Pauline (Atherton) Cochrane

Eric de Grolier: The Analytico-Synthetic Summarizer

Cochrane, P.A.: **Eric de Grolier: The analytico-synthetic summarizer.**
Int.Classif. 18(1991)No.2, p.78-86, 5 refs.
Appraisal of E.de Grolier's ability to summarize, to relate to important and relevant developments, and to systematize the results of conferences with special reference to five conferences held during the time between 1966 and 1990. The text is continued by one example, a reprint of the full text of one of his summaries, viz. the 'Synoptic Critique' to the Conference on Relational Factors in Classification, University of Maryland, June 1966. (I.C.)

Eric de Grolier has been asked many times to 'sum up' the conferences he has attended during his 60-year career in what he has called the 'transdisciplinary' work of information retrieval, terminology, structural linguistics, semiology, concept analysis, knowledge classification, etc. I have been fortunate to be in attendance at several of these meetings and every time I was impressed by his ability to 'sum up' in a constructively critical way, to give structure to his analysis, to provide historical perspective when necessary, and to provide warnings about problems and directional signals and recommendations for where to go next.

Such an ability as Eric de Grolier has exhibited is rarely praised or rewarded but it is very essential. When we present our papers at such conferences and there is too little time for critical comment we often go away relaxed and elated when in reality we should be puzzled and prodded into more study of the work of others and better designs for our research. Over and over again, Eric has provided such constructive criticism, reference to previous relevant research, and helpful suggestions for new research. We all owe him a debt of gratitude for his contributions in this area.

Recently I had the opportunity to read again what Eric had to say about five different conferences, spanning the years from 1966 to 1990. In every case he gave an 'abstract' of the conference which would be valuable in any database on the subject. To have the full text of Eric de Grolier's synoptic critique of the 1966 Symposium on Relational Factors in Classification (1), or the synthesis of the 4th FID/CR Study Conference at Augsburg in 1982 (2), or the 'wrapping up presentation' at the 1981 CONTA Conference (3) and the TKE'87 Conference in Trier (4), or the First International ISKO Conference 1990 in Darmstadt (5) is to have an instant recall of the conference if you attended, and to have a rich description of the issues, topics, and problems discussed if you did not attend. When the history of our field is written it is Eric de Grolier's summaries which will help document whether progress was made or not made and whether various players in this transdisciplinary field recognized each other's work or needed to follow new directions. His distillations and observations were always coherent, succinct and lucid, not necessarily unbiased but respectfully submitted regardless. Rarely did he go beyond three or four pages, but these pages are filled with helpful references and critiques of presentations and/or lists of struggles or problems we, collectively, will have to overcome. He was rarely wrong in these lists of problems and they alone are worthy of review at frequent intervals.

In 1982 he stressed the principles of complementarity and transculturality, two principles he comes back to over and over again in later conferences. As part of his summary he has usually included an analysis of the origin of papers to show the contributions of European countries, Asiatic countries, and the like. He warns against 'Ameroeurocentrism', the limits of automation, self-conceits, dogmatism, and a myopic concentration on bibliograpahic information when 'factographic' information is the more plentiful and useful. He often is the one to point up gaps or lacunae to insure that the next conference will cover what was overlooked. For example, in 1990, he decried the lack of reports of user studies, or probabilistic methods and in 1982 he appears to have planted the idea for the organization which was founded in 1989, namely The International Society for Knowledge Organization.

I attended two of the five conferences where Eric made the summary presentations. In both cases I know how difficult the task must have been. At the 1966 one his task was doubly difficult because his own work was on the line with that of Perreault, Newman, Farradane, and others. There were lengthy papers by Soergel and others that had to be reviewed. In his remarks he seemed comfortable with the philosophers, the linguists, the logicians, the psychologists, and expecially the documentalists and classificationists, something none of the representatives of these respective groups could do. And he brought a special message to us all by recalling Sapir's famous book on language and applying his classification

of languages to the artificial information languages (page 395). In such a way he was more than a synthesizer; he was also a teacher and a sage. Re-reading his excerpt from Sapir followed by his application of Sapir's classification to our field provided a cogent explanation for what has been happening to the indexing languages and thesauri as they have developed since that time.

De Grolier's contributions then, at these conferences, were not just for the moment, but for the historical record. They deserve to be compiled and indexed, readily retrievable by us today and by future generations of workers in this field.

References (all by Eric de Grolier):
(1) Synoptic critique (of the Conference on Relational Factors in Classifiction, University of Maryland, June 1966). Inform.Storage & Retrieval 3(1967)p.385-396
(2) Synthesis of the 4th FID/CR Conference. In: Perreault, J.M., Dahlberg, I. (Eds.): Universal Classification II. Subject Analysis & Ordering Systems. Proc.4th Int.Study Conf.on Classif.Res., Augsburg, 1982. Vol.2. Frankfurt: Indeks Verl.1983. p.163-167
(3) A Conference Resumé. In: Riggs, F.W.(Ed.): The CONTA Conference. Proc.Conf.on Conceptual & Terminological Analysis in the Social Sciences, Bielefeld, FRG, May 24-27, 1981. Frankfurt: Indeks Verl.1982. p.320-322
(4) Wrapping up presentation. In: Czap, H., Galinski, Ch. (Eds.): Terminology and Knowledge Engineering. Supplement. Proc.Int. Congress, Trier 1987. Frankfurt: Indeks Verl.1988. p.207-211.
(5) Conference Summary. In: Fugmann, R.(Ed.): Tools for Knowledge Organization and the Human Interface. Vol.2. Proc.1st Int.ISKO-Conf. 14-17 Aug.1990. Frankfurt: Indeks Verl.1991. p.248-251
Address: Prof.Pauline A.Cochrane, 23/5 Flower Road, Colombo 7, Sri Lanka

---

Editorial Note: *We would like to provide an impression of E.de Grolier's ingenious gift here described by reprinting - as an example - from the source indicated above under (1) with slight editorial amendments the text of the Synoptic Critique mentioned already. This will mean for many of us at the same time also an entry into a most interesting discussion of problems of which the majority seem to be still unresolved today. The responses made at that occasion have not been included (see however (1, p.397). We are grateful to Pergamon Press for kindly permitting this reprint.*

Eric de Grolier

SYNOPTIC CRITIQUE (of the Conference on Relational Factors in Classification)

"....I will again take up the summary which I gave you at the beginning of this conference instead of a criticism of Perreault; I shall try to see how far these questions have been answered. If they have not been answered, I shall try to see at least what sort of clarifications we could draw from our three days's discussions. ... I will change my scheme of classification a bit as a result of this conference, since its main result, for me, probably, was to have to prove the value of my ternary division - and to say that it was not good at all. You will remember, in my earlier presentation, that I established three categories: (1) the subject in itself; (2) the subject in passive relation with its milieu; (3) the subject in active relation with its environment - but this seems not to fit with what has been said here.

The first category is probably good, the one which considers the relational factors in themselves - in essence, as Perreault would probably say. Here we find (1.1) the definition of relational factors in terms of their limits. What can we consider to be relational factors? Do categories (1.2) enter into the picture or not, and to what extent? A third part would be to ask (1.3) whether the relational factors have (or have not) meaning in themselves, or in combination, or in content and so on? (1.4) Then, what is the number of relations? (1.5) What are their types? (1.6) What is their form?

The second category of questions, however, will be changed in comparison with what I said at the beginning: there will now be a perfect ternary subdivision: (2.1) the basis or foundation of relational factors; (2.2) their classification, comparison, and standardization. The third question, which was suggested by this discussion, I would call (2.3) their 'niceties' or their 'shades'.

The third part of this ternary grouping will be rephrased 'practical application'. That is, (3.1) (and this is an open question) if they have any use at all. (3.2) Second, the ways to express them. (3.3) Third, experimentation and testing. And, of course, there will be (in conclusion, as a sort of 'Perreaultism') something without which a good classical French exposé would prove to be of no value.

1. The Relational Factors in Themselves

1.1 The definition of 'Relational Factors'

Well, let's begin at the beginning. What is a relational factor? I must say that I was not very satisfied with the definitions I got in this conference. The best definition I ever found was one by Lévy, but one he did not give in this conference. The definition I refer to is included in a paper which he delivered at the Conference on Data Archives, which was sponsored by my organization, the International Social Science Council, in common with the International Committee for Social Sciences Documentation, in September 1964. I believe that has been published in a rather good issue of our journal, Social Sciences Information, which is entirely devoted to the question of data archives. Lévy, in his paper, equates relations with logical connectives and defines them tentatively as those elements of meaning which could without de facto limitations due to the meaning of connected terms link all index terms in the vocabulary. I think this is by far the best definition I have found.

We have other definitions given in this conference, for instance that of Perreault, who gave at least an implicit definition, emphasizing that "It is only in the transformation, by the use of syntactic elements, of mere strings or constellations of semantic elements, that indication comes to be information". By this it is probable that he intended by 'syntactic elements' relationships. I would object to the phrasing syntactic versus semantic, but that is another question.

Then on p.183 he defines the role, which is a bit more restrictve than the relational factor; but I think this definition could be attached to other relational factors also: "The role is that which attaches to each substantive element the delimitation of its function within the proposition". So here is a syntactic definition of relational factors.

Soergel also defines, or attempted to give a definition of p.232 of his report. He says also that the role indicators "may be viewed as very general concepts applicable over a broad range of fields of thought". Indeed, it is very near the definition of Lévy, but I think that the definition by Lévy is better phrased.

The remarkable fact is that if you take this definition of Soergel, and if you compare it with my own definition of categories which is given in this big book of mine (1), you will see - well, it's a repetition as usual: it is taken from my UNESCO study (2) and it says that categories are concepts (including the logical relationships) of a broad enough field of application to entitle them to be qualified as general. That indicates which terms in the language are applicable on a very wide range of fields, but no means are provided here to distinguish categories and relationships, which is bad.

Pagès also tried to make a definition and he made it by reference to logic. It is perhaps unnecessary to repeat what I said this morning on his definition. It seems to me that all these definitions introduce the syntactic element, but we will see more of this later.

1.2 Do Categories enter into the picture or not and if, to what extent?

Perhaps it would be better concerning categories if we see that the place of relational factors within the categories is not clear. It is not evident, let's say: some systems seem to confound them, to take them as the same thing, and some other schemes do not.

Perhaps it would be better to return to a previous terminology and to speak not of categories, of this kind of stuff which classificationists have the habit of calling categories, but to return to the old part-of-speech notion, or to a notion of form-classes.

Here again I will refer to a definition which was given by Lévy in a private letter (but a private letter which is published is no longer private), and which was reproduced without his permission in my AFOSR report. That is the letter of April 27, 1964. Lévy, one of Gardin's collaborators - at that time he was much less well known than now - wrote us that "In every (information retrieval) language with free syntax, as SYNTOL, and in the natural language, a partition of the lexical units is necessary for the univocal interpretation of the combination of units. The outcome and number of these grammatical categories and their nature is dependent upon numerous factors - number of the lexical units, semantic field covered, aims of the language, etc. - which to my knowledge have not been submitted to special investigations" - nor to my knowledge - . "A semantic base of the categories may be convenient, but it is not necessary. Thus in natural language the semantic characteristics of genders are for a long time not perceived".

I added a note to observe that it seems, however, that the example of gender is rather unfortunate as gender is probably not essential for the univocal interpretation of the combinations of lexical units. Nevertheless, this definition is perfectly clear and corresponds to what the new structural linguists call, generally speaking, form-classes, or syntactic categories, etc.

You will find a very good book which was published some weeks ago by Benveniste (3), a renowned French linguist, which is a collection of his articles and studies. He insists very strongly on the fact that so-called parts of speeches or form-classes are determined essentially by the syntactic roles which are assigned to each category.

In Bochenski's book (4) there is a good discussion of the syntactic categories which it is perhaps worthwhile to cite. One syntactic category is the class of expressions of language each of which can be exchanged with any other of the same class in a meaningful statement without depriving the statement of meaning. He says that this is more or less equivalent to the old parts of speech and that the number of primitive categories is fairly arbitrary.

1.3 Do Relational Factors have meaning in themselves, or in combinatioon, or in content?

Now, what is the use of categories in comparison to relationships properly speaking? Some systems of information retrieval are using both concurrently. For instance, Gardin's SYNTOL, as you know, uses parts of speeches with the aim of giving rules for orienting relations. Had I time enough I would discuss at length the rules for orientation with the aid of form-classes which are (or more exactly, are not), given in SYNTOL, but which were given very kindly by Lévy to me in a more or less internal report which, of course, I was very eager to reproduce as it was a very good piece of research; it is in my book (1).

Nevertheless, these parts of speech of SYNTOL exist. It seems that they have been more or less abandoned. There is a note in the last part (the hierarchies) of his lexicon, published for sociology, in which he says, "We have not classified the parts of speech as it seems that there are more urgent tasks in the field of sociology than to categorize terms".

Mrs. Atherton reported on an apparently somewhat similar scheme for the American Petroleum Institute, which uses, as far as I understand, concurrent relations (properly speaking) in the form of roles and categories of words.

Now there are other schemes, other schedules, other systems, which are more or less replacing relations with categories. For instance, I believe that the majority of facets in classification replace relationships, properly speaking, by categories. Not entirely, generally speaking, as there remains a sort of remnant of what is not expressed by the facets, which is now expressed by relationships. You remember the phase relations of the Colon Classification, and those in the English Electric System.

By the way, at the Elsinore Conference there was an excellent paper by Gardin on this problem. That is, the fact that the faceted classification replaces relationships as such by the faceted structure (5).

Now there is another theory - what I have said up to here is not theory, it is practice - that of Soergel's, namely that the categories would be formed by the relationships. That is what you find on p.231 of his report. I will not deal at length with this, I only mention it. In fact we will have more to say on categories when we will speak of the basis and foundations of relationships later.

1.4 What is the number of relationships?

Indeed, as I reminded you at the beginning, the mere fact that the number of relational factors is put in the range of thousands by Pagès and (orally) by Ceccato, or of hundreds by Ceccato (written), or of four or less by Gardin, must remind you and me that probably they are speaking not of the same but of different things. I think this was more or less clarified during the discussion with Pagès this morning.

In fact, there is probably an infinite number of relationships which are possible, or at least a very large number. They are not a finite class in the sense of the grammarians, on the one hand; they are expressed in many ways in the languages which we all speak, our mother-tongues. It can be expressed, though this has not been so much remarked, in just as many ways in the artificial languages that we devise for information retrieval and for information retrieval systems.

That is one of the lessons which I draw from Newman's presentation of POCS. But this is not the important thing. The important thing is that we are free to choose the number of relationships that we want to express directly by special means in our information language.

This is not new, of course, for those of us who are linguists, and especially structural linguists, and especially readers of the famous American linguist, Sapir. Nevertheless, it is a very important idea (which has been neglected by many authors who discussed how many relations there are) that we are free to determine what kind of relationships we need to express by grammatical or syntactical means. That is the core of the question. As the linguists know, every language has a different set of relationships expressed by grammatical means. It would be very surprising if the artificial information languages which, after all, are modeled on the natural languages, did not conform to this pattern.

Perreault has insisted on this point with a very good sentence on p.183 of his report. He says that "If the discourse can be reduced to its essential, or at least centrally thematic, pure conceptualities[2], then a similar reduction should be possible with the syntax".

Of course, this is quite evident. The maker of information retrieval languages, or artificial languages, generally speaking, is free to reduce the number of relations he thinks it necessary to express by special syntactical means and by special notation. Pagès says exactly the same on p.354 of his report.

1.5 What types of relationships exist?

Here we have to deal with - let's take first what is unimportant, contrary to the POCS system which takes the important things first, - the distinction between analytic and synthetic relationships. Well, Newman reminded you and me that he was one of the first, if not the first, to point out that this distinction was more or less artificial. You will find a remark concerning that point by me on p.170 of the Elsinore Conference.

Coyaud retains the distinction, but says that it was more or less an artificial one and that it could be dispensed with. I cited his passage in my book, and perhaps it is not wrong to cite it as he is not available in this country. He says that the difference between analytical and synthetic relationships is not necessarily a difference of nature but only a difference of view. I remarked before that I would drop the word 'necessarily' from this expression. You heard a very good discussion of this point by Soergel. He, in my opinion, definitely proved that this distinction is to be rejected. Lévy took an intermediary position in his report, p.324. He said that they would be compatible; I think this distinction is purely formal, with no foundation in the essentials.

But there is a much more important distinction between paradigmatic relations and syntagmatic relations. You will observe that in this conference certain authors avoided speaking of syntagmatic relationships and spoke essentially on paradigmatic relationships. In fact this curious language comes from the linguists. It comes from Fernand de Saussure and has been taken over by all the structuralists who followed him.

You will find in Benveniste's presentation an excellent definition of it. "The units of the language are placed in fact on two planes, the syntagmatic plane where they are envisaged in their relation of material sequence within the chain of discourse, paradigmatic when they are put in relation of possible substitution, each one at its

level and within its form-class. To describe these relations, to define these planes, is to describe the formal structure of the language".

Then you will find in Bochenski a very good figure, Fig.1, on p.33, concerning the three dimensions of the sign. I think it is taken from Charles Morris, but I am mot quite sure. He does not give his reference. It seems to me that I have seen this in a book of Charles Morris which was published by the University of Chicago Press in the *Encyclopedia of Unified Science* some thirty years ago (6). He gives three dimensions of the sign. Two are of interest to us here, the third is not. The three dimensions make a cubic structure. What Bochenski calls 'semantic' is the paradigmatic; what he calls 'syntactic' is in fact the syntagmatic. There has been discussion of this by Soergel. I was not so happy with his discussion of this paradigmatic-syntagmatic question. (Some people of the SEMA, which is an organization resembling the Rand Corporation, Société de Economie et de Mathématique Appliqué, presented a model for the description of documentary languages. They had some unfortunate expressions for defining paradigmatic and syntagmatic; unfortunately these were taken by Soergel as the target of his discussion.) It would probably be profitable to discuss at length this question of paradigmatic and syntagmatic organization. This is a rather vital point in the organization of the structure of an information language, but will have to wait for a later conference.

1.6 What is the form of relationships?

Here we had a discussion which was not very illuminating. I must say, as it was full of confusion on the monadic, dyadic, ternary, or *n*-adic relation of factors. Perreault said that the roles were monadic so that they were to be coupled with links. This is on the expression plane, not on the essential plane (p.184 of his report). Perreault is searching for essence and Lévy is searching for practical things; that is funny as he is French and Perreault is American; the Americans are reputed to be pragmatic, the French are reputed to be theorists. Lévy says that the relationships in SYNTOL are dyadic, binary - he prefers this term - by rule, by deliberate choice of the makers of this language and he considers the practical advantage of this formulation (p.316 of his report).

Soergel says (on p.223 of his report) something which is quite evident, that all relations of a higher order than two can be interpreted as dyadic. Of course, this is a reality. I will remind you that a long time ago, I believe it was thirty years ago, in a book entitled *Traité de Logique*, by Jean Piaget, there is in my opinion a completely clear discussion of the role of dyadic classifications. He says that in certain mathematical structures - very highly organized - you may have any number of classes, but that in all other structures which are not so wholly determined as properly mathematical being you can always reduce all classifications to binary ones. We are referred to a role of expression; not to an essence but to a manifestation, and to a practical application which we can see later.

2. Relational Factors in Relation to their Milieu

2.1 Foundations of Relational Factors

Here we have many schools of thought. We have the philosophical school, which is represented here by Perreault, who says that the basis of the relational factors is philosophical, and who obtains them by deductive process. We have two representatives, or three, perhaps, of the psychological school who are searching for the basis of relationships in psychology, Farradane and Ceccato. (We can so class Ceccato, who reduces philosophy to psychology, he is seeking the foundations not only of relational factors but of all thought in his famous triads.)

Pagès is a professional psychologist but is not to be classed as psychologistic because of that. There are some qualifications to be made here. I think that we could find something like the ontological fallacy; I will try to explain later. I think also that we could find what we could call a psychological fallacy, as when he says that the negative concept has no basis in psychology but only in philosophical speculation, or that comparison is not a direct process or psychological process but a name of an entity or a process. I think this would be very debatable from the point of view of psychology.

Now we have the two other authors here who might be classed in the mixed class of logico-linguists. Lévy, half, as he has also one foot in one class and another leg in the other class. One leg is in the logico-syntactic class. On p.320 he says that the relationships are of the logico-syntactic nature, which is half logic, half grammar. But later on, p.323-324 he leaves the 'L.L.Camp' for the 'P.P.Camp', the Pure Practice Camp. Pagès is true to the Logico-Linguistic, or Logico-Grammatical camp and, on p.353-354 of his report, he treats the relationships as logico-grammatical processes.

Now, I would like to remark that it is very easy, too easy, to consider the relations as the categories of some ontological Being, with a big 'B'. This is what I call the ontological fallacy. You will find it very clearly expressed in p.2 of the Bochenski book, where he gives an ontological classification of categories; Being, with a big 'B', comprises things, properties, and relations between them. You will see we are ontologists; as Jourdain was 'making prose without knowing it' in Molière, we are ontologists without knowing it when we speak of things, properties, and relations. Being might be conceived of as an essence or as an existence, a design. We have two aspects, and here again you find the old distinction between processes and entities. Finally, if a being is modified in some way - I cite Bochenski - that is if a thing is red, or if one geometrical figure has twice the area of another - we are confronted with a state of affairs. It is very remarkable to find in Bochenski more or less exaxtly the categories of Gardin who, however, says that it was obtained by pure inductive process. It means that the Aristotelian and

Thomistic tradition is very strong indeed in the occidental world, including even Gardin.

Well, this is the ontological fallacy. Why? It was proved by Benveniste - following, by the way, other people, but is was a very good proof - in a famous article which was called 'Catégories de pensée et catégories de langue'[3] (which I think would be a very good idea to translate into English and to put in something like a *Reader for the Prospective Librarian and Information Scientist):* The Aristotelian categories, which were supposed to be categories of Being, with a big 'B', of reality, are nothing other than linguistic categories taken from the peculiarities of the Greek language. Benveniste generalizes this. I think that in this audience, where some of you are familiar with the Whorfian terminology, this will sound rather familiar.

Benveniste says in an article on the nominal phrase[4]: 'In linguistics a thing cannot have a universalvalue in the opposition between process and object. Neither a universal value, nor constant criteria; and it even has no clear meaning. The reasons are that notions like process or object don't reproduce objective characteristics of reality, but are resultant from the expression of this reality, which is already linguistic. This expression cannot be anything else but particular. The distinction between process and object is evident only for him who reasons upon the basis of the classifications of his native language, which he transposes into universals of language. This person, if asked "What is the foundation of this distinction?" will rapidly admit that if 'horse' is an object, and 'to run' a process - well, this is so in French, of course, but 'run' in English could be a process as well as an object, because one is a substantive and the other a verb. So we must avoid, or try to avoid, considering that any kind of relational system, any kind of categories we devise is something which corresponds to the reality of the exterior world - if there exists one (which, if I were a Berkeleian I would disagree with, but I am not). They are based on the peculiarities of certain languages. Here you will recognize the familiar tune of Whorf.

In the presentation of POCS by Newman we see that the empiricists of the Patent Office in the United States discovered this fact long before it was discovered by either Whorf or Benveniste. In fact, empirically, in their classifications they refused the usual categories. On this point I would be very glad if in the *Reader for the Prospective Librarian and Information Scientist* which I propose, somebody would take over the six or eight pages of Bailey's report of 1912, which are, as far as I know until now, the best presentation in the documentary literature of this fact that the so-called categories merge or are indistinguishable for practical as well as for theoretical purposes.

2.2 Classification, comparison, and standardization of Relational Factors

Here we have much been given in this conference by Perreault, by Soergel, and by others.

First, concerning Soergel's tables. I will not expatiate on Soergel's classification which, in my opinion, is not good. It is not because it does not comform with mine, but because it is somewhat confused. I will reaffirm that it is better to separate things which must be separated; that is, not to confound relations like mode or like categories of time, space, or persons, and so on with the real relationships which are ordinarily treated as such. But if you follow the classification of Soergel you will see that he seems to put the most debatable relationships at the beginning or at the end. The core of his classification is more or less coherent and in conformity with the classification which I gave. It is a pure coincidence, since he did not have my paper, as far as I know. But Soergel's classification is a sort of catalogue without any discrimination between the important relationships and the unimportant relationships, between the relationships which are widely accepted by the majority of systems as relationships and those which are *genus intrare* - that is the Latin jargon of Molière. In the extended table of Soergel you have a sort of weighting by the sheer number of people who have adopted such or such a relation. This thus eliminates some of the relations which are considered important by only one or two systems. However, the number of systems considered is too low for a true statistical measure.

Now for Perreault. I think that he was a bit carried away by his own enthusiasm. He maintained that, for instance, his ternary system, or my system, were in agreement, more or less, with Gardin, and with Hjelmslev. That is not true. A pure coincidence that de Grolier, Gardin, Hjelmslev, and Perreault are all of the ternary type of classificationists.

PERREAULT: Gardin is dyadic, as you said yourself.

DE GROLIER: Well, Gardin is half dyadic, half triadic. That is, he likes dyads very much, but after that he likes triads.

The way Perreault constructs his table is not quite right: there are four relations, but since Gardin calls the predicative the same as the coordinative (the latter applied to a special form class), you can say that it is a triad based on two dyads. Mine too, as well as Hjelmslev's.

I took the pains to find the original Hjelmslev text (7), and what he calls there 'dependence' (mutual dependence), is that in which the one term presupposes the other, and vice-versa. What he calls 'determination' is unilateral dependence, where one term presupposes the other, but not vice-versa. And 'constellation' or free dependence is two terms compatible but neither presupposing the other - well, this triad is made by two dyads. He explains this at length on p.35 of his book and I will not repeat this.

The fact is that my classification is completely different from Hjelmslev's; it has not the same basis nor

the same object. Hjelmslev is classifying syntactic *things*, syntactic *forms*.

As for Gardin, you may read my report for AFOSR -and, of course, Perreault read it, but he simplified, and in my opinion he oversimplified, in saying that my classification was more or less the same as Gardin. It is not true; it has not the same basis. When you examine the two in detail you will see that my 'object in itself', which is probably not good, is part of the coordinative, but that another part of Gardin's coordinative is classed by me in 'object in active relationship'; the associative part matches all right, but the consecutive is divided and so on. So this apparent convergence is either due to mere appearance or to the application of this ontological linguistic fallacy which I mentioned before, or, still a valid hypothesis, to the use of common sources. That is the hypothesis which is presented by Coyaud and which I cite in my AFOSR Report, 1.192. That is, all the systems which have been presented until now have the same common ground which is half linguistics - the current practice of natural language - half ontology, and a bit of logic.

So this convergence has nothing to do with reality, but is has to do only with some peculiarities of the people who made the systems. This is not to say that this comparison is of no utility, but this is to warn you again of the ontological or linguistic fallacy, that is, of considering that these classifications have anything more than a relative value.

2.3 The 'niceties' and 'shades' of Relational Factors

I regret that I was unable to fulfill my intention of discussing Farradane's report; I am obliged to discuss it now, but very briefly. I think that we must keep in mind a warning which Newman gave us when speaking of POCS, that you must avoid niceties as far as possible to be able to avoid the non-mutually exclusive sets of relationships. I don't know if this is possible. It is an open question, for me at least: If anybody in this audience can give me the proof that we can't place all the relationships which exist, or may exist, in a set of mutually exclusive classes, I will be glad and will most sincerely thank you. Until now nobody, to my knowledge, has given this proof. Anyway we must try to combine relations, so far as possible, to organize a set of mutually exclusive classes. Lancaster gave very good proofs of the fact that this is not the case with certain systems like the Engineers' Joint Council, or Costello, or others.

I think that Farradane has not avoided niceties. I think indeed that it would be very difficult to avoid the subleties' giving rise to incoherencies and inconsistencies in use.

Now I would give another warning. Many people here seem to consider that language is unruly. Even Newman has said something on the unruliness of English. Well now, natural languages ar not unruly. They have rules which are highly complicated structures, which nobody until now has been able to discover in all their subtleties - but the rules exist. This has been proved among others by Whorf when he spoke of what he called the subtleties of English due to covert structures. The natural languages have an infinite number of covert structures which it is very difficult to distinguish and to apply. We must at least try to avoid our information languages' acquiring (through the use of subtle classifications) niceties of classifications, the same covert structure, the same complications, the same apparentlyunruled or unruly structure.

3. Practical Applications

3.1 Do Relational Factors have any use at all?

The first thing is: for what use to we need or not need - which is a big question - relationships, relational factors?

Soergel made an assertion which unfortunately he did not prove; on p.227 of his report he said that the expression of relationships was necessary for avoiding the shortcoming of subject headings; but this is merely an assertion without any kind of proof. I would not defend subject headings, but it has to be proved that the addition of relationships overtly expressed would add something to their practical value. It has not even been proved, as far as I know, that overt expression of relationships increases the value of descriptors if these descriptors are merely linked. The contrary has even been proved twice: once by Lancaster, and again by Lévy. (I would not say that this is a condemnation of the expression of relational factors by other overt structures. I would not say either that the aim of this method is to avoid shortcomings of subject headings or descriptors.) Gardin has often written that he added this feature of relationshsips to make his classificatory language more flexible; that is a good idea. He called his system 'free', syntactically free, I believe. He opposes this to the systems which are syntactically fixed, with fixed structures. It is true that the expression of the most important and most general relationships by syntactical means does impair the flexibility of the system.

Farradane, as I told you already, gives a very exciting possibility by enabling subtleties of meaning to be exhibited reproducibly; in fact, we can play with relationships. It can be a very nice toy, a very nice game, with which we rather old people can play like little children play with balls or animals or pets. In fact the relationship can be considered as a pet; but I don't know if this is a game we are playing. For my part, I doubt it. If we go that far, we will admit all the complexities, niceties, and subtleties of the covert structures of the natural languages, since the natural languages have precisely been made by man for expressing subtleties of meaning to be exhibited reproducibly. So we fall into another linguistic trap.

Now, Coyaud uses relationships in what he calls automatic analysis (or abstracting), which I call automatic (or semi-automated translation from more or less

ruly English or French (that of the abstract) to more or less rationalize and standardize information retrieval language. For that purpose the results are very encouraging. The results until now are that using only these semantic tools of notional networks (taken from Ceccato and more or less adapted to the Gardin-SYNTOL main scheme), Coyaud arrived at a result of around sixty-five per cent good results of automatized translation. This is not bad. There are tests and experiments in process for improving it by the use of rather simple syntactic means.

Bouillut used this technique for content analysis; in this country Stone and Scheuch and others at the University Consortium, so called, are trying to use these techniques in a special kind of content analysis: the analysis of data-archives information.

Soergel made an interesting proposal which, I don't know why, he was subsequently rather shy about; it was to apply these things, roles, to what I call the deviant cases: This is a rather interesting possibility: the use of these roles, or indications of relationships, when, and only when, and if, and only if, it is necessary for distinguishing a deviant from a normal case.

Now we have the people like Mooers, up to a certain point, and Dale, up to a certain point too, who think that after all these things have no use at all. That was for me the meaning of Mooers' discussion on Soergel, and Dale is apparently of the same opinion. I am not in agreement except that I recognize quite clearly that it is not necessary in many cases such as those you cited, in which it is unnecessary to enter into these subtleties or niceties which almost inevitably accompany the expression of relationships by formal, overt means.

3.2 Expression of Relational Factors

Now the ways to express them. Here I will be very brief as this was practically not discussed at all, though there was some discussion in Pagès report. There was, of course, the example of Farradane, who expresses some of the relationships by his operators, that is by special signs, and some others by categories. (I don't know what kind of notation he applies to these last, but this subject was apparently outside the main field of his interest. I rather regret it as I think that this is a very practical thing to envisage.) What are the best ways to express relationships, beginning with the mere implicit expression which I find in the POCS Schedules? It is not true that POCS is an unrelational system. It is one, but the relational structure is covered. It is hidden in the explanations; in the rules for hierarchies; in the subordination of classes. It is more or less a faceted classification. Here I am not in agreement with Lancaster. It is not so far from the English Electric Classification. It is not apparent in the notation, nor even in the tables; the people of the Patent Offrice seem to have a predilection to make very complicated schedules which are, as far as I am concerned, almost impenetrable, but there is a faceted structure inside of it.

Now you have the overt faceted structures. You have the roles - monadic, as Perreault ably defines them - accompanied or not with links. You have the binary relations of SYNTOL. You have the multi-arguments of Pagès, etc.

I will give you only one reference. In his famous book on language (8) Chapt.6, I believe, Sapir made a classification of languages which is ably reported by Benveniste. This is a very good classification, the best we have until now, which defines languages according to what he calls 'the concepts'. All languages have (1) basic concepts: independent words, objects, actions, qualities and so on; and (4) abstract relational concepts: purely formal relations which construct the syntax. Every language which is known contains at least these two classes of words. Other languages contain in addition what Sapir calls (2) derivational concepts by affixes and (3) concrete relational concepts. To these belong the categories of number, gender, and so on.

He classified the languages in four classes; Class A is that based on the use of categories 1 plus 4 (basic concepts plus abstract relational concepts). In this class you have, for instance, Chinese; this is the reason why Chinese is the language best adapted to information retrieval. In Class B you have languages like Turkish which employ categories 1 (basic concepts, 2 (derivation), and 4 (abstract relations). You have a Class C of which the only representative is the Bantu languages which contain the basic concepts plus the concrete relational concepts (the categories of number, gender, etc.). Finally, you have the large Class D to which belong the majority of our Western languages and the majority of the Hamito-Semitic languages like Latin, Hebrew, etc.; these employ categories 1, 2, and 3 (basic concepts, derivational concepts, and concrete relational concepts). French and English are between C and D.

Now you can have exactly the same classification according to the form of relational expressions in the artificial information languages; this is the proof that these languages are not outside the general family of language. This is also the proof, by the way, that this three days' discussion was directed at the core of the information languages, as you see that Sapir's classification of languages is based on the way languages express relationships. That is the core of the matter.

So we could try to classify the ways to express the relationships within the frame of reference of Sapir's classes and see which is best. In my opinion, the more we approach Class A (basic concepts plus abstract relationships), the better is the mode of expression. But this is an opinion still to be proved.

3.3 Experimentation and testing of Relational Factors

Now the last question, relating to experimentation and testing, will be treated very shortly since it was also outside the field of our meeting. We heard of some tests,

very few so far. Lévy's test for SYNTOL, which was negative. Melton's tests, negative, Lancaster's tests, negative. We await Farradane's test, which I hope will not be negative. We have some results of Cleverdon's tests, but the conditions in which they were conducted makes it very difficult to know if there is something to be derived from these results for judgment on relations.

Now here we have two theories: one, which says with Pagès that no direct, empirical validation of systems is possible. I hope that he will temper this in the future, otherwise one of my main claims, which is that we could prove or disprove the relative value of certain systems of classification, of information-language structures (and more precisely, the expression of relational factors) would be invalidated. I would be very unhappy.

Now, some tests I regret to say are impossible. I think that this is the case for the test Bregzis proposed on the Ceccato system. I would be very glad to be wrong, but I think you could not validate Ceccato's scheme by any application to information retrieval; this seems to me to be a formal impossibility. But I will not delve into this now.

Now you have an affirmation, which is Borko's, that no one system is intrinsically better than the other. Well, of course, we could discuss 'intrinsically', but I think that this affirmation is not proved. It would be very useful to prove or disprove it.

4. Conclusion

Finally, in conclusion, I will say that this field is most clearly an inter-disciplinary field; for advancing in it the contributions of many other disciplines are needed. I deny that documentation is a discipline; on this point I am absolutely against Ranganathan. It is not a discipline; it is a technique. As in every technical field which has some reputation, we have to draw on other scientific disciplines. We have to draw our terms from philosophy, and that was a very good beginning with our friend Perreault, and it is refreshing that a philosopher has been able to organize this meeting and to place the queen of the sciences, philosophy, in an assembly of such specialists from many countries of the world and from many disciplines as well.

I would regret that linguists were so few here, as we would have profited very much from their presence. Nevertheless, we have some logicians, some psychologists, and this proved the interest of this meeting to be of an inter-disciplinary nature. The National Science Foundation must be congratulated for that realization too, and the University, of course. Here I remember that I am an FID official and so I give flowers to everyone. After a good deal of frozen water, which I put on the head of everybody, now it seems I change my role. So my role is to give flowers, too. Also, I think this meeting has proved the interest of the specialized approach contrasted with the general approach which was that of the First International Conference at Dorking and of the Second at Elsinore. This is not to say that the generalized approach was necessarily wrong. It was useful; now we need specialized approaches and the consideration of specialized parts of the general structure, and the more specialized, the better it will be.

Now I will end. You know that Cato the Old finished every discourse with 'delenda est Carthago' - Carthago must be destroyed - and I finish all my interventions everywhere by saying that it is necessary to continue research. This is, of course, addressed to the valiant representatives of the monitoring agencies, as it is a plea for money. I think that we must continue research on what I call the two planes of research: (a) the practical application of relational systems; the practical devising of them; testing; and experimenting with them; and (b) the research into the foundations of the relational factors' structure, the study of which I have more or less treated today.

I will finish with a little joke. I will return to a suggestion of Pauline Atherton's: she said that classification was dead and so, let us revive classification. You know that many years ago Nietzsche said that God was dead. This was, of course, momentous: God was effectively dead, up to a certain point. But he was reinvented, not so many years later, under the new name of 'Point Omega' by Father Teilhard de Chardin. Immediately, for all the atheists, who are rather common in Europe, from the mere fact that God was rechristened 'Point Omega', it became a reputable thing to be, not an atheist, but an admirer of Teilhard de Chardin, thus admitting the existence of Point Omega. It is the same with us. We have changed words, in a way. We don't speak of classification but of 'relational factors'; that is much more fashionable. We don't speak of library systems, we speak of 'information control' that is much, much better. We don't speak of subject headings, we speak of 'information languages'; we acquire on the spot the dignity of something much better. In this way we could say that if classification is dead, all right, let's revive it.

Notes

1 I know why he did not want it to be published. First, because he was not the only person responsible. Second, the paper was in a rather tentative stage, and the determination of these categories or parts of speech was apparently a rather difficult task, and even more difficult to explain in a rational manner.
2 I don't know what 'pure conceptualities' are, but that is another matter. Let us agree on something which could be termed 'pure conceptualities'.
3 See Benveniste, (3) p.63-74
4 See Benveniste, (3) p.91-167

References

(1) Grolier, E.de: On the theoretical basis of information retrieval systems. Paris 1965. AFOSR Report AF61(052)-505), 1.154

(2) Grolier, E.de: A study of general categories applicable to classification and coding in documentation. Paris: UNESCO 1962.

(3) Benveniste: Problèmes de linguistique générale. Paris: Gallimard 1966.

(4) Bochenski, I.M.: The methods of contemporary thought. (tr.P.Caws) Dordrecht: D.Reidel 1965. p.44

(5) Atherton, P.(Ed.): International Study Conference on Classification Research, Elsinore 1964. Copenhagen: Munksgaard 1965. p.161-168

(6) Morris, Ch.: Foundations of the theory of signs. Chicago: Univ.of Chicago Press 1938.

(7) Hjelmslev, L.: Prolegomena to a theory of language. (tr.F.J. Whitfield) Madison: Univ.of Wisconsin Press 1961.

(8) Sapir, E.: Language; an introduction to the study of speech. New York: Harcourt, Brace & World 1949.