著者キーワードのMeSHへの変換

Conversion of Keywords Author-Assigned to MeSH

谷 垣 勲*

滋賀医科大学附属図書館

Tanigaki, I. : Conversion of Keywords Author-Assigned to MeSH. *Igaku Toshokan* 1990 ; 37(2) : 71-77.

Keywords reflect the major topics of the article. They assist indexers in cross-indexing. This survey discusses instructions concerning keywords in information provided to authors and correlation of keywords assigned by author with MeSH.

I. はじめに

わが国の科学技術論文や科学レポートの記載要領の基準であるSIST（科学技術情報流通技術基準：以下SISTという）は、キーワードを「論文や記事の主題を表わすために付与された語句」と定義し、「検索や索引作成の際の手掛かりとして各論文に付与することが望ましい」¹⁾とキーワード付与の目的を述べている。Council of Biology Editors Style Manualでは「キーワードに選ばれた用語は論文の主要概念を表わす。これらの用語は、標題、抄録、本文から採られ累積索引の見出し語になる。」²⁾と述べている。

キーワードは、論文内容を濃縮して主要な概念を表現する。抄録もキーワードと同じく論文内容を要約して情報を伝える。両者とも、論文内容を速く的確に把握させ、本文を通読する必要があるかどうか判断するめやすを与えるものである。キーワードは抄録より目に止まりやすく、一段と内容を集約している。通俗的に言えば、論文のキャッチフレーズの役割をになっている。もちろん、学術論文のキーワードは商業コピーとちがって、斬新さ、珍しさで人を引きつけるものではなく、多くの研究者と共通理解が得られる普遍性の方が求められる。

医学分野に限らず、学術情報は加速的に増加してい

る。学術論文といえども、著者は他の研究者の目にとまることを望んで、その研究成果を発表しているはずである。多くの論文の中から、自分の論文をみつけさせる役目を果たするのがキーワードであり、内容を読む気にさせるのがタイトルと抄録であると言えるだろう。

そこで本論文では、こうした機能をもつキーワードが論文著者である研究者と学術雑誌編集者にどのように認識されているか、とりわけ著者が付与したキーワードが、代行検索者（図書館員）にはなじみ深いMeSHとどのような関係にあるかを考察した。

II. 投稿規定におけるキーワード指示

医学雑誌の編集者たちは、キーワードをどのように認識しているかを以下にたどった。

1978年1月、カナダのバンクーバーに19の生物医学雑誌の編集者たちが集まり、投稿規定の標準化について話し合った。ここで話し合われた標準フォームは、1979年、British Medical Journal, Lancet, Annals of Internal Medicineに「Uniform Requirements for Manuscripts Submitted to Biomedical Journals」（以下「Uniform Requirements」という）として発表された³⁾。バンクーバーに集まった編集者グループは「International Committee of Medical Journal Editors」に発展して、1981年2回目の会議を行い、翌年、第2版を発表し⁴⁾、1988年に第3版を出した⁵⁾。

「Uniform Requirements」は、標題、抄録とキー

* Isao TANIGAKI : Library, Shiga University of Medical Sciences. Seta Tsukinowa-cho, Otsu City, Shiga 520-21, Japan.

ワード、本文、謝辞、レファレンス、図表の説明の標準について説明している。

キーワードについて、1979年の第1版では、「Key (indexing) term : Below the abstract, provide and identify as such, three to 10 key words or short phrases that will assist indexers in cross-indexing your article and that may be published with the abstract. Use terms from the Medical Subject Headings list from Index Medicus Whenever possible.」としている。下線部が第2版では、「Use terms from the Medical Subject Headings list from Index Medicus *when possible.*」と変わり、第3版では、「Use terms from the Medical Subject Headings (MeSH) list of Index Medicus ; *if suitable MeSH terms are not yet available for recently introduced present terms may be used.*」となった。「Uniform Requirements」は、キーワードにMeSHを勧めている。索引の助けに、3～10個のキーワードを論文につけるように指示している。第3版では、キーワードにあてはまるMeSHがなければ、現在使われている用語でもよい、と変わった。このような編集者会議の取り決めをうけて、各々の医学雑誌はキーワードをどのように指示しているのか、投稿規定を調べてみた。

調査対象は、1988年滋賀医科大学附属図書館で所蔵していて、「Information to Authors 1980-1981」(Urban & Schworzenberg, 1980)に投稿規定がのっている雑誌、118誌である。

81年は「Information to Authors 1980-1981」を、88年は1988年発行の1号を調査したものである。118誌の中で、投稿規定にキーワードに関する指示をのせているのは、81年が46誌、88年が56誌だった。キーワードの付与を指示する規定があるが、数が明らかでない雑誌が81年では13誌、88年で12誌あった。キーワードに関する規定がある雑誌の中で、MeSHから選ぶように指示しているのは、81年が9誌で、キーワード規定のある雑誌全体の約20%になる。88年は13誌、約23%である。

著者キーワードを導入していない雑誌も多い。81年

が73誌、88年が62誌と少なくなったが、それでも全体の半分以上である。

医学雑誌の世界では、キーワードはまだ定着していない。キーワードを指示する雑誌が少ないということは、キーワード付与の歴史が浅くて考え方も新しく、キーワードの概念が確立していないからだ。

著者の付けたキーワードの意味が錯綜すれば、読者の情報利用に混乱を生じる。ことばの概念をめぐる著者、編集者、読者の間に共通の認識が生まれるシソーラスの提供が必要である。著者キーワードを導入しただけではキーワードの品質は維持できない。既存のシソーラスに適合したキーワードを使用すれば、キーワードの表現する概念の整合性が保たれ共通認識が保証される。

医学分野のシソーラスとなるとMeSHである。「Uniform Requirements」は、キーワードをMeSHから選ぶようにしている。ところが、投稿規定でMeSHをキーワードに指示している雑誌は少なかった。著者が付けるキーワードは自然語をベースにしている。一方、MeSHは統制語のシソーラスである。研究内容を表現する用語に、専門領域の共通語をあつめたシソーラスを用いるのはよいが、それでは用語が不足している。「Uniform Requirements」も3版で、MeSHのほか最新の医学用語を使ってもよい、となった。統制語のシソーラスであるMeSHで表現できない医学概念を表わすために、自然語を付け加えるように変わったのである。また、一般にシソーラスには最新の用語は追加しにくい。こういったことも理由に考えられる。

Ⅲ. 著者キーワードとMeSHの対応関係

著者が付けたキーワードは簡単な変換で統制語のディクショナリーに利用できるという意見がある。著者が付けたキーワードは統制語にどの程度置き換えられるのだろうか。この点について、医学論文に付けられた著者キーワードがMeSHにどのように変換できるか対応関係を分析してみた。

1. 変換に関する諸論

著者が付与したキーワードを統制語に変換することについて、高橋らは、日本原子力学会誌と Journal of Nuclear Science and Terminology の掲載論文を対象に、インデクシングと著者の付与したキーワードとをディスクリプターに変換して、ヒットする文献の一致度を調べた。一致度は0.75以上と推定され、著者の付与したキーワードはインデクシングにきわめて有効であると述べている。彼らは、こういう結果から、「著者の付与したキーワードを二次情報システムに適用する場合、そのシステムが自然語ベースであれば直接そのまま、あるいは特有のディスクリプターがベースであれば、単純な変換だけで容易に利用できる。」と結論している⁶⁾。また、海老沼らは、「統制語にある特定のディスクリプターに自然語を単純に変換できることが高い再現性を保証する。」と述べている⁷⁾。このように、索引作業に携わっている人の側から、著者キーワードを統制語に変換する可能性が探究されている。

一般に、自然語を統制語に変換できる条件は、両者が同義概念であることだ。同義概念であることが、変換の可能性を高くする。Montgomery らは、ディスクリプターに同義語リストを用意したら、自動索引が可能になるとしている⁸⁾。ただ、Montgomery らの研究は同義語の定義があいまいだという批判があり⁹⁾、どんな場合を同義語と認定するか、自然語と統制語の表現形式を分析する必要がある。松村らは、抄録と標題にある自然語と MeSH ディスクリプターの対応関係を分析したが、この点について、「完全一致する語は、標題では34%、抄録では40%にすぎず、同義概念の表現形式を考慮しないと情報検索にもれが生じる。」と指摘した¹⁰⁾。

本調査では、変換の方法として、シソーラス登録語の関連語関係をたどって統制語を見つける方法を考えてみた。シソーラスは、用語の意味範囲や使い方を規定しており、他の用語との関係もわかる。今回は、MeSH の注記、相互参照、ツリーストラクチャー、サブヘディングとの組合せから対応する MeSH をさがってみた。また、もう一つの方法として、著者キー

ワードと MeSH が対応するいろいろな表現形式についても考察した。

2. 調査方法

まず、Anaesthesia と Anesthesiology の1989年1号に掲載された原著27論文に付いている著者キーワードを調べた。次に、BRS の MEDLINE で、これらの論文に付与された MeSH ディスクリプターを出力した。著者のキーワードは161、MeSH ディスクリプターは337だった。

原則として、コロンやセミコロンで区切られた著者キーワードは1つの概念とした。しかし、それでは成り立たないと判断したキーワードは MeSH を参考にして複数のキーワードで1概念とした。「Kidney : failure」, 「Anesthetic techniques : epidural」などである。著者のキーワードでは、人間や動物の対象生物、年令、性別、研究援助機関のキーワードはなかった。

3. 対応関係の分類

3.1 同義語の表現形式

同義語の表現形式は、松村らや O'Connor の分類を参考にして次のように行った。

① 完全一致型

著者キーワードと MeSH とが一字一句一致する。コロン、セミコロン、カンマは無視した。英式綴りと米式綴りの違いも無視した。

② 一部型

複数語の著者キーワードが部分的に MeSH に含まれている。その逆の場合も入る。

③ 転置型

著者キーワードに対応する MeSH の語順が転置型になっている。

④ 複合型

著者キーワードが複数の MeSH で表現される。その逆、また両方が複数の場合も含む。

⑤ 異なる語

語尾変化や品詞が異なる。別の語で表現されている。

3. 2 MeSH 関連語関係

MeSH の関連語関係は次の 4 つに分けた。

① 注記型

アルファベットの注記から対応する MeSH がわかる。

② 組合せ型

著者キーワードが MeSH とサブヘディングの組合せに対応する。

③ 相互参照型

see, see related の参照で対応する MeSH がわかる。

④ 階層概念型

著者キーワードに対応する MeSH がツリーストラクチャー上で、上位か下位の概念にあたる。

4. 調査結果

161の著者キーワードのうち、同義語の表現形式で対応する MeSH があったのが38で全体の約24%だった。MeSH の関連語関係から、MeSH に変換できたのが33、約19%だった。このどちらにも属さないが、対応する MeSH があった著者キーワードは4つである。全部で75の著者キーワードが MeSH に変換できたことになる。著者キーワードと MeSH の対応関係は1対1とは限らないので、変換できた MeSH は64になった(表1)。

同義語の表現形式では、「完全一致」が一番多く17あった。以下、KWは著者キーワードを表わし、DEは MeSH を表わす。

(KW) PRILOCAINE → (DE)PRILOCAINE

(KW) RECEPTORS ; MUSCARINIC

→ (DE) RECEPTORS, MUSCARINIC

などのように化学物質が多い。また、17のうち15が単

一語同士の対応だった。

一部型は12だった。

(KW) INCREASED INTRAOCULAR PRESSURE

→ (DE) INTRAOCULAR PRESSURE

のように MeSH に形容詞がない例や、

(KW) KIDNEY : FAILURE

→ (DE) KIDNEY FAILURE, CHRONIC

と MeSH に語の付加があったり、

(KW) METABOLISM, HYPERGLYCEMIA

→ (DE) HYPERGLYCEMIA

と一部が MeSH に対応する場合があった。

転置型は1つだった。

(KW) VISUAL EVOKED POTENTIALS

→ (DE) EVOKED POTENTIALS, VISUAL

MeSH では、ある概念の医学用語は転置型にして1か所でみつかるようにしている。

複合型は4つの著者キーワードが5つの MeSH に対応していた。

(KW) MEMBRANE

(KW) PROTON PERMEABILITY

→ (DE) CELL MEMBRANE PERMEABILITY

のように、2つの著者キーワードで MeSH 1つに対応する例と、

(KW) NEUROMUSCULAR RELAXANTS

→ (DE) CURARE-LIKE AGENTS

(DE) NEUROMUSCULAR JUNCTION

と、1つの著者キーワードが MeSH 2つに対応している例があった。

異なる語では著者キーワードは4つであった。

(KW) TRANSURETHRAL RESECTION OF THE PROSTATE

→ (DE) PROSTATECTOMY

と複数語の著者キーワードが単一語の MeSH に対応している。

また、

(KW) ANESTHESIA : OBSTETRIC

表 1. MeSH に変換できた著者キーワードの数

同義語の表現形式	38
MeSH 関連語関係	33
その他	4
無変換	86
合計	161

表2. 同義語の表現形式から MeSH に変換できた著者キーワード数

完 全 一 致	1 7
一 部 型	1 2
転 置 型	1
複 合 型	4
語 が 異 な る	4
合 計	3 8

→ (DE) ANESTHESIA, OBSTETRICAL
のように語尾変化している場合である (表2)。

MeSH の関連語関係から, MeSH に対応したうち、
注記型に分類できた著者キーワードは6つであった。

(KW)SPINAL CORD ISCHEMIA

→ (DE) ISCHEMIA/complication

(DE) SPINAL CORD/blood supply

(ISCHEMIA の注記に cord with
specific organ/blood supply と
ある。)

というような例が注記型に入る。

組合せ型で MeSH に対応した著者キーワードは9
つである。

(KW) FETUS : DRUG EFFECTS

→ (DE) FETUS/drug effects

(KW) BLOOD, HEMOGLOBIN : OXYGEN

→ (DE) OXYGEN/blood supply

などである。

相互参照型は9つあった。see 参照で対応する
MeSH に変換できた著者キーワードが8つ, see re-
lated 参照が1つあった。下線の語が対応する MeSH。
以下同じである。

(KW) SUXAMETHONIUM

→SUXAMETHONIUM see SUCCINYL-
CHOLINE

(KW) SPIN LABEL

→SPIN LABELS see related ELECTRON
SPIN RESPONSE

階層型も MeSH に対応した著者キーワードは9つ
だった。上位概念の MeSH に対応したのは7つあり、

表3. 関連語関係から MeSH に変換できた著者キーワード数

注 記	6
組 合 せ 型	9
相 互 参 照	9
階 層 関 係	9
合 計	3 3

下位概念には2つだった (表3)。数字はツリーナン
バーである。

(KW) EXCERCISE (G11.427.394.380)

→ (DE) EXCERTION (G11.427.394)

(KW) SEIZURE (C10.597.742)

→ (DE) CONVULSION (C10.597.742.286)

著者が付けるキーワードは, 論文の重要な概念を表
わしているはずである。少なくとも著者はそう考えて
いる。MEDLINE には, 論文の主要な概念を表示す
るメジャーディスクリプターがある。著者キーワード
に対応した MeSH が, メジャーディスクリプターかど
うか調べてみた。著者キーワードに対応した MeSH,
64のうち, メジャーディスクリプターは48で, マイナー
ディスクリプターは16だった。

MeSH に変換できなかった著者キーワードが半数
以上あった。「LEFT VENTRICLE」といった MeSH
にない医学用語があったり, MeSH にあっても別の
概念の用語に変わっていたりしていた。また, 下位概
念の著者キーワードに MeSH がついて上位概念が対
応しないこともあった。図1は著者キーワードから作っ
た麻酔薬語のグルーピングの図で, 図2はツリースト
ラクチャーの麻酔薬の階層図である。著者キーワード
の「ANESTHETICS, INTRAVENOUS」, 「ANES-
THETICS, VOLATILE」, 「ANESTHETICS,
GASES」の下位になる麻酔薬は, ツリーストラクチャー
では, 「ANESTHETICS」の下位概念にはいって
いる。このため, MEDLINE では個々の麻酔薬に MeSH
が付与され, この3つの麻酔薬名は MEDLINE では
採用されていない。MEDLINE は最も特定できる下
位概念の用語に MeSH を付与している。

図1. 著者キーワードから作った麻酔薬階層図

図2. ツリーストラクチャーの麻酔薬の階層関係
(下線語は著者キーワードに対応した MeSH)

5. 考察

同義概念の表現形式で MeSH に変換できた著者キーワードのうち、完全一致するものは17で、全体のわずか11%である。ほとんどが単一語だった。複数語の用語で完全一致したのは、1つだけだった。複数語は表現形式を分析しなければならない。表現形式から同義語と認められたものを入れると、著者キーワード全体の38, 24%になる。著者キーワードに限らず、自然語を統制語に変換する場合は、同義語の表現形式を分析しないと、変換できる可能性は低くなる。

著者キーワードの33, 全体の19%が MeSH の関連語から変換できた。今回の調査に限って言えば、関連語からの変換も同程度の可能性があるとと言える。MeSH の関連語をたどれば、対応する MeSH にたどりつくわけで、ある意味では「原始的な変換方法」である。ただし、著者のキーワードがシソーラスの参照語や注記語になっていなければならない。

また、著者キーワードに対応した MeSH のうち、メジャーディスクリプターが48, マイナーディスクリプターが16だった。著者とインデクサーの間では論文の重要概念に対する認識が一致している。

MeSH に対応しなかった著者キーワードが半数以上あった。MeSH になかったり、あっても別概念の MeSH が採用されている。論文内容を分析して、索引した結果である。また、逆に、MeSH になかったり、MEDLINE の索引に一致しない著者キーワードが多いということである。麻酔薬の例のように、より特定の用語をキーワードにすれば、MeSH に採用されやすくなる。

IV. おわりに

自分の論文を読むために、キーワードは文字通り鍵となる。バンクーバー会議に集まった編集者たちはその重要性をアピールしてきたが、まだすべての研究者や編集者に認識されているとはいえない。

シソーラスにはタイムラグがあり、著者が自然語を好む理由も理解できるが、代行検索に従事している者にとっては、MeSH を用いた検索の方が精度が高く、

できれば MeSH をキーワードに使ってほしいところである。広く多くの人に読んでほしいと思う研究者は最新の MeSH をキーワードに採用するのが最善であると思う。

以上、著者キーワードが MeSH にどのような対応関係で変換できるか、分析を試みた。今回は、27論文の少ないキーワードで検討した結果であり、直ちに一般論にはできない。また、表現形式や MeSH 関連語からの調査で、論文の中身には立ち入っていない。さらに、多くの論文で検討する必要がある。

引用文献

- 1) 科学技術庁科学技術振興局編 SISTハンドブック. 日本科学技術情報センター, 1988: 252, 267.
- 2) CBE Style Manual Committee, eds. Council of Biology Editors style manual. 4th ed. Arlington: Council of Biology Editors, 1978: 30.
- 3) International Steering Committee. Uniform requirements for manuscripts submitted to biomedical journals. Ann Intern Med 1979; 90 (1): 95-9.
- 4) International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. Ann Intern Med 1982; 96 (6, pt. 1): 766-71.
- 5) International Committee of Medical Journal Editors. Uniform requirements for manu-

scripts submitted to biomedical journals. Ann Intern Med 1988; 108 (2): 258-65.

- 6) 高橋智子他 学術誌へのキーワード付与制の要請. 第18回情報科学技術研究会論文集. 日本科学技術情報センター, 1982: 251-7.
- 7) Ebinuma Y, et al. Promotion of keywords assignment to scientific literature by contributors. Int Forum Inf Doc 1983; 8 (3): 16-20.
- 8) Montgomery R, Swanson DR. Machinelike indexing by people. Am Doc 1962; 13: 359-66.
- 9) O'Connor J. Correlation of indexing headings and title words in three medical indexing systems. Am Doc 1964; 15: 96-104.
- 10) 松村多美子, 緑川信之 情報検索における統制語と自然語. 図書館情報大学研究報告 1989; 8 (1): 49-63.

参考文献

- 1) 東海安興 オンライン情報検索における医学雑誌付与キーワードの役割. 医学図書館 1985: 32 (3): 250-65.
- 2) 山村良彦 科学技術情報利用におけるキーワードの役割. 金属表面技術 1987; 38 (7): 303-6.
- 3) 山崎昶他 学術雑誌の著者付与キーワードの検討. 電気通信大学学報 1987; 38 (理工学編); 25-31.