

実践医療用語における語構成要素の意味ラベルについて

相良 かおる
西南女学院大学
sagara @ seinan-jo.ac.jp

1. はじめに

着任した職場において看護教員が口にする「患者さんに介入する」や「緩和ケア」という表現に違和感を覚えた筆者は、臨床現場で使われる看護用語に興味を持った。

同じ頃、厚生労働省は、2006年迄に全国400床以上の病院の6割に電子カルテシステムを導入するという目標を掲げた。また国内外の看護系学会では、看護用語の標準化についての検討がなされ、幾つかの標準看護用語のモデルが公開されたものの、これらがどの程度普及するかは不明であった。

筆者は、電子カルテの普及により医療記録が電子化され、機械処理が可能となれば、患者診療用途（一次利用）に加え、統計資料、臨床研究や疫学研究、教育訓練、そして診療情報管理などの2次利用が要望されるだろうと考えた。そこで、電子化された医療記録データの自然言語処理を支援するために看護実践医療用語を見出し語とした分ち書き用辞書 ComeJisyo を2008年11月に公開した。その後、随時更新し、2020年4月公開の ComeJisyoSjis-2 の登録語数は、医師経過記録より抽出した実践医療用語を含め約11万語となった。

なお、門外不出の医療記録に含まれる実践医療用語の語構成や語種構成の実態は不明であり、よって本辞書では、語の単位認定の規則を定めず、看護師、管理栄養士、医師としての臨床経験者が一つのまとまった語としたものを登録している。

その結果、本辞書には助詞などを省略した臨時一語を含む合成語が多く登録され、語彙研究の言語資源としての利用も可能となった。そこで、2018年より日本語学の研究者、そして看護師および臨床医の経験者と共に実践医療用語の合成語7,194語の語構成の解析とこれらの意味の解析に着手している。

本発表では、本合成語7,194語から抽出した語構成要素6,442種類に付与した86種類の意味ラベルについて述べる。

2. 用語の定義

筆者は、日本語学の専門家ではない。従って、抄録を書く際などには、「複合語」と「合成語」の違いや、「語」と「単語」の定義などを複数の辞書や事典、日本語学の専門書などで調べ、普遍的な定義がないことに戸惑う。幸い共同研究者には、日本語学の研究者がいることから、草案の段階でこれらの用語が見直される。

定義の曖昧な語の使用は、時折誤解を生む。

本研究の成果物の利用者は、医療従事者、医療の専門科目の教育者、医療情報を扱う関連分野の専門家、そして言語の研究者などを想定している。

そこで、分野の異なる利用者のために本研究で用いる用語を以下のように定義する。

実践医療用語：

専門分野で使われる「専門用語」には、①学術上の専門用語（学術用語）と、②それ以外の専門用語がある。本研究では、医療施設で使われる医療記録に含まれる①と②を合わせた用語を「実践医療用語」という。

語構成要素：

合成語を構成する要素で、本研究では「合成語を医療の観点から意味的にまたは統語的に分割可能なすべての部分文字列」と定義する。分割できない合成語については元の合成語を語構成要素とする。

合成語：脳幹多発性硬化症

語構成要素：脳幹、多発性、硬化症、多発性硬化症

意味ラベル：

語構成要素に付与する意味を表すラベルで、共同研究者らで命名したものである。

「回旋位」のように他の語構成要素と結合することで異なる意味になるものには「・」を用いて「状態・位置」のように複数の意味ラベルを列挙する。また、「胃粘膜下」のように位置と身体部位の両方の意味を持つものには「/」を用いて「位置/身体部位」のように表記する。

なお、ここでの「意味」には、用語の意味を表すもの（例：「衛生用品」）と、医療者の使用面での心的な捉え方を表すもの（例：「病因」）が含まれる。詳細は後述する。

3. 研究方法

3.1 言語資源

実践医療用語を語構成要素に分割する際、医療の知識がなくても語の境界を見つけ易くするために、分ち書き用実践医療用語辞書 ComeJisyoSjis-1 の登録語 111,664 語から、一般的な語（『分類語彙表増補改訂版』[1]収録の語）を含む合成語 7,194 語を選定し本解析データとした。

3.2 研究デザイン

筆者の研究目的は、医療記録データの自然言語処理を支援することである。そして、本研究の第一目標は、医療記録に含まれる合成語を構成する「実践医療用語語彙表」を作成することであり、第二の目標は、実践医療用語の語構成に関する言語学的な知見を得ることである。そして第三の目標は、得られた知見を含む成果物を医療実践、医療教育の領域で、出来得れば言語学研究の領域においても利用可能な形で公開することである。

専門用語の意味の分類に関する意味理論には、ターミノロジー学、形式意味論、語彙意味論、認知意味論などがあり、それぞれに制約がある[2][3][4]。

一方、実践医療用語には、程度や数量、身体的部位や体位、時間的情報、患者の主観的状态などを表す表現が含まれ、既存の医学および看護学の学術用語を対象としたシソーラスの分類カテゴリーをそのまま利用できないことが分かっている[5]。また、「ポリビア／出血熱」のように単語の意味「地名」と「病名」を合成しても合成語の意味「ウイルスによって引き起こされる出血性熱性疾患」にならないものがあることも分かっている。更に、看護の知識を持たない筆者の「介入（強引に関わる）」と、看護学での「介入（看護の対象者に行う問題解決や回復への援助）」、医学での「介入（研究目的で、人の健康に関する様々な事象に影響を与える要因の有無又は程度を制御する行為）」には意味のズレがあることも分かっている。従って、実践医療用語の意味の分析に適した意味理論を見付けられず、また本研究の目的は、新しい意味理論を構築することではない。

そして本解析データは、実践医療用語の母集団を代表する標本ではない。では、本研究の研究デザインは何かと考えると、以下に示す臨床研究における(1)の中の比較対象が明確ではない「仮説を見つけるための記述研究」に当てはまる。

- (1) 記述研究…仮説を立てる
- (2) 分析研究…仮説を分析する
- (3) 介入研究…仮説を検証する

3.3 分析方法

本節では、語構成要素の抽出方法と意味ラベル付与の手順について述べる。

手順 1. 機械的分割

- (1) 形態素解析器 MeCab0.996[6]と見出し語約 87 万語の解析用辞書 UniDic[7]により、合成語を短単位に分割
- (2) 品詞の解析誤りを修正し、「接尾辞」は語を直前の語に、「接頭辞」は直後の語に連結
- (3) 連続する「カタカナ」のみの語を連結

機械的短単位分割：肝内 | 門脈 | 下 | 大 | 静脈 | 短絡

手順 2. 意味的分割（一般的）

分析者 5 名により手順 1 の語を分担して意味的に妥当な語にまとめる。表 1 の意味カテゴリー[8]を参考に語構成要素の意味に着目し、意味ラベルを付与した。

表 1 意味カテゴリー

自然物	動植物	物品	食品
道具	薬品	力	人間
機械	衣料	部分	家具
資材	地類	容器	建物
空間	形状	数量	動き
状態	時間		

なお、分析者の専門領域は、日本語学が 3 名、情報科学が 1 名、看護学が 1 名である。

語構成要素列：肝内 | 門脈下 | 大静脈 | 短絡

意味ラベル列：身体部位 / 位置 | 身体部位 / 位置 | 身体部位 | 状態

手順 3. 医療的観点による分割（看護学）

臨床看護経験者 2 名により看護の観点から見直しを行い、意味により分割した「短い単位」による語構成要素列（短）と統語的な纏まりを考慮した「長い単位」の語構成要素列（長）を作成した。

語構成要素列（短）：肝内門脈 | 下大静脈 | 短絡

意味ラベル列（短）：身体部位 | 身体部位 | 状態

語構成要素列 (長) : 肝内門脈下大静脈 | 短絡

意味ラベル列 (長) : 身体部位 | 状態

手順 4. 医療的観点による分割 (医学)

臨床医により医学の観点から見直しを行った。

語構成要素列 (長) : 肝内門脈下大静脈短絡

意味ラベル列 (長) : 身体部位・病態

4 結果

4.1 語構成要素数の概要

合成語 7,194 語より分割された語構成要素は 6,442 要素であり、これらに付与された意味ラベルは 86 種類であった。また、複数の意味ラベルが付与された語構成要素は 868 要素 (表 4 と表 5 の総計) あった。

4.2 意味ラベルの概要

意味ラベル 86 種類を、「具体」と「抽象」で大別し、「具体」に属する意味ラベル 20 種類をまとめたものが表 2 である。表 3 は「抽象」に属する 66 種類をまとめたものである。表の中のゴシック体は、手順 3 の看護学および手順 4 の医学の観点から追加・変更されたものである。表 2 では、医学の観点から「物質」が「体内物質」と「体外物質」に分けられ、「薬品」は、「薬品」と「医薬品」に分けられた。またウイルスは生物とは言えないことから「病原体」が追加され、「人間」に加えて「ヒト (生物学上の標準和名)」が追加された。

表 2 意味ラベル (具体)

ヒト	光	人間	動物
医薬品	施設	組織	病原体
衛生用品	場所	体外物質	物質
器具	食品	体内物質	物品
機器	身体部位	動植物	薬品

表 3 では、「状態」が「状態 (一般的な状態)」, 「症状 (本人の訴えによる状態)」, 「病態 (体の中で起こっている機序)」に細分化された。また「指標」, 「手技」, 「病因」, 「患部」, 「患者属性」が追加された。

手順 2 の実践医療用語を使ったことのない分析者による分類では、国語辞書に記載されている普遍的な意味による意味ラベルが付与されたのに対し、実践医療用語の使用経験を持つ分析者による見直しでは、使用面での意味による意味ラベルの追加・変更がなされた。例えば、手順 2 において表 1 の意味カ

テゴリに則ると、語構成要素「高齢者」の意味ラベルは「人間」であったが、手順 4 の医学の観点からの見直しにより、意味ラベルは「ヒト・患者属性」となった。同様に「血压」や「体温」の意味ラベルは「数量」から「指標」に変更された。

表 3 意味ラベル (抽象)

エネルギー	軌跡	手段	費用
サービス	距離	種類	病因
位置	教育	順序	病態
医療行為	傾向	症状	病名
運動	形状	状態	部分
音	形態	色彩	変化
課題	経過	心理	保健衛生
学術	検査	数量	保留
患者属性	行為	制度	方向
患部	行動	性質	方法
感覚	作用	生理	法規
関係	指標	精神	様相
関係 (文法)	時間	増減	例示
基準	自然	程度	属性
基礎	社会	動き	調節
機能	主体	熱	
規則	手技	能力	

4.3 多義の意味ラベル

表 4 多義の意味ラベル「・」

	意味ラベル	頻度	相対頻度
1	病名・病態	171	23%
2	病名・状態	69	9%
3	体内物質・体外物質	66	9%
4	状態・病名	63	8%
5	種類・病態	61	8%
6	病因・病態	46	6%
7	状態・経過	39	5%
8	ヒト・患者属性	32	4%
9	病因・経過	31	4%
	小計	578	77%
	総計	753	100%

接続する他の語構成要素により意味が定まる多義の語構成要素 753 要素 (全語構成要素 6,442 要素中、11.7%) に付与された「・」で列挙された意味ラベルの組は 53 種類あった。表 4 は、これらの内、頻度が 20 以上の語構成要素をまとめたものである。「病名・状態 (例: 血管損傷)」と「状態・病名 (例: 横断骨折)」の違いは、主観的な判断によるものである。

表5は、語構成要素に複数の意味を持ち「/」により列挙された意味ラベルと頻度をまとめたものである。

表5 多義の意味ラベル「/」

意味ラベル	頻度	相対度数
位置/身体部位	111	97%
状態/動き	2	2%
方向/動き	1	1%
形状/物質	1	1%
総計	115	100%

5. 考察

今回、手順3「看護師」と手順4「医師」による見直しを行い、使用面での心的な捉え方を知ることができ、更に医療従事者の中でも職種により捉え方に違いがあることが分かった。以下に使用面における心的な意味について考察する。

実践医療用語を使用しない筆者らが「人間」の意味ラベルを付与した「巨大児」は、看護の観点から「状態」に変更され、「女性」、「ドナー」は、医師の観点から「患者属性」に、そして「患者」は、「ヒト・患者属性」に変更された。また、「状態」を付与した「呼吸困難」は、看護の観点から「症状・状態」に変更され、「低血糖」と「感染」は医師の観点から「病態」と「病態・病因」に変更された。

これらのことから、患者を観察し援助(care)を行う看護の観点からは援助を要するか否かの判断を示す「症状」という捉え方があり、診断し治療(cure)を行う医師の観点からは「病態」や「病因」という診療の流れに沿った捉え方があることが分かった。

また、「ウイルス」の意味ラベルは「病原体」であるが、「ウイルス性」の意味ラベルは「病因」となり、その結果、同義語である「ウイルス/感染症」の意味ラベル列は「病原体 | 病名」、「ウイルス性/感染症」の意味ラベル列は「病因 | 病名」となる。

このことから、「病原体」が付与された語構成要素の語末に『性』が続くと「病因」となり、その後ろに「病名」が続く合成語(例:「アメーバ性/肝膿瘍」)の統語構造は、「病因」による「病名」と推測できる。さらに「病原体+病名」となる合成語(例:「腸炎ビブリオ/食中毒」)の統語構造もまた「病因」による「病名」であろうという予測が可能となる[9]。

6 おわりに

本研究では、看護および医学の実践医療用語の使用面における心的な意味が明らかとなった。しかし「病態」や「指標」は、医療の知識のない者には馴染みのない語である。そこで再度、意味ラベル名の妥当性を見直した上で、これらをまとめ、以下のよう一般的な意味ラベルを付与する予定である。

状況：(状況, 症状, 病態)

物質：(体内物質, 体外物質)

また「病名・状態」や「状態・病名」のラベルの順序は、主観的な判断によるものであり、これらについても検討が必要だと考えている。

なお、本研究の成果は『語構成要素語彙試案表』として公開する予定である。

謝辞

共同研究・分析者である小野正子氏、高崎智子氏、東条佳奈氏、麻子軒氏、山崎誠氏に深謝いたします。

本研究は、科学研究費補助金「語形成および意味的情報を付加した実践医療用語辞書の構築」(JP18H03499)の助成を受けています。

参考文献

- [1] 国立国語研究所. 分類語彙表 増補改訂版. 大日本図書, 2004.
- [2] 尾関周二, クリスティアン・ガリンスキー編著. ターミノロジー学. 文理閣, 1987.
- [3] 金水 敏, 今仁生美. 現代言語学入門 4 意味と文脈. 岩波オンデマンドブックス, 2016.
- [4] 郡司隆男, 阿部泰明, 白井賢一郎, 坂原 茂, 松本裕二. 言語の科学 4 意味. 岩波書店, 2004.
- [5] 相良 かおる, 小野 正子, 上野 恵子. 医療用語のシソーラス作成にむけた予備的調査. 西南女学院大学紀要, 2015, Vol.19, p.109-118.
- [6] MeCab : <https://taku910.github.io/mecab/> (参照 2021-1-12)
- [7] UniDic : <https://unidic.ninjal.ac.jp/> (参照 2021-1-12)
- [8] 石井正彦. 現代日本語の複合語形成論. ひつじ書房, 2007. 180-183.
- [9] 麻子軒, 相良かおる, 高崎智子, 東条佳奈, 山崎誠. 意味ラベルを用いた「一性」を含む病名の言い換え. 人文科学とコンピュータシンポジウム論文集. 2020, No.2, p.283-288.