

BSH と NDLSH の統合的活用の可能性 両者の相違の積極的活用法

谷口 祥一（慶應義塾大学文学部） taniguchi@z2.keio.jp

1. はじめに

BSH(基本件名標目表)第4版と NDLSH(国立国会図書館件名標目表)は、わが国の代表的な件名標目表である。NDLSHの個々の件名において示された、対応する BSH 件名の指示によって、多対多(ただし、殆どは1対1)の対応づけで両者は結びつけられている。本研究は、両件名標目表の統合的活用に向けて、両者の相違の状況を確認し、それを積極的に活かす活用法(具体的には両者を統合化した表示法)を検討する。なお、両者の比較を行った先行研究には、北らによるもの¹⁾、村上によるもの²⁾がある。

ちなみに、BSH、NDLSHともに上位語・下位語の関係による階層構造をもつが、複数の異なる上位語をもつ場合などがあり、全体としてはグラフ構造をなす。

2. 両件名標目表の対応づけの状況と孤立件名の状況

BSHは機械可読形式データを入力し³⁾、NDLSHはWeb NDL Authoritiesから一括ダウンロード用データを2013年8月7日時点で取得した⁴⁾。両者ともに、細目付きの件名は一部を除いて含まれていない。すなわち、実際の書誌レコードに出現する細目付きの件名の大半は、今回の照合からは除かれている。また、取得したデータに含まれている参照や細目のみのものを、照合には適しないと判断し除いた。その結果、BSHは8,034件名、NDLSHは19,195件名であった。このうち、上位語も下位語ももたない孤立した件名が、BSHでは433(5.4%)、NDLSHでは4,626(24.1%)含まれていた。

そこで、BSH、NDLSHそれぞれをベースにしたとき、a)NDLSHの個々の件名において示された、対応するBSH件名の指示による対応づけの有無と、b)上位下位関係をもたない孤立した件名か否かとの2つの区分に分けてクロス集計し全体の状況を

まとめた(表1)。

BSHをベースに集計したとき、NDLSHとの対応づけがある件名は7,016(87.3%)、対応づけのない件名が1,018(12.7%)となった。さらに前者7,016は、上位下位関係をもたない孤立件名241と上位下位関係のいずれか(または両方)をもつ6,775とに分かれる。同様に、後者1,018は、上位下位関係をもたない孤立件名192と上位下位関係のいずれか(または両方)をもつ826とに分かれる。

NDLSHについても同様に集計した結果、BSHとの対応づけがある件名は6,861(35.7%)、対応づけのない件名が12,334(64.3%)と、後者の割合が大きいことが確認された。併せて、孤立件名が相当数含まれること(対応づけがないもののうち、3,343)を確認した。

表1. 対応づけの有無と孤立件名とのクロス集計

BSHをベースにした集計			
	非孤立 件名	孤立 件名	計
NDLSHとの対応づけあり	6,775 (84.3%)	241 (3.0%)	7,016 (87.3%)
対応づけなし	826 (10.3%)	192 (2.4%)	1,018 (12.7%)
計	7,601 (94.6%)	433 (5.4%)	8,034 (100%)

NDLSHをベースにした集計			
	非孤立 件名	孤立 件名	計
BSHとの対応づけあり	5,578 (29.1%)	1,283 (6.7%)	6,861 (35.7%)
対応づけなし	8,991 (46.8%)	3,343 (17.4%)	12,334 (64.3%)
計	14,569 (75.9%)	4,626 (24.1%)	19,195 (100%)

表 2. BSH と NDLSH の階層構造の照合結果

A. BSH をベースにした照合: BSH → NDLSH			
件名の出現総数: 10,421		(件名の異なり数: 7,601; 孤立件名は除外)	
最上位語(248)ごとに照合	①基本的照合	②3階層の範囲での照合	③上位語の継承を組み入れた照合
上位下位関係の不一致	6,164 59.1%	6,314 60.6%	6,473 62.1%
		①との相違 +150 (内訳: -118, +268)	①との相違 +309 (内訳: -0, +309) ②との相違 +159 (内訳: -0, +159)
B. NDLSH をベースにした照合: NDLSH → BSH			
件名の出現総数: 19,643		(件名の異なり数: 14,569; 孤立件名は除外)	
最上位語(1,533)ごとに照合	①基本的照合	②3階層の範囲での照合	③上位語の継承を組み入れた照合
上位下位関係の不一致	2,515 12.8%	2,850 14.5%	3,295 16.8%
		①との相違 +335 (内訳: -302, +637)	①との相違 +780 (内訳: -0, +780) ②との相違 +445 (内訳: -0, +445)

なお、BSH と NDLSH との対応づけのある件名数が、いずれをベースに集計するかによって異なるのは、両件名間の多対多の対応づけに起因する。採用した NDLSH と BSH の対応づけの組は 7,021 であり、ここから BSH から NDLSH への対応づけは平均 1.001 (=7,021÷7,016)、NDLSH から BSH への対応づけは平均 1.023 (=7,021÷6,861) となる。また、孤立した件名を除いて計算すれば、前者は平均 1.036、後者は平均 1.259 となる。

3. 階層構造の一致の程度

次に、両件名標目表の対応づけがなされた部分について、BSH 件名の上位語・下位語関係に対して、対応づけられた NDLSH 件名間で同じく上位下位関係にあるのかプログラムで照合し、不一致の箇所を検出した。同様に、NDLSH 件名をベースにした階層構造の照合を行い、不一致の程度を明らかにした。これらの結果を表 2 に示してある。

照合処理には、以下の方式を採用した。

1) BSH 件名と NDLSH 件名との間には、数は少ないが、多対多の対応関係がある。そのため、対応づけられた複数ある件名のうち、いずれか一つが上位下位関係にあれば合致したものとする。

2) すべての BSH/NDLSH 件名に対応づけがあるわけではない。そのため、対応する件名をもたない箇所では不一致は発生しない(いかなる件名とも合致する)とみなすこととした。

まず、BSH をベースとした照合(表 2 の A)では、最上位語(248 件名)ごとに照合を行った結果、複数の上位語をもつケースがあるため総出現件名数 10,421(件名の異なり数 7,601)となる。それぞれの件名の上位語・下位語関係に対して、個々の件名に対応づけられた NDLSH 件名が同じく上位下位関係(または同一件名への対応づけ)にあるのか照合し、上位下位関係にない不一致の箇所を検出した。その結果、不一致の箇所は 6,164、全体の 59.1%であった(表 2 の A.①)。なお、この照合においては、BSH の各最上語から辿った処理とし

ており、BSHの孤立件名は除外されるが、NDLSH側は孤立件名であるか否かは問わず、その両方が対象となる。

照合の範囲を直近の上位下位という2階層から、3階層に広げ照合したときには、不一致は6,314(全体の60.6%)に増加した(表2のA.②)。これは、上位下位関係にある件名間では対応するNDLSH件名が上位下位関係(または同一件名)になくとも、そのさらに上位語に対応するNDLSH件名とは上位下位関係(または同一件名)にある場合を合致するとみなした照合である。不一致の数が増加したのは、こうした照合によって、件名をもたない箇所では不一致とはみなさないという適用事例が減少した結果である。表に示した通り、実際には階層構造の合致箇所が118増加したが、逆に不一致箇所が268増加している。

併せて、対応する件名をもたないときには、上位の件名に対応づけられている件名を継承して擬似的に付与する方式を試みた。その結果が表2のA.③であり、すべてのBSH件名で対応するNDLSH件名をもつことになり、そのため不一致の箇所が大幅に増加している(6,473、62.1%)。

表2のBには、NDLSHをベースにした階層構造の照合処理の結果を示した。最上位語(1,533件名)ごとに照合を行った結果、複数の上位語をもつケースがあるため総出現件名数19,643(件名の異なり数14,569)となる。階層構造の不一致の箇所が2,515(全体の12.8%)と、BSHをベースにした場合よりも大幅に少ないのは、対応づけられたBSH件名をもつものが限られており、対応する件名をもたない箇所では不一致は発生しないとみなした結果である。3階層に照合の範囲を広げた場合(B.②)、対応する件名をもたないときに上位語の継承を組み入れた場合(B.③)それぞれにおける不一致箇所の増加傾向は、BSHの場合と相似する。

4. 統合した表示の試み

両者の階層構造の不一致を積極的に活用し、それぞれの件名標目表が用語(名辞)および上位下位関係の追加によって、より

豊かな統制語彙として活用できる可能性を示すべく、両者を統合した表示法(個々の件名に対する表示および階層構造表示)を複数検討した。図1には、BSH件名の表示において、その対応するNDLSH件名の上位語・下位語を併せて表示する方法を適用した事例を、図2にそれらNDLSH件名の上位語・下位語に代えて、それぞれに対応するBSH件名を元々の上位語・下位語に加え表示する方法を適用した事例を示した。

また、図3にはBSHの階層構造表示に、対応するNDLSH件名とその上位語を加えて表示した例を示した。なお、"(A)"は当該NDLSH件名がBSH上位語に対応するNDLSH件名と上位下位関係を構成していない、すなわち階層構造が不一致であることを示している。図4は、図2で採用した方法から得られる階層構造の表示であり、件名「教育工学」とその下位語、および「教科書」とその下位語が、BSHによる元々の階層構造中の位置づけに追加されたかたちで出現している(破線による囲み部分が該当)。これは同一用語に対する、他の可能な解釈(用語のスコップと他の語との関係づけ)を表すものと捉えることができる。また、こうした統合化により、対応づけがあるが孤立した件名については、対応づけられた先の件名がもつ上位下位関係を介して擬似的な階層関係をもつ可能性が開ける。他方、孤立した件名の解消に向けた作業(人手による作業を含む)の必要性も再認識された。

引用文献

- 1) 北克一, 芝勝徳. 我が国における主題索引ツール統合の試み. TP&D フォーラムシリーズ. 3, 1994, p. 6-31.
- 2) 村上幸二. BSH(基本件名標目表)とNDLSH(国立国会図書館件名標目表)の計量的比較分析. 紀要(奈良文化女子短期大学). 40, 2009, p.135-147.
- 3) 日本図書館協会. BSH4-Computer File(基本件名標目表第4版機械可読データファイル). 2002.
- 4) 国立国会図書館. Web NDL Authorities. 一括ダウンロード用ファイル. <http://iss.ndl.go.jp/ndla/download/>(参照2013-08-07)

ID:BSH400212900 (NDSLH00567159)
 HD:教育学
 YM:キョウイク コウガク
 UF:Educational technology
 TYPE:10
 TT:教育学 BSH400212100
 BT:学習指導 BSH400127400
 BT:教育学 NDSLH00567145
 NT:教育機器 BSH400212500
 NT:C A I BSH401093200
 NDC8:375
 NDC9:375.11
 NDLC:FC51
 LCSH:sh85041150

図1. BSH 件名に NDSLH 件名を統合化した表示(1)

ID:BSH400212900 (NDSLH00567159)
 HD:教育学
 YM:キョウイク コウガク
 UF:Educational technology
 TYPE:10
 TT:教育学 BSH400212100
 BT:学習指導 BSH400127400
 BT:教育学 BSH400212100
 NT:教育機器 BSH400212500
 NT:C A I BSH401093200
 NDC8:375
 NDC9:375.11
 NDLC:FC51
 LCSH:sh85041150

図2. BSH 件名に NDSLH 件名を統合化した表示(2)

教育学：

- ・ 教育指導 [BT:教育学]:
- ・ ・ 学習指導 [BT:教育指導]:
- ・ ・ ・ 学業不振 [BT:学習指導]:
- ・ ・ ・ 学習法 [BT:学習指導]:
- ・ ・ ・ 技術・家庭科 [BT:学習指導; 家庭科学]:
- ・ ・ ・ 教育課程 [BT:学習指導]:
- ・ ・ ・ 教育学 [BT:学習指導]:
- ・ ・ ・ ・ 教育機器 [BT:教育学]:
- ・ ・ ・ ・ C A I [BT:教育学]:
- ・ ・ ・ 教科書 [BT:学習指導]:
- ・ ・ ・ ・ 往来物 [BT:教科書; 日本一教育]:
- ・ ・ ・ ・ 教科書検定 [BT:教科書]:
- ・ ・ ・ 教材 [BT:学習指導]:

図3. BSH 件名に NDSLH 件名を統合化した階層構造表示(1)

教育学 [BT なし]

- (A) ガイダンス (教育) [BT なし]
- 学習指導 [BT: 教授法; ガイダンス(教育)]
- (A) 学業不振児 [BT: 問題児]
- (A) 学習法 [BT なし]
- (A) 技術・家庭科 [BT: 教科]
- (A) 教育課程; 一貫教育 [BT: 教育課程]
- (A) 教育学 [BT: 教育学]
- (A) ティーチングマシン [BT: 教具]
- (A) コンピュータ教育 [BT: 情報教育]
- (A) 教科書 [BT: 教材]
- 往来物 [BT: 教科書]
- (A) 教科書検定・採択制度 [BT なし]
- (A) 教材 [BT なし]

教育学：

- ・ 教育学： (BT: 学習指導; *教育学)
- ・ ・ 教育機器： (BT: 教育学)
- ・ ・ ・ C A I： (BT: 教育学)
- ・ 教育指導： (BT: 教育学)
- ・ ・ 学習指導： (BT: 教育指導)
- ・ ・ ・ 学業不振： (BT: 学習指導)
- ・ ・ ・ 学習法： (BT: 学習指導)
- ・ ・ ・ 技術・家庭科： (BT: 学習指導; 家庭科学)
- ・ ・ ・ 教育課程： (BT: 学習指導)
- ・ ・ ・ ・ 教育学： (BT: 学習指導; *教育学)
- ・ ・ ・ ・ 教育機器： (BT: 教育学)
- ・ ・ ・ ・ C A I： (BT: 教育学)

- ・ ・ ・ 教科書： (BT: 学習指導; *教材)
- ・ ・ ・ ・ 往来物： (BT: 日本一教育; 教科書)
- ・ ・ ・ ・ 教科書検定：
- ・ ・ ・ 教材： (BT: 学習指導)
- ・ ・ ・ ・ 教科書： (BT: 学習指導; *教材)
- ・ ・ ・ ・ ・ 往来物： (BT: 日本一教育; 教科書)
- ・ ・ ・ ・ ・ 教科書検定： (BT: 教科書)

図4. BSH 件名に NDSLH 件名を統合化した階層構造表示(2)