

Annals of Library and Information Studies Vol. 70, June 2023, pp. 66-73 DOI: 10.56042/alis.v70i2.1045



# The formation of the subject Cell Biology from the perspective of linguistics: An analytical study

Debabrata Maity<sup>a</sup> and Bidyarthi Dutta<sup>b</sup>

<sup>a</sup>Central Library, Khejuri College, Baratala, Khejuri, West Bengal-721431, Email: maitydebabrata6@gmail.com, <sup>b</sup>Department of Library and Information Science, Vidyasagar University, Midnapore, West Bengal-721102, Email: bidyarthi.bhaswati@gmail.com

Received: 30 April 2023; accepted: 25 June 2023

Cell Biology (CB) deals with the structure and function of cells, and it is an important branch of biology. A total of 1918 keywords were extracted from the titles, abstracts, and objects' captions of 25 top-cited research articles in the domain of CB, and they were linguistically analyzed. It is found that CB is closely interrelated with Chemical Sciences, Medical Sciences, and Biological Sciences, and the mode of formation of CB is distillation of Kind 2—expanded.

Keywords: Cell Biology; Linguistic analysis; Object analysis

#### Introduction

The discipline of Cell Biology (CB) deals the study of the structure and function of the cell-the fundamental unit of life. The concept of cell theory mainly emerged from a definition of the cell, given by Schleiden and Schwann. According to them, "all living creatures, both simple and complex, are made out of one or more cells, and the cell is the structural and functional unit of life"<sup>1</sup>. CB has various branches such as Cytoecology, Cytochemistry, Cytopathology, Cell Physiology, Cytotaxonomy, Cytogenetics, and so forth. Shah<sup>2</sup> mentioned briefly about the top sixteen techniques used in CB in present days, which are immunofluorescence microscopy, iconexchange chromatography, affinity chromatography, partition and adsorption chromatography, gel filtration chromatography, radioactive tracer technique, (RIA), enzyme immunoassay, radioimmunoassay spectroscopy, nuclear magnetic resonance spectroscopy (NMR), optical rotatory dispersion (ORD) and circular dichrois (CD), infra-red spectrophotometry, atomic absorption/ flame spectrometry, flow cytometer, applications of flow cytometer and cell sorter, and noninvasive scanning of soft tissues.

Library and Information Science (LIS) is a blending of two disciplines, that is, Library Science and Information Science, and it "can be studied using techniques from the humanities, social sciences, and pure science"<sup>3</sup>. One of the basic thrust areas of LIS research is knowledge organization, which includes core activities like classification, document description and indexing. These activities must be executed by the subject specialists in libraries, archival centers and databases for better retrieval and dissemination of necessary information or documents. In libraries and related centers, any document is classified by analyzing its subject content. Specialists use library classification schemes like *Dewey Decimal Classification, Universal Decimal Classification, Colon Classification, Library of Congress Classification* and so forth to fulfill the purpose.

To develop any classification scheme, the knowledge about "the mode of formation and the structure of the subjects in the universe of subjects and of the isolate ideas in the universe of isolate ideas" is mandatory<sup>4</sup>. Any subject can be analyzed into a set of components expressed by a set of terms. It is possible to draw a variety of relations between any two components of a subject or an isolate, and the type of relationship among the components of a subject identifies the mode of formation of the subject<sup>5</sup>.

There are several modes of subject formation in the universe of subjects in LIS study, but the relationships among the components of the subjects are not yet analyzed linguistically. The present study describes the mode of formation of CB through linguistic analysis.

#### **Review of literature**

The credit for starting the concept of modes of formation of subjects goes to Ranganathan, as in 1948 he introduced a paper entitled 'Development and structure of the universe of subjects' in the postgraduate library science curriculum of the University of Delhi<sup>6</sup>. Later, the concept was continued to develop mainly by the associates and schoolmen of Ranganathan at the DRTC in Bangalore and a few from abroad. Ranganathan, in 1950, pointed out that "subjects in the universe of knowledge can be formed by means of four modes of formation; these are loose assemblage, lamination, dissection, and denudation"<sup>5</sup>.

Neelameghan<sup>7</sup> drew attention to the formation of primary basic subjects by fission. He also discussed the arrangement of basic subjects<sup>8</sup>. Gopinath and Seetharama<sup>9</sup> defined and explained seven modes of formation of subjects including loose assemblage. lamination, fission, fusion, distillation, agglomeration, and cluster (121-135). In 1993, McGarry<sup>10</sup>, in his book The Changing Context of Information, mentioned two separate modes of subject formationprocreation and annexation. Psycholinguistics, Sociolinguistics, and Neuro-Linguistics are examples of the subjects formed by the procreation mode, whereas Commercial Geography, Political Geography, and Medical Geography are examples of the subjects formed by the annexation mode.

Satija, Madalli, and Dutta<sup>11</sup> described the analogical and instrument-based subject formation modes. According to them, "some subjects find parallels in other disciplines." Social Darwinism, Social Physics, and Social Entropy are some subjects formed by the analogical mode. Secondly, the subjects "grown into full discipline by gathering around a machine or device" are instrument-based subjects. Dutta and Dutta<sup>12,13</sup> explained how the subject formation process could be analyzed linguistically. The subject formation process was compared with universal linguistic forms by Neelameghan<sup>14</sup>. He showed that the formation of a generic framework for structuring subjects has a parallel in the search for universal linguistic forms such as that expounded in the works of distinguished linguists like Chomsky, Fodor, Katz *et al.* Neelameghan<sup>15</sup> also developed a generalized facet structure of subjects Fig. 1.

Microscopy, Microbiology (both developed depending on the machine microscope), and subjects based on computer or mobile-related technologies are examples of instrument-based subject formation. All the modes described above are not interpreted by linguistic analysis, and nowhere in the above said literature CB is drawn as an example of any mode of formation of subject. In this paper, an attempt has been made to bridge these gaps.

#### Linguistic analysis

From the perspective of structural aspect, a word may be categorized in three groups: 1) Root word 2) Stem word and 3) Compound word. The compound words can be in four categories<sup>16</sup>, i.e. (1) exocentric compound (2) endocentric compound (3) copulative compound and (4) appositional compound. The primary lexical unit of a word is known as root word. It carries the most significant aspects of semantic content, which is not further divisible into smaller constituents. The stem indicates a part of word that may be either prefixed or suffixed to a root word to add a new form to its meaning.

For instance, if 'child' is a root word, then it has several stem words like 'childish', 'childhood', 'children' etc. Here, the central theme of all stem words is same as the root word, but the forms are a bit different. The attachment of more than one primary lexical units or roots creates compound words. The compounds may also contain stems in different forms. Now, let us say a compound word that is formed from two different primary lexical units or roots, viz. A & B. Let us denote the said compound word, i.e., the combination of A & B as (A + B).

Suppose the meaning of A is M, and that of B is N. Then following four cases may arrive:

Case 1: (A + B) = N, where  $N \equiv B$ , i.e. (A + B)indicates a particular kind of B, which means N only, but not (M + N). Examples of such kind of compounding are darkroom, football, roommate, birthday, bedtime etc. For instance, the word darkroom

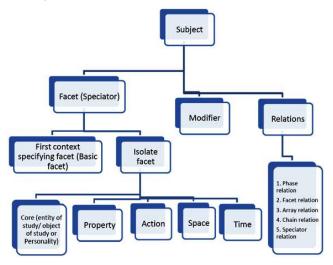


Fig. 1 — Generalized facet structure by Neelameghan

is formed from two root words dark and room (may be reckoned as A and B respectively), but the complete word darkroom indicates a special type of room, but not any kind of darkness. This kind of compounding is known as 'endocentric compounding'. The meaning is contained here within the components.

Case 2: (A + B) = P, where  $P \neq M$ ,  $P \neq N$ ,  $P \neq (M + N)$  or any other combination of M & N, i.e. (A + B) indicates a special kind of an unexpressed semantic head, which neither expresses A, nor B, but a completely new theme. This kind of compounding is known as 'exocentric compounding'. Examples of such kind of compounding are breakfast, pickpocket, paperback, egghead, ladyfinger etc. For instance, the word breakfast is formed from two root words break and fast (may be reckoned as A and B respectively), but the complete word breakfast indicates a completely new concept (a meal) that has no conceptual vicinity with either break or fast. Here the meaning is not contained within the components.

Case 3: (A + B) = (M + N), i.e. (A + B) denotes the sum of what A and B denotes. This kind of compounding is known as 'copulative compounding'. Examples of such kind of compounding are bookstore, eggshell, sleepwalk, eyelid, newspaper etc. For instance, the word bookstore is formed from two root words book and store (may be reckoned as A and B respectively), while the complete word bookstore also indicates nothing but the storehouse of book. Here also the meaning is contained within the components.

Case 4: (A + B) = R, where R = M = N, i.e. (A + B)denotes different descriptions of the same referent. This kind of compounding is known as 'appositional compounding'. Examples of such kind of compounding are managing-director, founder member, player-coach, student-worker, singer-actor etc. For instance, the word managing-director is formed from two root words managing and director (may be reckoned as A and B respectively), while the complete word managingdirector indicates one person (not two) who simultaneously acts both as manager and as director.

In this study, the selected 1918 keywords were analyzed to find the root words and then the compound keywords were categorized as mentioned in the methodology.

## **Objectives of the study**

The present study includes the following objectives:

 To identify the nature of compound words used in CB;

- To see the nature of root words used in the keywords of CB;
- To observe how the root words used in CB are scattered by origin in different disciplines;
- To find out the core disciplines involved in the growth of CB; and
- To identify the mode of formation of CB.

## Methodology

At first, 25 top-cited research articles (except review articles, reports, commentaries, etc.) in the field of CB were collected as the sample size for the present study from the Web of Science database. During the search process, the term "Cell Biology" was put within a double inverted comma in the search box, and the time span of 1980–2014 was fixed. The criterion was set to show the retrieved results as per relevance. An article having at least one *object* (i.e., any of the non-textual elements like tables, diagrams, figures, charts, photographs, maps, etc.) with a proper caption was considered as a sample. A total of 168 objects were found in the sample articles. Then the following steps were done:

• The captions of the objects of the articles were analyzed to cull out keywords. Though a single keyword may appear more than once in a particular object's caption within a single article, it was noted only one time. The same keyword was noted a second time if it occurred once or more than once in another object's caption in the same article; noted a third time if it occurred again in an object's caption except the previously analyzed first and second objects within the same article; and the process continued. In all, 1593 keywords were derived from the captions of the objects of the articles, and additionally, 325 keywords were derived by analyzing the titles and abstracts of the articles, figuring a total of 1918.

• Acronyms of keywords were expanded (e.g., Fluorescence-Activated Cell Sorting for FACS). Numerical figures (including 1, 2, 3, I, II, III, etc.), qualifiers (including low, high, derived, etc.), jargons (including A23187, B-100, M71/2, etc.), function words (including and, of, for, in, etc.), and nontechnical words (including induced, activity, associated, etc.) were removed from the keywords to obtain focal words (may be called focal terms).

• The focal words of the keywords were separately arranged together with their frequency of occurrence. Each of the focal words was analyzed further to reach its root(s) and/or root word(s). In the present study, the term "root word" has been used to mean both root and root words. Here we found 832 root words. During the root word analysis, compound words (a total of 278) were identified and listed separately with an indication of their semantic type.

• The obtained root words were listed, and they were further searched in some tertiary sources of information (see Notes) to find out their subjectspecific meaning and original context, and tabulation work was done for further analysis and interpretation. A root word may occur in more than one subject or discipline with different contextual meanings, but the emphasis was given only on that subject-specific meaning which is more contextual to use in CB, and the root word was tabulated under that specific subject. For example, the root word "cell" is used in more than one discipline with different meanings-in General Biology, "the basic structural and functional unit of living organisms; ... " in General Physics, "a device for converting chemical energy into electrical energy; . . ." in Chemistry, "short for electrolytic cell; ..." in Ecclesiastical Terms, "a small religious house dependent upon a larger one"<sup>17</sup>, and so forth. But it is easily judged that the meaning of the word "cell" in the context of General Biology is used in the field of CB. So the root word "cell" was counted under the subject General Biology. Secondly, if it was seen that a root word has contextually related meaning from more than one discipline or subject, then the root word was counted under the discipline or subject that is old by origin. In the present study, the terms "subject" and "discipline" have been used in the same sense.

### Analysis

## Nature of compound words

The compound words found in the present study were categorized further according to their semantic type, and the result is presented in Table 1. According to the study, 74.82% of compound words are endocentric, 14.03% are exocentric, and 11.15% are copulative.

Depending on the subjects of the component root words and the position of the component root words

Table 1 — Semantic classification of the compound words used in CB					
Sl. No.	Semantic Type	Frequency	Percentage		
1	Endocentric	208	74.82		
2	Exocentric	39	14.03		
3	Copulative	31	11.15		
4	Total	278	100.00		

in a compound word, the subject combination pattern of each compound word was identified. Though it was found that the 278 compound words are formed of 113 distinct subject combination patterns, only the top 20 patterns are presented in detail in Table 2. These top 20 patterns are responsible for forming 155 (55.76%) compound words. The combination of Common Root Word and Common Root Word produces 33 (11.87%) compound words and ranks at the top, followed by the patterns formed of General Biology and Common Root Word (19, 6.83%), Common Root Word and General Biology (13, 4.68%), Common Root Word and Anatomy (8, 2.88%), and so on.

Subjects of the root words forming the compound words were analyzed further, and the result is given in

Table	2 — Top twenty subject combinati compound words used		forming the
Sl. No.	Subject Combination Pattern	Frequency	Percentage
1	Common Root Word-Common Root Word	33	11.87
2	General Biology-Common Root Word	19	6.83
3	Common Root Word-General Biology	13	4.68
4	Common Root Word-Anatomy	8	2.88
5	Common Root Word-General Chemistry	7	2.52
6	General Chemistry-Biochemistry	7	2.52
7	General Medicine-Common Root Word	7	2.52
8	Common Root Word-Genetics	6	2.16
9	Organic Chemistry-Organic Chemistry	6	2.16
10	Biochemistry-Biochemistry	5	1.80
11	Common Root Word- Biochemistry	5	1.80
12	Common Root Word-General Medicine	5	1.80
13	Common Root Word-Physiology	5	1.80
14	General Chemistry-General Chemistry	5	1.80
15	Biochemistry-Common Root Word-Biochemistry	4	1.44
16	Biochemistry-General Chemistry	4	1.44
17	General Biology-General Chemistry	4	1.44
18	Organic Chemistry- Biochemistry	4	1.44
19	Physiology-Common Root Word	4	1.44
20	Physiology-General Biology	4	1.44
21	Others (=93)	123	44.24
Total S	ubject Combination Pattern = 113	278	100.00

Table 3. It is observed from the table that root words from 32 distinct subject categories occur 597 times to form the total (i.e., 278) compound words. With a frequency of 203 (34.00% of the total), the category Common Root Word occurs the most times, followed by Biochemistry (71, 11.89%), General Biology (67, 11.22%), General Chemistry (60, 10.05%), Organic Chemistry (37, 6.20%), and so forth. The time of occurrence for each of the 11 subjects, including Electromagnetism, Electronics, General Science, Geology, Horticulture, and so forth, is very low (0.17%).

### Nature of root words

The 832 root words as mentioned in the methodology section were broadly divided into two-

Table 3 — Finding the subject from compound word analysis					
Sl. No	o. Subjects	Times of occurrence	Percentage		
1	Common Root Word	203	34.00		
2	Biochemistry	71	11.89		
3	General Biology	67	11.22		
4	General Chemistry	60	10.05		
5	Organic Chemistry	37	6.20		
6	General Medicine	30	5.03		
7	Anatomy	28	4.69		
8	Physiology	22	3.69		
9	Genetics	11	1.84		
10	Zoology	11	1.84		
11	Pathology	9	1.51		
12	General Physics	7	1.17		
13	Microbiology	6	1.01		
14	Oncology	5	0.84		
15	Cell Biology	4	0.67		
16	Immunology	3	0.50		
17	Linguistics	3	0.50		
18	Mathematics	3	0.50		
19	Ambiguous Terms	2	0.34		
20	Astronomy	2	0.34		
21	Botany	2	0.34		
22	Electromagnetism	1	0.17		
23	Electronics	1	0.17		
24	General Science	1	0.17		
25	Geology	1	0.17		
26	Horticulture	1	0.17		
27	Inorganic Chemistry	1	0.17		
28	Optics	1	0.17		
29	Pharmacology	1	0.17		
30	Philosophy	1	0.17		
31	Physical Chemistry	1	0.17		
32	Surgery	1	0.17		
Total	= 32	597	100.00		

Subject-Specific Root Words (SRWs) (581) and Common Root Words (CRWs) (251). Though the non-technical words and function words like a, an, the, and so forth were omitted before analyzing the focal words, it was found that the focal words used in CB are composed of 69.83% SRWs and 30.17% CRWs.

#### **Distribution of SRWs in different disciplines**

The SRWs used in CB are distributed according to their subject of origin in Table 4. From the table, it is found that 14.11% of the total SRWs come from Biochemistry, followed by Anatomy (13.60%), Organic Chemistry (9.81%), General Chemistry (9.47%), General Biology (7.06%), General Medicine (5.34%), Zoology (4.99%), Pathology (4.65%), and so forth. Almost 70% of SRWs have their origin within these eight top-ranked subjects. So, these subjects may be identified as core or nuclear subjects (i.e., subjects of zone 1 as per the terminology used in Bradford's Law) responsible for the development of CB. Except for the mentioned 70%, the next 20% of SRWs (i.e., under the cumulative frequency of >70%-90%) have their origin in 11 subjects (from Genetics to Pharmacology in Table 4). These subjects may be termed as subjects of zone 2 or allied subjects of CB. The last 10% of SRWs originated from 34 subjects (i.e., subjects of zone 3 or alien subjects) with a very low frequency of occurrence. These three zones can be seen in Fig. 2.

It is found from the study that 2.75% of SRWs have their origin within the subject CB itself. SRWs having the source of origin in CB or Cytology are kept together under CB in the present study. Eight root words (1.38%) were kept under the class of *Ambiguous Terms* because the contextual meaning of these root words was unavailable both in general and subject-specific dictionaries available to us. So it

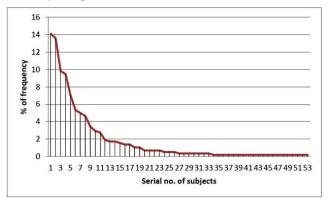


Fig. 2 — Subject-wise distribution of SRWs

Table 4 — Subject-wise distribution of SRWs						
Sl. No.	Rank No.	Subject	Frequency of SRWs	% of Frequency	Cumulative Frequency	% of Cumulative Frequency
1	1	Biochemistry	82	14.11	82	14.11
	2	Anatomy	79	13.60	161	27.71
	3	Organic Chemistry	57	9.81	218	37.52
ļ	4	General Chemistry	55	9.47	273	46.99
5	5	General Biology	41	7.06	314	54.04
5	6	General Medicine	31	5.34	345	59.38
, 7		Zoology	29	4.99	374	64.37
3	7 8		29	4.65	401	69.02
	8 9	Pathology				
)		Genetics	20	3.44	421	72.46
0	10	Physiology	17	2.93	438	75.39
1	11	Cell Biology	16	2.75	454	78.14
2	12	Microbiology	11	1.89	465	80.03
3	13	General Physics	10	1.72	475	81.76
4	13	Inorganic Chemistry	10	1.72	485	83.48
5	14	Botany	9	1.55	494	85.03
6	15	Mathematics	8	1.38	502	86.40
7	15	Ambiguous Terms	8	1.38	510	87.78
8	16	Immunology	6	1.03	516	88.81
9	16	Pharmacology	6	1.03	522	89.85
20	17	Electromagnetism	4	0.69	526	90.53
1	17	Hematology	4	0.69	530	91.22
2	17	Linguistics	4	0.69	534	91.91
23	17	Oncology	4	0.69	538	92.60
24	18	Atomic physics	3	0.52	541	93.12
25	18	Bacteriology	3	0.52	544	93.63
26	18	Electrical Engineering	3	0.52	547	94.15
27	19	Anthropology & Ethnology	2	0.34	549	94.49
28	19	Astronomy	2	0.34	551	94.84
.0 29	19	Geography	2	0.34	553	95.18
30	19	Optics	2	0.34	555	95.52
31	19	Philosophy	2	0.34	557	95.87
32	19	Surgery	2	0.34	559	96.21
		Telecommunication				
33	19		2	0.34	561	96.56
34	20	Aeronautics	1	0.17	562	96.73
35	20	Automotive Engineering	1	0.17	563	96.90
6	20	Dentistry	1	0.17	564	97.07
57	20	Ecology	1	0.17	565	97.25
8	20	Economics	1	0.17	566	97.42
9	20	Ethnography	1	0.17	567	97.59
0	20	Geology	1	0.17	568	97.76
1	20	Geometry	1	0.17	569	97.93
2	20	Horticulture	1	0.17	570	98.11
3	20	Molecular Biology	1	0.17	571	98.28
4	20	Neurology	1	0.17	572	98.45
5	20	Nuclear physics	1	0.17	573	98.62
6	20	Parasitology	1	0.17	574	98.80
7	20	Phylogeny	1	0.17	575	98.97
8	20	Physical Chemistry	1	0.17	576	99.14
9	20	Psychology	1	0.17	577	99.31
50	20	Textile	1	0.17	578	99.48
51	20	Toxicology	1	0.17	579	99.66
52	20	Trade Name	1	0.17	580	99.83
52	20	Video Technology	1	0.17	581	100.00

	Table 5—Dist	tribution of SRWs a	according to broad subject c	ategory	
Broad subject division	n Subjects under broad division	Frequency	Total no of frequency (%)	Cumulative frequency	% of cumulative frequency
	Biochemistry	82			
Chemical Sciences	Organic Chemistry	57			
	General Chemistry	55			
	Inorganic Chemistry	10	205 (25.20)		25.20
	Physical Chemistry	1	205 (35.28)	205	35.28
	Anatomy	79		377	
	General Medicine	31			
	Pathology	27			
	Physiology	17			
Medical Sciences	Pharmacology	6			
	Hematology	4	172 (20 (1)		64.89
	Oncology	4	172 (29.61)		04.89
	Surgery	2			
	Dentistry	1			
	Neurology	1			
	General Biology	41			
	Zoology	29			
	Genetics	20			
	Microbiology	11	123 (21.17)	500	
Biological Sciences	Botany	9			
	Immunology	6			86.06
	Bacteriology	3			80.00
	Ecology	1			
	Molecular Biology	1			
	Parasitology	1			
	Phylogeny	1			
Others		81	81 (13.94)	581	100.00

#### ANN. LIB. INF. STU.; JUNE 2023

is expected that they originate from newborn technical/subject-specific terms.

#### Identification of mode of formation

The subjects as presented in Table 4 were categorized further under the broad disciplines to which they belonged. The Universal Decimal Classification scheme was used for the purpose. Subjects dealing with the chemical properties of objects were put together under Chemical Sciences; subjects focusing on life, organs or parts of organs, and ecosystems were put together under Biological Sciences; and the subjects discussing human health-related issues were put together under Medical Sciences. Except for it, the rest of the subjects (including CB and Ambiguous Terms) were put together under 5 shows the result.

From Table 5, almost 86% of SRWs are distributed across three broad disciplines, including Chemical

Sciences, Medical Sciences, and Biological Sciences. Only 14% of SRWs have their origin in other subjects. So, it can be noted that the SRWs used in CB have been distilled mainly from the three broad disciplines—Chemical Sciences, Medical Sciences, and Biological Sciences, or CB uses the ideas that occur as "practice-in-action in subjects going with"<sup>18</sup> these three broad disciplines.

From the existing literature, we have already learned about the two types of distillation, namely, the distillation of Kind 1, and the distillation of Kind 2. When theories, ideas, and practice tools are distilled from diverse basic subjects to form a new primary basic subject, it is known as the distillation of Kind 1. Research Methodology, Management Science, and Conference Technique are some examples of distillation of Kind 1. Distillation of Kind 2 happens when idea(s) are distilled from the "subjects going with a particular basic subject only" to form a new basic subject. Statistical Calculus, Library Service, and Operation Research are examples of distillation of Kind 2. Another notable point includes, that the process of linguistic analysis of keywords was already implied in different literature on linguistics since long<sup>19-22</sup>, but it was hardly applied in the field of library classification till date. The different literature of linguistics includes the term semantics, which is the study and analysis of meaning of words. At this point, the concept of linguistic analysis of keywords emerges.

In the case of CB, we found that it is neither an example of distillation of Kind 1 nor of distillation of Kind 2, but has a position between the two types of distillation. As the subject-specific key terms describing the tools, techniques, or subject matters used in the domain of CB have been distilled out mainly from three broad disciplines and not from the only one basic subject applicable to the distillation of Kind 2, it is clear that CB is formed by the elaborate phase of the distillation of Kind 2 and can be termed as the distillation of Kind 2—expanded.

#### Conclusion

From the study, it is clear that CB is very much inter-related with some other scientific and medical disciplines like Biochemistry, Anatomy, Organic Chemistry, General Chemistry, General Biology, General Medicine, Zoology, Pathology, Microbiology, and so forth; which indicate that CB is multidisciplinary in nature. So it is expected that the development of or new innovations in these interrelated fields will also impact on the development and growth of CB.

The present work is done mainly by analyzing the captions of the objects of the sample articles. It brings forward the concept that the object's captions are good reflectors of the subject content of scientific articles, and they can be used as element metadata to describe the scientific articles. In future work, we will try to find out the thrust areas of CB through linguistic analysis.

#### Notes

1. Following websites are visited during root word identification and their subject-wise distribution:

- https://www.etymonline.com/
- https://www.thefreedictionary.com/dictionary.htm
- https://medical-dictionary.thefreedictionary.com/
- https://en.wiktionary.org/wiki/
- https://www.merriam-webster.com/

## References

- Cell Biology, Available at https://www.nature.com/ scitable/topic/cell-biology-13906536 (Accessed on 11 Oct 2021).
- 2 Shah R, Top 16 Techniques Used in Cell Biology (With Diagram). Available at https://www.biologydiscussion.com/ cell-biology/techniques-cell-biology/top-16-techniques-usedin-cell-biology-with-diagram/26521 (Accessed on 11 Oct 2021).
- 3 Bolin M K, Linguistics and LIS: a research agenda, *SLIS* Student Research Journal, 7 (1) (2017) 1-12.
- 4 Ranganathan S R, Prolegomena to Library Classification, 3<sup>rd</sup> edn (Asia Publishing House; Bombay), 1967, p. 351.
- 5 Kumar K, Theory of Classification, 4<sup>th</sup> rev edn (Vikas Publishing House; New Delhi), 1988, p. 201-212.
- 6 Satija M P, Subjects and disciplines: modes of formation of subjects. Library and Information Science: 02 KO & Processing: Classification, e-PGPathshala, INFLIBNET, 2021, p. 1-27. Course Material.
- 7 Neelameghan A, Primary basic subject by fission, *Library Science with a Slant to Documentation*, 10 (1973) 162-163.
- 8 Neelameghan A, Basic subjects and their arrangements, Library Science with a Slant to Documentation, 10 (1973) 207-221.
- 9 Gopinath M A and Seetharama S, Interdisciplinary subjects, and their classification, In Proceedings of the paper presented at the Third International Study Conference on Classification Research, Bombay, 6-11 January 1979, p. 121-135.
- 10 McGarry K, The Changing Context of Information, 2<sup>nd</sup> edn (LA Publishing; London), 1993, p. 164.
- 11 Satija M P, Madalli D P and Dutta B, Modes of growth of subjects, *Knowledge Organization*, 41 (3) (2014) 195-204.
- 12 Dutta B and Dutta C, A linguistic view of subject formation process as described by Ranganathan and others, *Annals of Library and Information Studies*, 61 (1) (2014) 56-64.
- 13 Dutta B and Dutta C, Concept of subject in the context of library and information science from a new angle, *Annals of Library and Information Studies*, 60 (2) (2013) 78-87.
- 14 Neelameghan A, Sequence of component ideas in a subject, Library Science with a Slant to Documentation, 8 (1971) Paper Q
- 15 Neelameghan A, Basic Subject, *Library Science with a Slant* to Documentation, 10 (1973) Papers F to N
- 16 Lieber R and Pavol S, ed, The Oxford handbook of compounding, (Oxford University Press; Oxford), 2009.
- 17 Cell, Available at https://www.thefreedictionary.com/cell (Accessed on 11 Oct 2021).
- 18 Neelameghan A, Primary basic subject by distillation, Library Science with a Slant to Documentation, 10 (1973) 168-172.
- 19 Akmajian A *et al*, Linguistics: an introduction to language and communication, (Prentice Hall of India; New Delhi) 2008.
- 20 Lehrer A, Semantic fields and lexical structure, (Amsterdam: North-Holland) 1974.
- 21 Palmer F R, Semantics, (Cambridge University Press; Cambridge) 1996.
- 22 Hintikka J, Aspects of Metaphor, (Springer: [s.l]), 1994