**Subject indexing in Italy:
recent advances and future perspectives**

Alberto Cheti
Associazione italiana biblioteche, GRIS

Anna Lucarelli
Federica Paradisi
Biblioteca nazionale centrale di Firenze
Firenze, Italy

Meeting: **200. Classification and Indexing**

Abstract

The aim of this paper is to present an overview of subject indexing in Italy, illustrating the progress made in recent years in creating effective tools for subject access that comply with international principles and standards and that are capable of dealing with the exponential growth and availability of information resources on the web, and meeting the information needs of a new generation of users.

Great attention is being paid today to classification and subject indexing languages based, as far as is possible, on shareable standards, and that allow interoperability and mappings. The "new challenge for libraries and standards" is to create future systems which have no boundaries or language limits, and are capable of working with heterogeneous resources.[1]

Recent years in Italy have seen many cataloguing initiatives bear fruit. For example, this year the new Italian cataloguing code (REICAT) has been published.[2]

---

[1] Patrice Landry, covered this subject in his paper *From traditional subject indexing to web indexing: the challenges of libraries and standards*, presented at the Conference "*Il mondo in biblioteca. La biblioteca nel mondo*", held in Milan, 12-14 March 2009.

[2] *Regole italiane di catalogazione (REICAT)*. Bozza complessiva. Versione riveduta e aggiornata. Gennaio 2009, http://www.iccu.sbn.it/upload/documenti/REICA_bozza_complessiva_genn2009.pdf?l=it. The definitive edition in printed format is expected by the end of June 2009.

1. The Italian tradition of subject indexing to date and the tools applied

1.1

Italian subject indexing has a long tradition that goes back to the beginning of last century. In the 1920s the Italian library community based their first subject cataloguing languages on the Library of Congress Subject Headings, but it wasn't until 1956 that the *Soggettario per i cataloghi delle biblioteche italiane*,[3] was created at the National Central Library of Florence (BNCF). The *Soggettario* based on the conceptual and structural model of the 5th edition of the Subject headings used in the dictionary catalogues of the Library of Congress, promoted the introduction of subject cataloguing throughout Italian libraries. In 1958, the Italian National Bibliography (BNI), edited by the BNCF, introduced the DDC as its system for organizing bibliographic records in its monthly issues.

As no DDC Italian editions existed at that time, abridged and adapted versions were used internally. The BNCF has continued to play a leading role in creating, editing and updating of indexing tools.

For the following fifty years the old *Soggettario* would serve as the standard for subject indexing practices for Italian libraries. It was essentially a structured list of subject headings and supplied a few examples of suitable combinations of terms (subject strings), but didn't provide rules regarding syntactic or semantic relationships between terms.

The principles on which it was based were theoretically sound, but its structure showed several problem areas. The old *Soggettario* was created before the international standards were published and therefore lacked syntactical instructions for the building of subject strings and mechanisms for vocabulary control and enrichment. Over time it became outdated, not only regarding vocabulary but also the entire language structure, due to its lack of explicit rules. Its application led to very diversified (if not inconsistent) indexing practices and solutions.

Despite these problems, the cataloguing practices of the BNI played an important role as an authority for the majority of Italian libraries, offering a model of indexing language. Between 1958 and 1999, the BNI regularly edited lists of new headings with the aim of both updating and enriching the *Soggettario*'s vocabulary.[4] Its structure, however, remained unaltered.

---

[3] *Soggettario per i cataloghi delle biblioteche italiane*, a cura della Biblioteca nazionale centrale di Firenze. Firenze, Stamperia Il cenacolo, 1956.
[4] Bibliografia nazionale italiana, *Soggetti. Liste di aggiornamento 1956-1985*, Firenze, Biblioteca nazionale centrale, 1987; Bibliografia nazionale italiana, *Voci di soggetto. Aggiornamento 1986-1998*. Milano, Editrice Bibliografica, ©1999.

1.2

Subject indexing practices in the 1950s used both pre-coordinate and post-coordinate languages. In the 1980s, the construction of thesauri that followed ISO standards, became widespread, notably in specialized documentary environments where pre-coordination was mainly in use. Unfortunately, these experiences were often not linked and the thesauri constructed were multidisciplinary and multilingual, but rarely general. Since then no complete survey on thesauri prepared and in use in Italy has been conducted. However we are hopeful that, sometime in the near future, the Associazione italiana biblioteche (AIB) will carry out such an initiative, in collaboration with other organizations such as ISKO Italia and Associazione Italiana per la Documentazione Avanzata (AIDA).

1.3

Another indexing tool that became widely used in Italy in the 1980s was the DDC, thanks to the publication of the first DDC Italian versions edited first by the AIB in 1986 and adopted by the Italian National Bibliography (BNI) in the same year. Since 2000 the Italian versions have been edited by the National Central Library of Florence. It is also important to underline that the BNCF, since the publishing of the Italian Edition of DDC21 has taken on the task of editing all future Italian translations of the DDC, thus carrying out its institutional duty to provide official tools for semantic cataloguing.[5]

An analysis of the use of the DDC in Italy has recently been conducted by the BNCF in April and May of this year (the first of its kind). The Central Institute for Union Catalogue (ICCU) that manages in Italy the SBN network supplied essential information. Analysing the results of the detailed questionnaire showed that the DDC is the most used classification system, rarely working hand in hand with other international classification systems in use mostly in university and research libraries. The data indicate that the DDC is not uniformly applied: not all libraries use the same (latest) edition nor the same full or abridged format. Some use both systems at the same time, depending on the material to be indexed. The DDC is predominantly found in those libraries using it for cataloguing purposes and not just for shelving. Our survey was performed with the aim of planning initiatives that would improve the functionality and use of the DDC such as online tools

---

[5] The BNCF has been given this commitment since 1998. The first DDC Italian version edited by the BNI was the 21st Edition (published in 2000), followed by the 14th (published in 2006) and the 22nd (published in January 2009). Thus the BNI is obliged to permanently study the DDC and its application, understand its richness and its difficulty, maintain continuous contact with the OCLC editorial committee and present itself as a reference point for Italian libraries.

compatible with the new-generation OPACs (as the positive experience with the Deutsche Nationalbibliothek has shown).[6]

Therefore, in summing up the history of classification and indexing in Italy to date, we see that a considerable variety of tools and practices were used until only recently. The old *Soggettario* and the Italian editions of the DDC have both acted as important reference points and have allowed for sharing of methodology. However, full adherence to international standards has only come about recently. Progress has been gradual but significant. One project that has contributed to this development with important innovations has been the Research Group on Subject Indexing (GRIS), formed by the AIB, that produced a *Guidelines for subject indexing*, published in 1996.[7]

The theoretical fundamentals and the norms elaborated by GRIS, contributed to the impulse for the renewal project of the old *Soggettario* (commissioned and supervised by the BNCF) from which the *Nuovo soggettario* was born in 2007. The *Nuovo soggettario* is an indexing language, consisting of a set of rules and a controlled vocabulary in the form of a general thesaurus.[8]

2. Background of the *Nuovo soggettario* and the GRIS contribution

2.1

Derek Austin, when introducing PRECIS, warned those readers already expert in the use of subject indexing systems such as *Sears* or *LCSH*, "to set this experience to one side for the present, and approach PRECIS, if not with a *tabula rasa*, then at least with an open mind".[9] With this in mind, GRIS began its work at the beginning of the 1990s. It momentarily put aside the old *Soggettario* and, with an "open mind", looked not only at PRECIS but at other systems and methods, at thesauri, bibliographic classifications, chain indexing systems, and at the opportunities of standardizing according to international principles, as well as at modern theories on subject indexing. Each one of these possibilities offered a possible contribution to the development of the new subject indexing method in Italy. The first objective of GRIS was to improve on the traditional methods of subject indexing already developed and in use in Italy and not make a clean sweep. A new approach was

---

[6] M. Heiner-Freiling, *DDC German. The project, the aims, the method*, in *Moving beyond the presentation layer*. Published as supplement of "Cataloging and Classification Quarterly", v. 42, n. 3/4, p. 147-163; see also: http://melvil.d-nb.de/#MelvilClass.

[7] Associazione italiana biblioteche-GRIS-Gruppo di ricerca sull'indicizzazione per soggetto, *Guida all'indicizzazione per soggetto*. Roma, AIB, rist. 2001, http://www.aib.it/aib/gris/gris.htm.

[8] Biblioteca nazionale centrale di Firenze, *Nuovo soggettario. Guida al sistema italiano di indicizzazione per soggetto. Prototipo del Thesaurus*. Milano, Bibliografica, ©2006 (stampa 2007).

[9] D. Austin, *PRECIS. A manual of concept analysis and subject indexing*. 2. ed. London, The British Library, 1984, p. 2.

proposed which allowed for the creation of new criteria and rules that would modernize the traditional tools and improve the quality of subject indexing.[10]

Austin stressed the following characteristics of PRECIS:

- the use of an *open-ended vocabulary*: new terms in current use can be admitted into the indexing vocabulary and used as subject headings as soon as they have been encountered in documents, without waiting for them to appear on a list of authorized headings;
- the construction of coextensive subject strings, in which different terms are interrelated to form a type of *summary* of the subject, instead of multiple subject headings, used as independent access points and often containing information that is too general for the specific subject.

These aims characterized the GRIS project. It is clear that to achieve these two objectives implied the use of control tools that were rigorous, logical and generally applicable on both a semantic and syntactic level, i.e. the necessity of precise, shareable rules. The mission of GRIS was to achieve these aims and its first step was the publication of its guidelines.

The "open mind" that Austin advised was also mentioned by that pioneer of modern theories of subject indexing, Brian Vickery, in his address at an ASLIB meeting over fifty years ago: "The different techniques of indexing differ in *form*, in the way the principles of subject analysis are applied. Each form has own particular advantages and field of use - and much work remains to be done to define the appropriate field in which each technique is applicable. But this evaluation will be easier if we first gain a clear understanding of the principles which are to be applied".[11] This unitary vision of the principles of subject indexing and bibliographic classifications, shared by other experts, such as Eric Coates and Austin, paved the way for the research work done by GRIS.[12]

The new approach set down in the GRIS guidelines took effect where the old *Soggettario,* with its traditional subject indexing, left off. The renewal process had two aspects:

- a conservative aspect, that dealt with the preservation of the terminology of the old *Soggettario,* as a starting point for the new vocabulary, and;
- an innovative aspect, seen in the adoption of the most evolved model proposed by the GRIS guidelines.

---

[10] Associazione italiana biblioteche, GRIS – Gruppo di ricerca sull'indicizzazione per soggetto, *Per un metodo coerente di indicizzazione per soggetto*. "AIB notizie", 4 (1992), n. 6/7, p. 10-12 (p. 10).

[11] B. C. Vickery, *Developments in subject indexing*. "The journal of documentation", 11 (1955), n. 1, p. 1-11 (p. 10).

[12] E. J. Coates, *Classification in information retrieval: the twenty years following Dorking*. "Journal of documentation", 34 (1978), n. 4, p. 288-299; D. Austin, *Basic concept classes and primitive relations*. In *Universal classification. Subject analysis and ordering systems. Proceedings 4th Internatl. Study Conference on Classification Research, 6th Annual Conference of Gesellschaft für Klassifikation e. V., Augsburg, 28 June - 2 July 1982*. Frankfurt, Indeks Verlag, 1982, vol. 1, p. 86-94.

2.2

The main conceptual model to which GRIS, and afterwards the *Nuovo soggettario* referred to is the analytico-syntethic model conceived by Ranganathan and developed by the British Classification Group (CRG). This model, used in facet classifications, efficient at defining the macrostructure of a verbal indexing language, was also accepted by various normative documents such as *Guidelines for the establishment of monolingual thesauri* (ISO 2788-1986), *Principles underlying subject heading languages (SHLs)* (IFLA, 1999) and *Structured vocabularies for information retrieval. Guide. Part 2: Thesauri* (BS 8723-2:2005). Another important reference has been the criteria and procedures for vocabulary control and construction of semantic relationships, according to the various standards ISO, BS and NISO.[13] Today the main indexing systems such as LCSH, SIPORBASE and RAMEAU meet these criteria and procedures, but there has not been the same concordance on their syntactic aspects.

Maria Inês Cordeiro, presenting the results of the *Working Group on Principles Underlying Subject Heading Languages* wrote: "it appears that subject systems are currently more advanced and concordant to each other in technical devices for terminological control and semantic (paradigmatic) structure, but less structured and similar in respect to syntagmatic aspects of knowledge representation. In this regard, it seems that this needs to be explored to acquire common understanding and compatibility, if libraries want to pursue pre-coordination".[14]
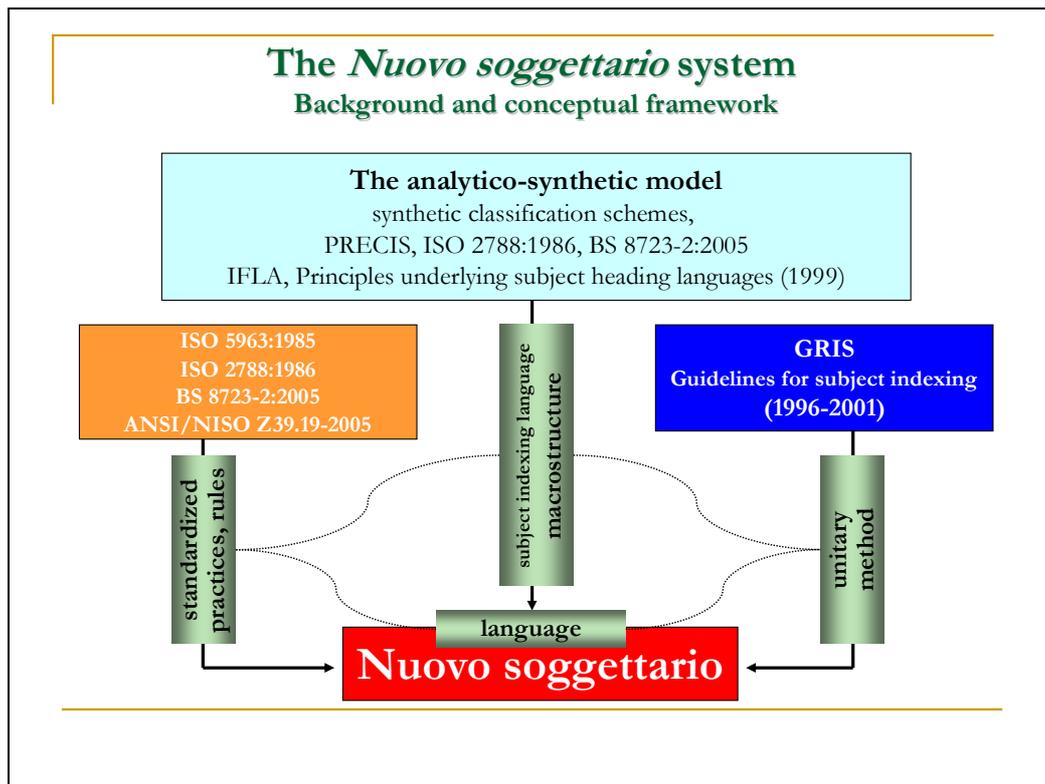
The GRIS conceptual and theoretical choice, adopted for the *Nuovo soggettario* was directed towards the *role analysis* that favors *precision*, *consistency* and *predictivity* within the subject strings.[15] The roles provide a model of knowledge representation, on a syntactic level, that is consistent with the model of summarization of document content established by the ISO 5963 standard. However, the *Nuovo soggettario* offers syntactic rules simpler than those suggested by the GRIS guidelines.

The following chart shows an overall view of the theoretical and methodological background of the *Nuovo soggettario* system:

---

[13] ANSI/NISO Z 39.19 -2005, *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*.

[14] M. Inês Lopes, *Principles underliyng subject heading languages: an international approach*. 61st IFLA General Conference, August 20-25, 1995. http://ifla.queenslibrary.org/IV/ifla61/61-lopm.htm (accessed June 13, 2009).
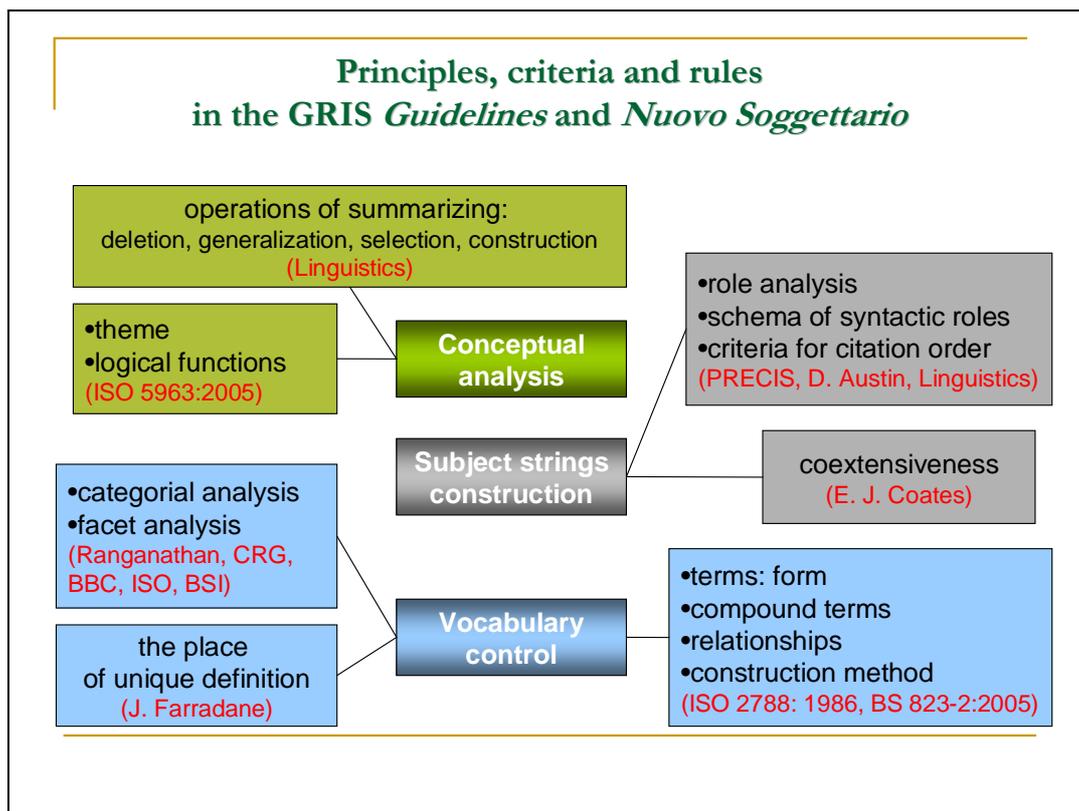
[15] *Role* is the individual component of the syntactic structure of subject strings, namely the position given to a term within the strings. This position represents the logical function of a concept in defining the subject (GRIS, *Guida all'indicizzazione per soggetto*, p. 8; Biblioteca nazionale centrale di Firenze, *Nuovo soggettario*, p. 119).

The *Nuovo soggettario* system
Background and conceptual framework

The analytico-synthetic model constitutes the macrostructure of the indexing language (vocabulary and syntax, the two distinct, but when necessary interrelated, components). The ISO norms provides guidelines, standard procedures and rules for examining documents and vocabulary control. The GRIS Guidelines offers a comprehensive method. However, the *Nuovo soggettario* represents a real indexing language with a set of practical instructions added to criteria, procedures and rules.

2.3

The chart below shows the general principles, criteria and rules on which the GRIS *Guidelines* and *Nuovo soggettario* are based. These have been grouped according to the three fundamental operations of subject indexing.

**Principles, criteria and rules
in the GRIS *Guidelines* and *Nuovo Soggettario***

operations of summarizing:
deletion, generalization, selection, construction
(Linguistics)

•theme
•logical functions
(ISO 5963:2005)

Conceptual analysis

•role analysis
•schema of syntactic roles
•criteria for citation order
(PRECIS, D. Austin, Linguistics)

Subject strings construction

coextensiveness
(E. J. Coates)

•categorial analysis
•facet analysis
(Ranganathan, CRG, BBC, ISO, BSI)

the place
of unique definition
(J. Farradane)

Vocabulary control

•terms: form
•compound terms
•relationships
•construction method
(ISO 2788: 1986, BS 823-2:2005)

Regarding the conceptual analysis the following elements are evident:

• explicit reference, included in the GRIS Guidelines, to macro-operations that, when necessary, can assist indexers in basic theme reconstruction;[16]

• the notion of "theme", used in ISO 5963, for defining the document subject;

• the list of logical functions or elements (factors) that have been recognized as essential in analysing a subject and that define its extension or exhaustiveness (e.g. a transitive action, an object, an agent, an instrument, etc.).

The syntactic model is based on the notion of *coextension* (E. Coates) and on the *role analysis* method (PRECIS), that implies a role schema and some criteria for the citation order of terms within a subject string. These criteria demand that each term in the string would be directly related to its neighbour (*one-to-one relationship*) and imply that preference would be given to those concepts that play the logical function of "object" of a transitive action, "agent" of an intransitive action, or "possessor" of its "part/property" (*principle of logical dependency*).

Finally, regarding the principles and methods for vocabulary control: the analysis of the basic meaning of a term (the place of *unique definition*), the assignment of category and facet (*categorial*

[16] Teun A. van Dijk, *Complex semantic information processing. In: Natural language in information science. Perspectives and directions for research. Results of a workshop on linguistics and information science, Biskops-Arnö, Sweden, 3-5 May 1976*. Edited by Donald E. Walzer, Hans Karlgren, Martin Kay. Stockholm, Skriptor, 1977, p. 127-163.

analysis and *facet analysis*), the choice for the form of the terms in compliance with the three-relationship model (equivalence, hierarchical and associative relationships); the method of the construction of the semantic relationship network, including both the inductive method for the development of vocabulary and the deductive method which works on revision of the general structure of categories and facets.

2.4

In the *Nuovo soggettario* the distinction between syntactic and semantic relations, between subject strings and concepts, and between vocabulary and syntax depends on the two models used in the control of the relationships: *roles* and *categories*. The syntactic role assigned to a term determines its position in the subject string and varies according to the context. The category assigned to a term, based on its meaning, is, however, permanent and determines its placement in the structure of the vocabulary. What are the advantages of this approach? As declared in the Library of Congress's report on *pre and post-coordination*, the main advantage of pre-coordination is *flexibility*: "Pre-coordinate terminology is more flexible in terms of system design. Pre-coordinate indexing can be disassembled for use in post-coordinate systems. Pre-coordinate indexing can be useful for either browse or keyword searching ". Pre-coordination also "enables clearer indications of relationship of topics when that is the point of the work".[17] What is important is the requirement of *specificity* and *expressivity* of the language; thanks to syntax, the language is capable of expressing relationships that are more specific than those expressible by the operators AND, OR and NOT. To this we can add the requirement that Elaine Svenonius calls *contextuality*, that is to say, the capacity of the context to specify the meaning of the term, meaning that: the "context, called in indexing the *subject statement*, and represented in the catalogue by the subject string, enables not only different functions or roles of a term in relation with other terms of the context, but also partially different knowledge of the same term, that can be considered as aspects, components, or modifications of its meaning".[18] A final requirement is *precision* or *recall* in the retrieval of documents.

Which form of pre-coordination is more efficient in satisfying these requirements? The dilemma "pre-coordination or not" should be transformed, as suggested by Elaine Svenonius, into the question "to what degree and in what manner should terms be pre-coordinated". In fact, the solution of the dilemma "depends at least in part on the kind of pre-coordination that is done. The distinction between an enumerative (or largely enumerative) index language and a synthetic one becomes

---

[17] Library of Congress, *Library of Congress Subject Headings. Pre- vs. Post-coordination and Related Issues*. Report for Beacher Wiggins, prepared by the Cataloging Policy and Support Office, March 15, 2007, p. 7. http://www.loc.gov/catdir/cpso/pre_vs_post.html (accessed June 13, 2009).

[18] A. Cheti, F. Paradisi, *Facet analysis in the development of a general controlled vocabulary*. "Axiomathes", 2008, n. 18, p. 223-241 (p. 234) (DOI 10.1007/s10516-008-9033-4).

important, as will be seen, in the evaluation of the relative merits of pre and post-coordinate languages".[19]

The *Nuovo soggettario* has adopted, in fact, a synthetic pre-coordination: indexing can produce a subject string that best represents the subject of the document – "the subject-heading made to measure, the subject heading co-extensive with the subject of the book", as defined by Coates[20] – beginning with the terms of a controlled vocabulary, arranged according to syntactic rules. The benefits of an analytico-synthetic language are not only seen on a syntactical level. There are also advantages to the vocabulary. Indexers can work with simple units, that is to say, indexing terms subjected to semantic control criteria. In this way a systematic and coherent development of the vocabulary is supported.

Distinguishing between the semantic and the syntactic level also allows the *Nuovo soggettario* to be used for post-coordinate indexing. The two levels are integrated into the system as distinct components and, when necessary, separable. Naturally, it is necessary that these language characteristics are valued in relation to both OPAC functionalities and search modalities. Lois M. Chan is correct when stating that a pre-coordinate language "is of little value in environments where browsing the index or the thesaurus is not an option, because users typically do not search by string".[21] Thus, the modality of *two-stage research* (from the terms to the subject strings and from the strings to the documents) should be considered an indispensable option when indexing with a pre-coordinate language, although coupled with other access modalities.

Therefore, a cost-benefit analysis of subject cataloguing should keep in mind, on one hand, the quality of the language and the form of pre-coordination adopted (in terms of flexibility, specificity, contextualization etc.) and, on the other hand, the functionality implemented in the catalogues and their suitability at rendering visible and efficient the characteristics and the potential for expression of the indexing language in use.

3. *Nuovo soggettario*

Italian libraries are moving in the direction of adopting the *Nuovo soggettario*, an indexing language capable of effectively satisfying what is today required of subject indexing languages: to

---

[19] E. Svenonius, *Precoordination or not?* In *Subject indexing: principles and pratices in the 90's. Proceedings of the IFLA Satellite Meeting held in Lisbon, Portugal, 17-18 August 1993*. Edited by Robert P. Holley, Dorothy McGarry, Donna Duncan, Elaine Svenonius. München, Saur, 1995, p. 231-255 (p. 235-236).

[20] E. J. Coates, *Subject catalogues. Headings and structure*. London, The Library Association, 1988 (1. ed., 1960), p. 33.

[21] L. M. Chan, *Thoughts on LCSH* (July 31, 2006). In: Library of Congress, *Library of Congress Subject Headings. Pre- vs. Post-coordination and Related Issues*, Appendix 1, p. 16.

comply with shared principles and standards, to sustain interoperability and mappings, to develop in a multilingual direction.

Thanks to the separation between the semantic and syntactic components, the *Nuovo soggettario* can function as either a pre-coordinate or post-coordinate indexing language. As mentioned in the Library of Congress Report, the consideration that a language be able to provide these two approaches is important. It is also acknowledged that the two concepts of *recall* and *precision*, used also in bibliographic information retrieval from the web, are not outdated concepts but continue to be important in reference to various research contexts and opportunities provided by the latest-generation OPACs.

The *Nuovo soggettario* provides an Italian vocabulary, intended for Italian libraries, but it is also capable of being developed in a multilingual direction given that "the monolingual subject access is the major constraint for researchers wishing to conduct simultaneous subject queries across many catalogues"[22]. It is also applicable in a variety of information environments and not for exclusively bibliographic and library applications and is suitable for the indexing of heterogeneous resources (a requisite now widely recognized as important). In fact, the latest version of the new database of the SBN Index provides software modules for the bibliographic treatment of different types of material, such as graphic and cartographic materials, printed and manuscript music, online and digital resources, photographs, etc.).[23]

As it has been already underlined, the *Nuovo soggettario* is made up of different but complementary components:

- a printed guide (which can also be accessed in PDF format through the user interface) consisting of rules governing both vocabulary control and the construction of subject strings;
- a general thesaurus, i.e. the controlled and structured vocabulary accessible on the web on subscription. It now comprises approximately 33,000 terms;[24]
- a *Manual* for cataloguers: i.e. a set of practical instructions to help in the application of the syntactical rules governing full categories of terms and/or specific bibliographic cases. It is available through the user interface;

[22] P. Landry, *From traditional subject indexing to web indexing: the challenges of libraries and standards*, cit.
[23] http://www.iccu.sbn.it/genera.jsp?s=59&l=it (accessed June 13, 2009); for more information on the SBN Web, see G. Contardi, *The Italian National Library Service (SBN): a cooperative library service infrastructure and the bibliographic control*, paper to be presented at the 75th IFLA General Conference, Milan, 23-27 August 2009.
[24] For the Thesaurus features, see L. Franci, A. Lucarelli, M. Motta, M. Rolle, *The Thesaurus of Nuovo soggettario: structural features and web application projects*, paper to be presented at the Satellite Meeting "Looking at the Past and Preparing for the Future", Florence, 20-21 August 2009.

- the subject string archive, i.e. the list of the subject strings built on Thesaurus terms and syntactic rules. The BNCF's OPAC includes the subject headings created by the BNI cataloguers, who have been using the *Nuovo soggettario* since September 2007.

The analytico-synthetic model is oriented towards producing specific coextensive strings that closely reflect the conceptual content of the document being catalogued, as shown in the following figure.

Bibliothèque du voyage français en Italie
  Fasano : Schena ; Paris : Lanore.
  v. ; 21 cm
  Contents:
  1 : Le voyage français en Italie des origines au 18. siècle : bibliographie analytique / Vito Castiglione Minischetti, Giovanni Dotoli, Roger Musnik
  2 : Le voyage français en Italie au 19. siècle : bibliographie analytique / Vito Castiglione Minischetti, Giovanni Dotoli, Roger Musnik
  Subject(s)

- **BNI**
1. **Italia** – **Descrizioni** [**dei**] **Viaggiatori francesi** – **Sec. 18.-20.** – **Bibliografie ragionate**

- **BNF**
1. **Voyageurs français -- Italie -- Bibliographie analytique**
2. **Récits de voyages français -- Bibliographie analytique**
3. **Italie – Descriptions et voyages -- Bibliographie analytique**

- **World Cat**
1. **Traveler's writings, French -- Italy -- Bibliography**
2. **French -- Travel -- Italy -- History -- Bibliography**
3. **Italy -- Description and travel – Bibliography**

Nevertheless, single coextensive strings not always represent the best choice when facing complex subjects, therefore over-complex solutions ("engineered") should be avoided.[25] Coextension, however, in most cases, produces a good solution from the viewpoint of expressiveness and communication of the conceptual content of the document and from an information retrieval viewpoint.

Many Thesaurus terms have syntactic notes, i.e. specific instructions on the way a specific term should be used within the subject strings. The purpose of the syntactic notes is to facilitate, when necessary, the application of the general syntax rules, and to assist cataloguers in the construction of subject strings. It is a set of instructions drawn up in compliance with the method of "role analysis", which includes all the necessary indications for a given term, e.g. its syntactic value, logical

---

[25] As Mary Dykstra pointed out in her discussion at an international Seminar on the Feasibility study for the renewal of the *Nuovo soggettario*, held at the BNCF, in Florence (Italy), April 2001.

function, syntactic role and citation order. The syntactic note serves as a link between the two separate but complementary components: the syntactic and the semantic.

The above-mentioned features enable this tool to be used in both general and specialized information environments (e.g. government, public, academic and research libraries). Regarding its applicability for indexing of special material, the BNCF has recently used it for subject indexing of photographs belonging to a rare photograph collection (the photograph archive of "Il mondo", an illustrated magazine edited by Pannunzio). The figure below shows a recorded photo from the BNCF's OPAC with its subject heading:

Sophia Loren divertente interprete del film italo americano La baia di Napoli : Roma international foto press, [1960].
1 fotografia : b/n ; 22x19 cm + 1 c. (Fotografie dell'archivio de "Il mondo")

Tit. dattiloscritto sulla c. alleg. - Pubbl. in Il mondo, a. 12 (1960), n. 22, con il tit.: Hollywood : i lustrini di Sophia Loren nel film La baia di Napoli

1. **It started in Naples <film>** – **Set cinematografici** – **Attori** [:] **Loren, Sophia** – **1960** – **Fotografie**

The *Nuovo soggettario* Project is progressing according to the recommendations of the preceding feasibility study at speeds reflecting the available financial resources (70% of which come from private resources). Once new funds are raised, the project infrastructure will be finished, including the creation of the necessary software for cooperation between libraries. By the end of 2010 we foresee it becoming a permanent system, housed at the BNCF and regularly maintained and updated by a dedicated staff.

4. Future cooperation

The plans for future cooperation being laid in Italy at the present time are looking, above all, at developing a collaborative use of the recently created indexing tools and at opening up to international experiences. As far as national cooperation is concerned, the BNCF has already set up conventions with various organizations that use the *Nuovo soggettario* and with whom it can share its evolution on a methodological and technical level. The aim of these initial experiences in cooperation has been to enrich the Thesaurus thanks to suggestions of specialized terminology by the partners, to test interoperability with other thesauri and share resources. These collaborations satisfy different needs, with the aim of evolving the *Nuovo soggettario* in the direction of a public service on a national level. The BNCF is presently cooperating with seven partners. The applied model could gradually diffuse over time to include other interested institutions. As well as collaborating with the National Central Library of Rome, the BNCF collaborate mainly with academic and research libraries and documentation organizations. We are collaborating with the libraries of Pisa University, the library of the Bocconi University (Milan), the ITTIG (Institute of Legal Information Theory and Techniques, Italian National Research Council), and the company Idest, a private producer of bibliographic records, sharing the production of the BNI children's books series with the BNI editorial office.[26] Some of these partners manage their own specialized thesauri that will be integrated into the general Thesaurus of the *Nuovo soggettario*. Methodological protocols and electronically usable procedures and forms have been set up with them to submit proposals of new terms or changes for inclusion in the Thesaurus, on a regular basis, in a similar way to what is currently done with SACO partners.[27] The contents of these collaborations are not homologated between them and, as well as the submissions of new terminology, they are based essentially on: training of the cataloguers, current and shared usage of the new language, even if with diversified levels of indexing, and access to the Thesaurus data base. In certain cases, the collaborator also supplies a preliminary stage of vocabulary control (e.g. scope of a certain term, sources used to back it up, etc.) and in other cases, supplies scientific or linguistic consulting on specialized terminology, as in the case of the ITTIG, for legal terminology. There are many positive aspects to this BNCF's initiative: the variety of institutions involved, the multiplicity of content of each agreement; the training offered by the BNCF as part of the setting up of the collaboration. The delineated model deals with actual problems (both on scientific and technical levels) such as the integration between specialized and general thesauri, their interoperability and the preparation of controlled vocabulary for semantic access to non-book material. The Thesaurus of the *Nuovo*

---

[26] For more information on this collaborative work, see F. Paradisi, *Collaboration between producers of bibliographic data. Achievements and planned initiatives at the Italian National Bibliography* http://archive.ifla.org/IV/ifla74/papers/162-Paradisi-en.pdf.
[27] http://www.loc.gov/catdir/pcc/saco/sacopropform.html.

soggettario will not be the only adoptable Italian tool, but it could become a sort of meta-thesaurus that could function as a reference structure, enabling mappings across different vocabularies.[28]

Thanks to the ICCU participation, this small-scale cooperation model we are testing should be widened to the majority of Italian libraries participating in the National Union Catalogue (SBN), also through different descriptive levels. A real advantage will be the reduction of subject cataloguing costs through cooperative efforts. We hope, therefore, to move in two directions in the future:

- to decentralize the sharing of methods and procedures of the Thesaurus enrichment/development;
- to centralize the maintenance of the system and the validation of the subject headings, thanks to a national subject authority system still to be created.

These initiatives will be possible on the condition that the team responsible for the *Nuovo soggettario* will become an official reference centre, as occurs in France, for example, at the Centre nationale RAMEAU.[29] They should be housed in a permanent structure with adequate human, financial and technological resources that can be dedicated to the maintenance of the system, to its development and to the validation of the subjects proposed by the partners.

Analogous plans to strengthen cooperation should be prepared for the diffusion of the use of the DDC in our libraries, with the aim of continuing editing of the Italian versions and developing online products, no doubt necessary for the future.

---

[28] BS 8723:2005-2008 - *British Standards Institution. Structured vocabularies for information retrieval. Part 4.* London, British Standards Institution.
[29] http://rameau.bnf.fr/informations/convention.htm.