Published in LSP & Professional Communication Volume 1, Number 1, April 2001 - ISSN 1601-1929

Diachronic Analysis of the Visuals in the Research Paper: A Corpus-Based Study of the Strategies and Semiotics of Visual Representation in Nutrition Biochemistry

Anne Magnet
University of Burgundy
Dijon, France

1. Introduction

There has been relatively little examination of the visuals used in the scientific research paper, as opposed to studies of written texts, in the field of ESP. Johns (1998:183) explains :

> There are at least two reasons for this concentration upon the written word. One, certainly is that applied linguists and compositionists are more interested in, and comfortable with, written language. Most of us are trained in the humanities, where words are central to disciplinary values and argumentation. A second, related reason is that in some academic classrooms and disciplines, such as literature, graphs, charts and other visual representations are not central to disciplinary values.

However, Huckin (1987), Bazerman (1989), Berkenotter & Huckin (1993), Myers (1995), and more especially Miller (1998) and Johns (1998) have committed themselves to the study of visuals in the scientific article, as part of their professional involvement, in the disciplinary culture. It could however be argued that non-verbal communication has enjoyed an explosion of interest and use in the second half of the XXth century, due to the exponential development of image technology. Visuals should therefore be present in the research paper as they are in any other media, and there is certainly no neglect of visual representation in other disciplines. Yet, the issue of how scientists use non-verbal elements emerges as a pertinent area of research in ESP and more specifically in the study of the research paper since :"visual elements in the form of figures, tables occupy from one third to one half of the page of a typical research article" (Miller, 1998 : 29) and since, over the years there has been a noticeable increase in the size and importance of visual elements. (Bazerman, 1989, Huckin, 1987, Myers, 1994).

The aim of this paper is thus to report the results of a corpus-based analysis of visual representation in the research paper, in the form of figures, tables, charts, graphs, diagrams and photographs with a diachronic prospect. This work was undertaken in the light of two theoretical foundations offered by the sociology of scientific knowledge represented mostly by the socio-constructivist authors (Latour & Wolgar, 1979, Knorr-Cetina, 1981, Latour, 1990, Lemke, 1998) and a socio-linguist (Martin, 1996) and the social semiotic theory of representation initially borrowed from linguistics. Quantitative data are briefly reported and the basic concepts of social semiotics are recalled before being used as tools for the present analysis and the discursive strategies underlying the use of visual representation in the research paper are also disclosed.

2. The data sample

The data for this study was drawn from a corpus built around the single biochemical research theme of *Essential Fatty Acids* (EFA), as investigated by a community of biochemists and nutritionists from the earliest paradigm set up on the theme in 1929, which established that the consumption of fat is necessary to the survival of mammals. The selection of 72 articles was carried out through multiple processing :

- Sorting of the authors : only papers published by laboratories located in English-speaking countries were chosen.
- Sorting of the journals : only articles published in the top 5 Biochemistry and Nutrition Journals, according to the Science Citation Index (SCI) were kept.
- Representativity of the articles : the 240 papers selected by the ESP linguist were submitted to 2 scientific specialists for selection on a content-based procedure (for details on the constitution of the electronic corpus : Magnet, 2001a, forthcoming).

The most represented journals in the corpus are : *The Journal of Nutrition, The American Journal of Clinical Nutrition, The British Journal of Nutrition, Lipids, The Journal of Lipid Research, the Journal of Biological Chemistry*, which all enjoy high ranking according to the SCI. The 72 articles were then grouped into periods named «before 1960», « the 60s», « the 70s», « the 80s» and « the 90s», to give the corpus a 5-entity outlook and thus a more user-friendly approach, knowing that the results reported here can only give evidence of trends in the use of visuals over the 70-year period studied, within the specific field of nutrition biochemistry.

3. Discursive strategies underlying the use of visuals in scientific discourse

Diachronic studies are making a shy come back in linguistics and emerge in ESP (Salager-Meyer, 1997, 1999) to show that if the Scientific Experimental Article is to be understood and analyzed as a genre, it should not be inferred from it, that such a format is a stable genre, with unchanging and immutable laws and conventions. A diachronic study may make it possible to account for the dynamic process involved in scientific writing within a specialist community. The scientific research paper is not only made up of continous verbal information. The text is interspersed with graphic representation such as tables, curves, diagrams and possibly photographs, mathematical formulae or equations. Non-verbal features are widely used by specialists and one single question to be raised about them is : «what drives scientists to choose *visual* rather than *verbal* communication skills in a paper, since either is available ?». The quick answer points to genre conventions, which may satisfy a «how-question», but not the «why-question», to paraphrase philosophers of science (Lambert & Shurz, 1994).

Another possible explanation is to interpret non-verbal representation in terms of efficiency. Lemke (1998:87) states that science essentially requires the use of various media to operate :

> The concepts of science are not solely verbal concepts, though they have verbal components. THey are semiotic *hybrids*, simultaneously and essentially verbal, mathematical, visual-graphic and actional-operational. The actional, conversational and written textual genres of science are historically and presently, fundamentally and irreducibly, multimedia genres. [...] Language as a typologically oriented semiotic resource is unsurpassed as a tool for the formulation of difference and relationship, for the making of categorical distinctions. It is much poorer in resources for formulating degree, quantity, gradation, continuous change, continuous co-variation, non-integer ratios, varying proportionality, complex topological relations of relate nearness or corectedness, or nonlinear relationships and dynamic emergence (which I refer to as the *topological* dimensions of meaning).

Graphic representation transposes scientific discourse towards different symbolic media. The decision to use it or not, the choice of its positioning through the text or in various sections of the paper, the specific type of representation selected as well as the information delivered or supported, all play active part in the «construction of the scientific fact». The use of visuals in the research paper can be interpreted as one of the discursive strategies displayed by the specialist in the composing process of the research paper (Latour, 1979). This amounts to identifying the specialist as an active participant in the making of science and not as a faithful but passive observer of reality. The scientist actually «constructs» the visual elements in the same way as he/she «builds an ideal experiment» reported in the research paper.

However these items should *not* be understood as *images* or *illustrations*, as is the case in other media. Although these terms are used frequently by semiologists, they will not be used here since they could be misleading in relation to the nature of the scientifc paper. In fact, visuals are to be considered as the *hard core* of the experimental article in science, which allows researchers to «create order from disorder» (Latour, 1979: 235), or to «rearrange initial chaos represented by the data produced in the laboratory during the numerous experiments which will not all be reported, or more widely represented by the sum of knowledge of the community on the same theme» (Prigogine & Stengers, 1984 : 239-266). This is nicely summed up by Miller (1998: 31): «when messy nature is reduced to figures and numbers, it can be described with more precision and confidence». Latour (1990:39), describing the activity of a research laboratory, underlines the absolute necessity for the research team to produce graphic inscriptions from the manipulation of reality :

> No matter what scientists talk about, they start talking with some degree of confidence and being believed by colleagues, only once they point at simple geometrized two-dimensional shapes. The 'objects' are discarded or often absent from laboratories. Bleeding and screaming rats are quickly dispatched. What is extracted from them is a tiny set of figures. This extraction.... is *all that counts*. Nothing can be said about the rats, but a great deal can be said about the figures.

The biochemists and nutritionists who wrote the corpus articles on the theme of Essential fatty acids (EFA) are not different from other scientists and they seem to follow a similar mode of action : what counts is the extraction of quantitative data thanks to which *visual persuasion* tools can be produced.
They are not only to be understood as an essential component of the research paper but more so as the hub around which the text is organized. Huckin (1987:5) notices that some expert readers begin by deciphering the visuals before deciding to read on :

> When I asked the scientists to demonstrate how they customarily read a newly published article in their field, they all displayed a reading pattern dominated by the search for new information. First they read the title, then the abstract, then they looked for most important data, usually in graphs, tables, drawings and other visual aids.

Let us now proceed with the evolutionary description of the space, number and nature of visuals and the analysis of these tools in the corpus presented above.

4. Evolution of non-verbal elements in the research paper on the EFA theme (1929-1999)

4.1 Space assigned to visuals

Visuals take up from one third to one half of each page of scientific papers in the journals *Science* and *Nature* (Miller, 1998:29). Lemke (1998) reports an average use of six visuals per page in *Science.* Huckin (1987), Bazerman (1989) and Myers (1984) all noted increased size and number for visuals in scientific information. The increased use of visuals over the years is however not observable in our corpus.

Table 1 : *Evolution of the Text/Visuals/ Bibliography Distribution in the Research Paper on the Theme of EFA from 1929 to 1999*

Corpus (72 articles)	*Text (% of article)*	*Visuals (% of article)*	*Bibliography (% of article)*	*Article (average length in words)*
Before 1960	64.5%	27.8%	7.7%	4056 ± 2320
1960s	61.7%	30.2%	8.1%	3490 ± 1030
1970s	65.4%	26.3%	8.3%	3818 ± 1241
1980s	58.4%	30.1%	11.5%	3593 ± 772
1990s	60.3%	25.5%	14.2%	5071 ± 1613

Indeed, visuals already represented 27.8% on average of each article in the period before 1960. The space occupied by these elements in relation to the paper as a whole varied between 25 and 30% and still represented more than a quarter of the article in the 90s' corpus. If no significant fluctuation is observed, it is due to the already large space allotted to these elements from the beginning of the corpus. It is interesting to notice that in the 60s and 80s, this percentage went up to more than 30% of the article, in spite of the urge by publishers to limit the use of visuals, obviously for cost reasons :

> Authors are urged to economize on space used for tables and figures. These should fit one column width (2 5/8 inches) or when necessary, two column widths (5 1/2 inches). **A charge will be made** for that space used for tables and figures which exceeds one-half of the space used for the manuscript exclusive of tables and figures. *The Journal of Nutrition, Guide for Contributors to the Journal, volume 7, 1960.*

The corpus presented here is thus largely characterized by a strong presence of non-verbal elements. The increasing importance assigned to the bibliography section and references in the article led to the shortening of the written text (only 58.4% of the total article in the 80s), but was without effect on the space allocated to visuals (30.1% of the total article in the 80s). This heavy use of non-verbal elements needs to be taken into account as one of the tools for the construction of the scientifc fact through the scientific paper. The results shown in Table 1 go beyond the examination by Myers (1995:113), and therefore justify his point in an even stronger manner :

> Scientific articles may have as much as a fifth of their space taken up by 'pictures'. Clearly these 'pictures' are doing more than just illustrating, supplementing and breaking up the dense blocks of text and attracting the attention of any reluctant readers.

4.2 Number of visuals in the research paper

If the space granted to non-verbal elements has remained high and relatively stable in the course of the last seventy years, it might also be revealing to investigate whether their number in each article has evolved or not. Indeed, the space used may not actually reflect the number of visuals, owing to the variations observed in the size and scale chosen by the authors or the publishers of the journals.

***Table 2** : Evolution of the Number of Visuals in the Research Paper on EFA from 1929 to 1999*

	Before 1960	1960s	1970s	1980s	1990s
Number of Visuals (mean /article)	7.2 ± 3.2	6.5± 2.9	6.15± 3.12	5.7± 5.8	7.2± 2.69
Tables	3.8± 2.78	3.3± 2.4	4.9± 3.2	4.2± 1.81	3.7± 2.05
Figures/Diagrams	3± 2.79	2.7± 1.18	1.38± 1.68	1.3± 1.20	3 ± 2.49
% of Tables relative to all Visuals	52.7%	50.7%	79.6%	73.6%	51.3%

Our corpus exhibits an average of 7.2 ± 3.2 visuals per article in the period before 1960 and 7.2 ± 2.69 in the 90s, with a slightly lower number in intermediate decades (around 6 per article). These results agree with those observed by Lemke (1998) in the Journal *Science*. The biochemists and nutritionists who wrote the corpus articles seem to be representative of the larger scientific community in their use of what seems a rather conventional number of visuals in the research paper. Through this, scientists appear to follow a basic strategy in their activity : what counts is the extraction of figures, thanks to which they achieve visual persuasion . The various modes of visual inscription selected by the specialists to convince their peers of the validity of their scientific claim now deserve investigation.

4.3 Nature of the Visuals in the research paper

Once evidence of the essential and constant presence of visuals throughout the corpus has been brought, a more intriguing question needs to be addressed concerning the types of visual representation chosen by the community of specialists on EFA, the reasons for these choices and the function of visual representation in this specialized discourse. Tufte (1983:13) made an inventory of the possible types of graphical displays and attributed various roles to them :

Graphical displays should

- show the data,
- induce the viewer to think about the substance rather than about the methodology, graphic design, the technology of graphic production, or something else,

- avoid distorting what the data have to say, present many numbers in a small place,
- make large data sets coherent,
- encourage the eye to compare different pieces of data,
- reveal the data at several levels of detail, from a board of overview to the fine structure,
- serve a reasonably clear purpose : description, exploration, tabulation, or decoration,
- be closely integrated with the statistical and verbal description of the data set.

This investigation applies to the delivery of an informative or didactic message but it does not take into account the polemic and agonistic value of scientific discourse. Within the framework of a socio-constructivist approach, visuals can be understood as a way not only to show but to demonstrate, not only to describe but to bring evidence:

> Visuals in academic articles provide data to convince the reader of the validity of the findings and allow the readers to see how the data were obtained and to interpret the data themselves. These visuals are impregnated by theory to show not only that they are anchored in the literature but that they have wider implications (Miller, 1998:31).

In our corpus, visuals take the form of (in the decreasing order) :

- Tables, reporting mathematical results derived from selected experiments (not all) carried out in the laboratory.
- Diagrams, representing mathematical data turned into graphs to emphasize the correlations or the changes.
- Bar charts, which demonstrate the evolution of the parameters chosen and enable their comparison.
- Photographs, (in the period before 1960) providing evidence of the physical and physiological consequences of EFA deficiency at the level of tissues or cells. Later, photographs became scarce and were used to highlight the high degree of precision obtained in the separation and isolation of various molecules through chromatography.
- Autoradiography or prints obtained on a photographic emulsion by radioactive products.

Visual representation in our corpus is mostly in the form of tables and figures which are mainly bar charts (as early as 1930) and curves. These two types of visuals (tables and figures) represent at least 92% of all non-verbal items, the remaining part being photographs or diagrams. Unlike other fields of research in science (Bazerman, 1989:172-173), nutrition, and the specific theme of EFA in particular, does not show any shift from the use of tables to that of diagrams. Tables displaying mathematical results remain the leading non-textual medium (3.8±2.78 tables before 1960 and 3.7±2.05 in the 90s). The number of figures, mainly bar charts and curves used in the research paper has varied along the years but we see in the 90s (3±2.49) a return to the numbers used before 1960 (3±2.79). However this relative stability should not be interpreted as a sign of stagnancy in the strategies underlying scientific activity at work in the research paper. Closer investigation needs to be brought to the contents of these preferred visual aids.

4.4 Role assigned to visuals in a socio-constructivist approach

4.4.1 Role versus nature of the visuals

Before 1960, the production of quantitative data was limited, due to the difficult production of reliable data for technical reasons. This was of course not only characteristic of the research on EFA, but of all biochemical investigations. The development of laboratory technologies made available to biologists and biochemists should have led, in the research paper, to an extensive display of the results, obtained through sophisticated methods of isolation and purification of the molecules studied, i.e. : essential fatty acids. However, around the 70s, what became essential to scientists was the *number of papers* published in order to promote their careers which were assessed in these terms. The tendency grew towards splitting the results of a single experiment into several articles. In other words, it became more rewarding to publish results in a number of articles containing each an average of 3/4 tables rather than a single article containing 10 tables. Besides, in the *Guide to Authors* of the Journals mentioned above, the publishers kept urging scientists to limit the number of figures and photographs for obvious financial reasons. Moreover, the cost of publication born by the publishing laboratory strictly depends on the number of pages of the paper, which is also a good incentive to keep space-consuming visuals within sensible limits. Cost constraints also account for the absence of colour in the visuals of most scientific journals (*Science* and *Nature* being strong exceptions in this respect).

The predominance of tables, followed by diagrams and bar charts in our corpus is also due to the specificity of the theme. EFA only draw their existence by being brought to the fore through laboratory assays and tests and therefore by their quantification, after separation or isolation from other tissue or blood components. And even more particular to the theme, due to the relation of competition between the various fatty acids observed in the metabolism of mammals, any variation of one parameter is followed by changes in others, which drives specialists to produce the whole set of values and not only those on which the experimenter focuses. The field experts consulted to build the corpus were of the opinion that the tables presented in a scientific paper are used as a reference, as a control by future experimenters. If values obtained in a laboratory under the same experimental conditions differ too much, they will be checked over and over in order to consider a possible controversy over one of the community's paradigms. This would result in a new scientific claim. The large amount of mathematical data shown in tables, induced by the specificity of the theme, and also an illustration of general science in the course of the XXth century explains the recourse to tables as a privileged data display tool. Tables represent an *intermediate form* between text and pure graphics. A table does not call for compulsory linear deciphering, and yet it tends to imply the conventional reading move from left to right, line after line.

The display of numerous data in tables also confers another specificity to the EFA specialist community. Unlike other scientists who start their reading of a scientific paper by browsing through the visual aids, biochemists and nutritionists specialized in EFA, tend to read the Abstract and the Results section which bring to the fore the most striking facts, before perusing the visuals.

4.4.2 Role versus positioning of the visuals

The diachronic prospect of this study makes it possible to view the strategies developed by a specialist community as a dynamic process. To our knowledge, no investigation of the distribution of visuals through the various sections of the article has been done. If visuals are to be considered as the *hard core* of the paper, their arrangement throughout the article and among its various sections may not be haphazard and the study of this arrangement is probably of interest in the understanding of general discursive strategies.

It should first be recalled that authors rarely decide where visuals will be placed within the article. Indeed, they are asked to produce all the visuals on separate sheets and to note where they wish to have these elements inserted. The final decision falls to the publisher of the journal. Should the role of visuals be only to give the results a mathematical or graphic inscription, they would only appear in the Results section of the article, but Table 3 displays a wider distribution of non-verbal items throughout the article, which calls for more investigation.

Table 3 *: Evolution of the Distribution of Visuals in the Research Paper on EFA (expressed in %) from 1929 to 1999*

Section of the article	Before 1960	1960s	1970s	1980s	1990s
Introduction	1.4%	0%	0%	5.8%	2.5%
Materials & Methods	32%	9.2%	12.5%	15.8%	15.9%
Results	41.7%	52.4%	67.5%	43.4%	55.8%
Results & Discussion	20.8%	33.8%	8.75%	10.8%	0%
Discussion	4.1%	4.6%	11.25%	24.2%	25.8%

First and contrary to what might be expected, the Introduction sometimes contains visuals (in 10 articles/72, before 1960, in the 80s and 90s). In a few cases, the table providing the compositions of diets and their EFA contents in a kind of experiment called «mediated dietary experiments» by the specialists, has obviously been displaced from the Materials & Methods section to the Introduction for editorial reasons. But in other articles, the presence of a diagram providing a schematic visualization of biological metabolisms in the Introduction serves a particular purpose. It represents the condensed *macrotext,* that is the knowledge shared and referred to by the specialist community at a given time (Martin, 1996:17). The use of such a figure in the Introduction enables scientists to efficiently focus the reader's attention on a selected paradigm. It bars the reader's way to other possible choices of scientific interpretation on the specific theme of the article. This observation needs to be confirmed in other fields of research. Visuals can be construed in a scientific environment as a way to *materialize* and thus *reify* objects and processes. This accounts for the constant presence of visuals in both sections : (1) Materials & Methods and (2) Results. The contents of the first two tables, generally found in the Materials & Methods section, but sometimes placed at the beginning of the Results section, aim to assure the reader that the experiment has been done according to the procedural conventions of the community. They exhibit the data likely to be checked by other experimenters, such as the animals' weights, the composition of the diets tested, the length of the dietary protocols, etc. The last 2 or 3 tables placed in the Results section represent the *nucleus* of the paper, and the parameters reported justify the originality of the experiment. Thus, the importance or dramatization of the data is constructed and accentuated through the succession of visuals. Table 3 indicates a particularly strong presence of visuals in the Results section (41.7% before 1960 to 55.8% in the 90s). The increasingly heavy use of tables and bar charts in the Results section through the corpus also explains the regression of written text in this section. Visual representation of the results is more efficient and requires less effort in the selection of values for the process of *foregrounding*. A negative correlation is observed thanks to the diachronic investigation of the topographical use of visuals in the research paper. Actually, of the total number of visuals in any research paper, the percentage found in the Materials & Methods section halved from 32% before 1960 to 15.9% in the 90s. Over the same period, the percentage of visuals in the Discussion section increased dramatically from 4.1% before 1960 to 25.8 % in the 90s. This evolution might be due to the shortening of the text in the Results section, which renders the insertion of visuals difficult in this part of the article, but it may also be justified in terms of discursive strategies. The increased presence of visuals in the Discussion section and more particularly that of tables which display the most important parameters justifying the strongest scientific claim, stresses the fact that visuals actively support the construction of the argumentation in the strategy of persuasion. They are placed where they are the most effective, that is where they can be used as convincing pieces of evidence. They are used to prompt the community to ratify the answer to the hypothesis raised, and incorporate it into the macrotext, which is the major function of the Discussion in the article.

5. Semiological approach of non-verbal elements in the research paper

5.1 From object representation to process epitome

Survey of the corpus gives evidence of a striking aspect in the evolution of the use of visuals over a 70-year period : the move from object representation in the visuals to that of processes. Photographs were sometimes included in the articles published before 1960 (from half a page to a complete page). These photographic visuals represent one of the *external allies*, (Latour,1987) and are used as a guarantee (Bastide, 1985:137) :

> Dans le double mouvement des illustrations au texte et du texte aux illustrations, la photographie joue le rôle de garantie et assure l'enchaînement des différentes étapes du dispositif de visualisation, de sorte que celui-ci apparaisse linéaire et homogène.

In the 40s and 50s, articles also included pieces of realia such as follow-ups of patients or drawings of the body parts of children affected by EFA deficiency. This type of visual representation gave a genuine touch of true life to the scientific report. This type of authentic document extracted from real life disappeared almost totally from the corpus after the 60s. Photography became a rare visual tool in the corpus (5 occurrences for the rest of the corpus) and its nature changed. From the genuine representation of reality through visual exhibition of lesions due to EFA deficiency, it shifted to the manipulation of reality, as shown in electron microscopy photographs of cells or radioactive tracings printed on film, which require the use of sophisticated laboratory techniques. From the evidence of an *object* brought through photography, the laboratory work presented in the article has become the display of a process. The photographed physiological results of EFA deficiency have also been replaced, since the 60s, by the production of mathematical averages of a large amount of data provided in the form of columns of figures in a table, considered as a much more convincing piece of evidence in the persuasion strategy. A photograph is to be interpreted as the substantiation of a theory from one single case and therefore raises the question of possible generalization. Mathematicized results processed so as to build up tables or charts are less easily questionable and represent a more reliable and resistant *ally,* in the construction of *hard* scientific facts (Latour, 1979) :

> Photographs, while persuasive, are seen as unique examples of the phenomenon studied. The combination of visual and iconic however, make for a powerful argument (Miller,1998:36).

> The differences between photograph and diagram can be resolved by associating the photograph with the unique, situationally specific, perspectival, instantaneously, and particular aspects of the thing under examination while the diagram brings into relief the essential, synthetic, constant, veridical, and universally present aspects of the thing 'itself'. (Lynch,1990:163).

5.2 From description to the building of relationships

In the articles from the beginning of the corpus up to the 60s, scientists literally described their experiments and for instance each rat was *named* (e.g.: ‡ rat 29 541) and the values relative to each rat were reported in tables or represented in charts *individually*. From the 60s on, only means of values for parameters measured in groups of generally 5 rats are provided. The aim pursued by scientists has become the production of numerous compressed data and the comparison and the correlation of the various parameters chosen, in order to establish strict relationships between them, as Bertin (1981:176) states : «The aim of tables and graphics is to make relationships among previously defined sets appear». The numerical data displayed are therefore a simplified representation of reality. Some mathematical means may even be the mean of means, as indicated in the caption of a table :

The results are given as the mean ± S.E. of the mean for 8 rats on each diet. (*Biochemica Biophysica Acta*, 1988:504)

or of several experiments in another caption :

Values are means from two experiments . (*Lipids* : 1980:843).

Parallel to the sophistication of the data provided in visuals, the written text *inside* tables has been receding to the benefit of symbolic and numerical notations. For instance, the notation of linoleic acid, an essential fatty acid, was first replaed in tables and diagrams with LA and then with 18:2 *n-6*, or *18:2* ω*6.,* which makes understanding of the tables for a non-specialist totally impossible. For instance, in a table taking up more than half a page, only 4 (compound) lexical items can be found (*The Journal of Nutrition*, 1980:1700, table 4), all the remaining signs being symbolic and numerical.

The shift from object description to relationship building can also be observed in the *titles* of tables and diagrams. Correlation appears through the heavy use of the term : *effect of ,* (39 titles of visuals in the corpus), followed by *relationship(s), influence of, correlation(s), comparison of, changes in , response to , percentage distribution of, differences in,* and also, to a lesser extent : *plot of observed changes, concentrations versus predicted changes*. The interest for the process, rather than for the object is illustrated by the coining of the titles of visuals such as : *incorporation of , time course of, regression lines representing changes in, inhibition of,* which all indicate interference of the specialists with the object studied.

5.3 Semiotics tested in the research article on EFA

5.3.1 The referential power of visuals

Using traditional semiological criteria to make an inventory of, classify and describe the sign language offered by non-verbal representation, it is possible to study the function of these items more effectively by trying to understand what they refer to. According to the Paris School of Semiotics and more specifically to Barthes's theory (1964), image in the wider sense uses *indexical* reference, that is built on a direct link with the *referent*; index is understood as the apparent sign that something exists. An index would thus be: «the sign through which a sequential or causal relationship is built between *signifier* and *signified* »(Dyer, 1982:125). The function of image is thus to demonstrate, to bring evidence. The process of autoradiography as used in the corpus of the 80s, and 90s is a classical example of this type of reference used in biology. It enables visualization of molecules through photographic impression of a radioactive isotope.

Reference can also be *iconic*, that is based on resemblance. The *icon* can be seen as the sign in which the 'signifier-signified' relationship is a link of similarity or homology. Resemblance can be real as in photographs or stylized as in conventionalized signs.

Reference can lastly be *symbolic*, that is, built on conventions accepted by a community. This is the case, in our corpus, of the notation system of fatty acids as already observed in Tables. The diachronic investigation of non-verbal elements in the corpus shows limited use of indexical reference in the form of photographs, occurring mostly before 1960. Once the very existence of EFA had been accepted by the community, the purpose of the experiments evolved into demonstrating the origin, the variations in contents and the role of EFA in the metabolism of mammals.

Iconic reference is rather scarcely used since it is poorly suited to the description of dynamic processes. To conclude, symbolic reference is by far the most widespread referential process behind visuals used in the corpus. It is increasingly present, since symbols are constantly being created by the community, which thus builds its own code of communication. This evolution is probably not proper to the research paper on EFA, but it is precisely the gradual change from indexical to symbolic reference over the period studied that largely contributes to making the scientific article esoteric.

5.3.2 Social Semiotics applied to the research article

Another way of looking at the semiotics of scientific discourse is represented by what Halliday & Martin (1993) called S*ocial Semiotics.* This analysis methodology developed from systemic linguistics can be applied to non-verbal investigation as Halliday (1985, 1993) and more recently Miller (1998:31-40) have done. Three functions of messages can be distinguished : - *ideational metaf unction*, 'which involves referential meanings either in the world or in our minds and is what most people associate the field with'; - *interpersonal metafunction*, 'which deals with the relationship between writer and reader'; - *textual metafunction*, 'which serves to organize the messages in the other metafunctions into a coherent whole'.

5.3.2.1 The ideational metafunction

In the light of such conceptualization, it appears that text and image are not interchangeable. One of the built-in advantages of visuals as compared to text is that they enable scientists to *show* what is difficult or too long to describe in or around an experiment. In the case of our corpus, the essential referential metafunction of non-verbal elements is to highlight the multiple relationships between variables observed or produced by the specialist.

The very high occurrence of symbolic reference in tables (e.g.: the *n-3* fatty acid family versus the *n-6* fatty acid family in tables) and even more in diagrams (e.g.: ,•, O,•) also allows the information to be densified or compressed, these two discursive processes being used again as allies in the persuasion strategy.

> By far the most prominant use of visuals in the academic texts is to highlight the logical relations of comparison among variables organized in such a way as to imply a cause and effect relationship. [...] The multiple comparisons that can be made are endless. In these visuals, the individuality of each creature has been suppressed to highlight the logical comparative relations of comparison among types.[...] Visuals with a very condensed aspect contain much more information per square inch than the written text. (Miller, 1998:36-37).

This evolution towards more symbolic reference clearly ends up in the 90s by discarding almost all text from diagrams and figures. A definite advantage of symbolic reference is the efficiency acquired by the reader who can peruse the important results at a glance.

5.3.2.2 The interpersonal metafunction

Visuals in scientific discourse enable the results to be checked by peers. Validation of the results depends more on what is *shown* than on what is *said.* The community behaves as doubting Thomas : seeing is believing, which infers that visuals are a sort of prerequisite to building credibility in the eye of the specialist community. Apart from presenting data, they also organize information. This is not too obvious in the articles before 1960, in which the experiment is totally reported, but from the 60s on, diagrams and especially bar charts were built with a view to supporting argumentation and anticipating dispute. It can be admitted however that the organization of information gradually became a convention, a characteristic of the genre. Curves and bar charts present the clear advantage of being able to display the «fourth dimension», that is *time,* even and perhaps above all, if it is the *reconstructed time* of the experiment, which occurs in the article. A figure may translate time into spatial dimension, and it is conventionally deciphered from left to right. It is therefore represented by convention on the abscissa of the curve. Representation of time brings dynamics into figures, which, in other types of visuals, is expressed as an assembly of arrows or a cascade of reactions.

5.3.2.3 The textual metafunction

This metafunction refers to two processes first described in systemic linguistics as *theme* and *rheme,* or more simply through a binary return move to *given* or *new* information. The layout of visuals itself may fulfill a theme function, just like a caption under a table or a figure will provide the orientation or theme of a visual. Indeed, the layout of a visual tends to orientate the reader. Kress & Van Leeuwen (1996 :187) define the status of information according to the positioning inside the visual, which states that 'left' is *given*, that is supposedly known by the reader, and 'right' is *new*, or at least not yet accepted as a paradigm and therefore to be considered with special attention :

> When pictures or layouts make significant use of the horizontal axis, positioning some of their elements left, and other different ones right of the centre, the elements placed on the left are presented as Given, the elements placed on the right as New. For something to be Given means that it is presented as something the viewer already knows, as a familiar and agreed-upon point of departure for the message. For something to be New means that it is presented as something which is not yet known, or perhaps not yet agreed upon by the viewer, hence as something to which the viewer must pay special attention. Broadly speaking, the meaning of the New is therefore 'problematic', 'contestable', the information "at issue"; while the Given is presented as commonsensical, self-evident.

The positioning of given/new information actually supports natural deciphering (at least in Indo-European languages) in a left to right movement. This is all the truer in the field of nutrition, where chronology plays an important part in the interpretation of the results and where therefore curves represent time on the abscissa, making reading from left to right compulsory. Chronological representations of experiments in curves represent almost a third of the corpus visuals. Yet, even in the visuals which do not use chronology in their layout, the left side reports the initial parameters or *controls*, whereas new or major information, or the solution brought to a question are placed on the right. Another interesting observation is the adequation between the table title key-words and the far right column results in tables. However, at the end of the 80s and in the 90s, with the development of symbolic reference, the EFA specialist community started using a layout based on a rising or descending order to report the results obtained on the various fatty acids. Fatty acids are thus ordered in the symbolic notation according to the number of carbon atoms, double bonds and the positioning of the last double bond (from 16:0 to 22: 6*n-3*). In that case, even when data obtained on a single fatty acid constitute the new or major information, this result is not displaced towards the right side of the tables or graphs. Another symbolic notation has developed from this logical account of data. The most significant results are signalled by a superscript letter [a], [b], or [c], which are generally explicited in the caption :

> [a] Significant effect of linolenate deficiency, when data from both groups of linolenate-supplemented animals are combined and compared with data from both groups of linolenate-deficient rats (The Journal of Nutrition, 1980: 1501-1503, tables 2 to 8).

Quick information retrieval from the research paper on EFA will therefore be efficiently performed either by skimming the right-hand side of visuals or by scanning data highlighted by superscript letters.

6. Conclusion

Visual understood as non-verbal elements are shown to be a constant convention in the discourse of specialists on the specific theme of EFA. The corresponding macrotext is characterized by a relative stability of the number of visuals and by privileged recourse to tables, curves and bar charts, in decreasing order of presence. Visuals represent one of the most obvious traces of the upstream work done in the laboratory before the writing of the article. Unlike visuals displayed in scientific vulgarization magazines, they are not meant to attract the reader, and besides their dryness, the tendency to accumulate more and more data and the rising use of symbols make their reading a daunting task, at least to the layman. The diachronic prospect shows that no true type of visual representation has appeared; the widespread use of computers has only made the production of data easier, quicker and has therefore led to the amplified treatment and display of quantitative measurements, while photography, a more expensive and less rewarding tool has almost totally disappeared. Visuals are not totally built on the facts established in the article. They also include some already validated scientific facts. They are supposed to be interpretable independently from the text. Yet the study of the visual/text relationships shows interaction and especially the anchoring function of the linguistic message to visuals. The consultation with specialists of the field investigated here, leads to the opinion that the densified information in the visuals prompts the specialized reader to change strategies. There is a trend towards using the author's mediation as exposed in the text of the Results section, which focuses on the major findings. The most obvious signs of evolution in the visuals is on one hand the number of symbolic and semi-symbolic systems and on the other hand, the shift from visuals that describe to visuals that relate.

These results could lead to more general conclusions as to the use of visuals in scientific communication as a whole, but they need to be confirmed by other studies of the same type in other fields of scientific research.

Appendix:

A complete list of the papers analyzed here can be found in the annexes of the PHD thesis entitled : “Etude diachronique de l'article scientifique de recherche en anglais : le cas de la nutrition”, defended at the University of Burgundy on january 10, 2000 by Anne Magnet, or can be obtained by contacting the author by e-mail to anne.magnet@u-bourgogne.fr. The total corpus has been computerized and is available on request.

Bibliography:

Barnes, B.. D. Bloor & J. Henry (1996) - *Scientific Knowledge : a Sociological Analysis.* London : The Athlone Press.

Barthes, R. (1964a) - «La rhétorique de l'image». *Communications.* **4**. 40-51.

Barthes, R. (1964b) - «Eléments de sémiologie». *Communications.* **4**. 91-135.

Bastide, F. (1985) - «Iconographie des textes scientifiques». In B. Latour & J. de Noblet (eds.). *Les vues de l'esprit. Culture technique.* **14**. 133-151.

Bazerman, C. (1989) - *Shaping Written Knowledge. The Genre and Activity of the Experimental Article in Science.* Madison. Wisconsin : The University of Wisconsin Press.

Bertin, J. (1981) - *Graphics and Graphic Information Processing.* (W. J. Berg & P. Scott. trans.) Berlin : Walter de Grutter.

Dyer, G. (1982) *Advertising as Communication*. London : Methuen.

Halliday, M. A. K & R. Hasan (1985) - *Language. Context and Text : Aspects of Language in a Social-Semiotic Perspective.* Geelong. Vic. : Deakin University Press (reprinted 1989. Oxford : Oxford University Press).

Halliday, M. A. K & J. R. Martin (1993) - *Writing Science. Literary and Discursive Power*. Pittsburgh. Pennsylvania :The University of Pittsburgh Press.

Huckin, T. N. (1987) - «Surprise Value in Scientific Discourse». Unpublished Communication at the C.C. C.C. Convention. Atlanta. Georgia. March 1987.

Johns, A. M. (1998) - «The Visual and The Verbal : a Case Study in Macroeconomics». *English for Specific Purposes.* **17**. 2. 183-197.

Knorr-Cetina, K. (1981) - *The Manufacture of Knowledge. An Essay on the Constructivist and Contextual Nature of Science.* Oxford : Pergamon Press.

Kress, G. & T. Van Leeuwen (1996) - *Reading Images : The Grammar of Visual Design.* London : Routledge.
Lambert, J. K. & G. Schurz (1994) - «Outline of a Theory of Scientific Understanding». *Synthese*. **101**. 65-120.
Larkin, J. H. & H. A. Simon (1987) - «Why a Diagram is (sometimes) Worth Ten Thousand Words». *Cognitive Science. **11**. 65-99.*
Latour, B. & S. Woolgar (1979) - *The Social Construction of Scientific Facts.* B. .Barnes & S. Shapin . (eds.) Beverly Hills. California : SAGE Publications. Inc.
Latour, B. (1990) - «Drawing Things Together». In M. Lynch & S. Woolgar (eds.). *Representation in Scientific Practice*. Cambridge. Massachussetts : The MIT Press. 19-68.
Lemke, J. L. (1998) -«Multiplying Meaning : Visual and Verbal Semiotics in Scientific Texts». In Martin. J. R. & Veel. R. (eds.). *Reading Science : Critical and Functional Perspectives on Discourse of Science*. London : Routledge. 87-113.
Lynch, M. & S. Woolgar (1990) (eds.) - *Representation in Scientific Practice.* Cambridge. Massachussetts : The MIT Press.
Magnet, A. (2001a, *in press*) -«Etude de l'article scientifique de recherche en anglais : Discours sur la méthode». In *L'anglais de spécialité en France : Mélanges en hommage à Michel Perrin.* Bordeaux : *ASp.*
Magnet, A. (2001b, *in press*) «Stratégies de validation du discours scientifique : analyse diachronique du rôle de la section Discussion dans l'article de recherche». *ASp* 26-29.
Martin, J. (1996) - «Les enjeux du discours scientifique : la stratégie de véridiction». *ASp.* **11-14. *13-31.***
Miller, T. (1996) - «Visual Persuasion : The Role of Visuals in Academic Articles and Popularizations». *Les Cahiers de l'APLIUT.* **15**. 3. 22-35.
Miller, T. (1998) - «Visual Persuasion : A Comparison of Visuals in Academic Texts and the Popular Press». *English for Specific Purposes*. **17**. 1. 29-46.
Myers, G. A. (1995) - «Words and Pictures in a Biology Textbook». In «Functional Approaches to Written Text». Miller T. (guest ed.) . *The Journal of TESOL-France*. **2**. 2. 21-38.
Prigogine, I. & I. Stengers (1984) - *Order out of Chaos: Man's New Dialogue with Nature.* London: William Heinemann.
Salager-Meyer, F. (1997) «Diachronic evolution of intertextual referencing in medical discourse» (1980-1990). *Interface : A Journal of Applied linguistics.*
Salager-Meyer, F. (1999) «Referential Behaviour in Scientific Writing : a diachronic study» (1810-1995), *English for Specific Purposes*. **18**. 3. 279-305.

Tufte, E. R. (1983) - *The Visual Display of Quantitative Information.* Cheshire. Connecticut: Graphic Press.

***

Published in LSP & Professional Communication Volume 1, Number 1, April 2001 - ISSN 1601-1929

ABSTRACT

Diachronic Analysis of the Visuals in the Research Paper: A Corpus-Based Study of the Strategies and Semiotics of Visual Representation in Nutrition Biochemistry

Dr. Anne Magnet
Université de Bourgogne
Dijon, France

Key-words:
Visuals / Research paper / Diachrony / Corpus-based /
Socio-constructivism Semiotics / Nutrition Biochemistry.

This article reports a diachronic analysis of the visuals as used in the scientific research paper over a 70-year period (1929-1999). The corpus-based (72-articles) study investigates the strategies of specialists in nutrition biochemistry publishing on a single biological theme : essential fatty acids (EFA). It provides data on the various kinds of visuals as well as an inquiry into the size, positioning, nature and function of visual representation within the scientific paper considered as a genre, over the years. Based on the theoretical principle of semiotics, it then proceeds with an analysis of the concept move which underlies the use of visuals, conceived as the *hard core* of the paper. Readability of a scientific document depends on the number of symbolic and semi-symbolic systems at stake. A scientific visual is constructed as the article itself : It is built as a stratagem, an ambush with no way out. Visuals are used to convince. They are also the only *evidence* of laboratory work. Semiotic analysis shows that the descriptive nature of the visuals used in the corpus changes after the 60s to become the display of multiple relationships. It also analyzes how the techniques of *foregrounding* and *backgrounding* can apply to the study of non-verbal items in the research paper.

Social semiotics gives another light on the use of visuals by showing that a scientific paper does not only contain *new* information. Based on the *theme/rheme* concepts, borrowed from systemic linguistics, the study of visuals in the corpus and especially that of tables, demonstrates that what is already known by the community (*given)* tends to be produced on the left-hand side of the tables, whereas the *new* information is placed on the right. There is also adequation between the title key-words of the tables and the right-hand side results found in these tables. However, due to the development of symbolic representation over the years, the results concerning the research theme in the 80s and 90s tend to follow logical

ordering. Thus, the results about essential fatty acids (EFA) are presented or *staged* in an order according to their chemical composition. Significant results in this case do not appear on the right, but are signalled by special symbols. Specialists can thus very efficiently find the results of interest through this deciphering process, while the layman is more and more at a loss when trying to understand the data represented in visuals.

This paper stresses the interest of a diachronic prospect in the study of the visual characteristics of the scientific article, by showing that if visuals are a constant feature, and even constitute the *hard core* of the article, they should not be considered as stable traits in the research paper contemplated as a genre and their evolution points to changing persuasion strategies pursued by the scientific community represented here.

***