看護介入分類 NIC のような行為を含む用語集から動作を表す動詞とその訳語を抽出することができる。これによって、動詞訳語のバリエーションが調べられる。

一方、訳語ではなく意味によって動詞を整理することを考えた場合、日本語と英語の動詞間でどのような違いがあるかに着目し、動詞を分類する方法がある。本知識ベースからはこのような意味に基づく分類のためのデータを抽出できる。この意味に基づく分類からは英単語だけでは発見しにくいオントロジ

一の構築が可能となると考えている。今後、検討中に看護用語オントロジーにおいてこの知見を活かしていきたい。

５．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔学会発表〕（計 ８ 件）
①柏木公一、電子カルテ時代の看護記録と看護情報:用語の標準化について、第30回医療情報学連合大会、2010年11月20日、浜松

②Kimikazu KASHIWAGI、Meaningful Use of Electronic Health Record、HIMSS Aisa10 Health IT Congress　2010年10月26日、Korea

③KIMIKAZU KASHIWAGI、Development and Application of Nursing Information System、International Symposium of Nursing Decision Making and Informatics Application、2010年7月30日、Taiwan

④柏木公一，香西ひろみ、看護プロファイルの統合、第 11 回日本医療情報学会看護学術大会論文集，2010年6月26日、仙台

⑤柏木公一、用語標準化の最前線:ISO/TC215 WG３の活動、第 14 回日本医療情報学会春季学術大会、2010年5月29日、高松

⑥柏木公一、電子カルテ時代の看護記録と看護情報:看護記録の改善、第14回日本医療情報学会春季学術大会、2010年5月28日、高松

⑦Kimikazu KASHIWAGI、A feature of Nursing Information Systems in Japan、Asia Pacific Association for Medical Informatics、2009年11月22日、Kobe

⑧Kimikazu KASHIWAGI、An Ontological Approach for translating the ICNP、10th International Congress on Nursing Informatics、2009年6月30日、Helsinki(Finland)

６．研究組織
(1)研究代表者
柏木公一（**独立行政法人国立国際医療研究センター・研究員**）
研究者番号：20334378