1982-2002: FROM THE PAST TOWARDS THE FUTURE

Maria Gabriella Ottaviani
University of Rome "La Sapienza"
Italy

This paper, after considering the reasons and aims that gave origin to the International Conference on Teaching Statistics (ICOTS), traces the line of thought along which the Conference developed from 1982 to 2002. This is done by applying textual data analysis to the titles of the papers published in the Proceedings of the first five Conferences, and to the titles that were on the International Programme Committee Web site on October 27, 2001. Knowing past and present enables one to present suggestions about possible future Conference scientific developments.

THE ICOTS CONFERENCE: REASONS AND AIMS

As V. Barnett wrote in the Proceedings of the First International Conference on Teaching Statistics (ICOTS), held in Sheffield (UK) in 1982: "This Conference arose from the special interests and actions of the International Statistical Institute (ISI)" (Preface, xi).

In reality the ISI interests in the improvement of statistics education began in 1949, immediately following the Second World War, with the founding of the Committee on Statistical Education within the ISI. Through this Committee, the Institute itself promoted the university training of statisticians at an international level, while in developing countries the ISI concerned itself with the education of official statisticians (Gani, 1979; Vere-Jones, 1995). The ISI began paying more attention to teaching statistics in schools following the mid-seventies. In those years, mainly in developed countries, socio-economic conditions improved, quantitative information given by governments and published in newspapers was ever more frequent, and personal computers became ever more widespread. Moreover, the teaching of mathematics in schools began to change, so that also statistics and probability could find a place within the mathematics programme in pre-university schools, and statisticians became conscious of the necessity to go deeper into the teaching/learning problems that teachers of mathematics had to face when dealing with teaching statistics in schools.

In 1976, the ISI re-emphasised its propositions for statistics teaching (Zarkovich, 1976). Knowing that the theme of teaching could offer the best possibility of influencing the future development of statistics, the ISI underlined the necessity to :

- produce comprehensible university textbooks with applications connected to “real life”;
- publish a Journal to assist teachers at different levels in terms of keeping them informed;
- organise general meetings, for those interested in the teaching of statistics, different from those formal ISI Round Tables organised up to that point for experts in statistical education.

From 1979-1987, the Education Committee, chaired by Professor J. Gani, succeeded in obtaining important results, through the creation of diverse “Taskforces”. Thanks to these, but also thanks to the willingness and involvement of those in charge of each taskforce, significant initiatives were taken. In 1979, at the International Centre for Statistical Education at Sheffield University, “Teaching Statistics” was first published, one of the most important didactic statistical Journals distributed all over the world. Furthermore the International Conferences on Teaching Statistics were initiated in Sheffield in 1982, thanks to L. Råde and V. Barnett. Since then they have continued once every four years, in accordance with the views expressed by the first ICOTS participants, who hoped that "this Conference will be the first of a series of such international conferences on teaching statistics to be held at four-yearly intervals" (Proceedings of ICOTS 1, Preface, xii). According to these wishes, ICOTS 2 was held in Victoria (Canada) in 1986 and ICOTS 3 in Dunedin (New Zealand) in 1990. The success of the ICOTS, the audience of which had passed from about 400 participants from all over the world in 1982 to about 560 in 1990, demonstrated that statistics teachers felt a strong need to unite, talk, and discuss the problems experienced in the course of their daily activities. In the meantime, it gave specialists in epistemology, psychology and statistical education the possibility to expound the results of their research and theories. These forces also emerged with the starting of the International Study Group for Research on Learning Probability and Statistics on the occasion of ICOTS 1 (1982) and

their importance was recognised by the ISI at the Cairo Conference in 1991, where the proposal to establish an International Association for Statistical Education (IASE) as a new Section of the ISI was approved by a unanimous vote of the ISI General Assembly. As a consequence of this, the statistical education activities were transferred to the IASE, among which one of the most important was to continue with the ICOTS program. A Transitional Committee, in charge from 1991 to 1993 and chaired by D. Vere-Jones, took the most important decisions about the venue and organisation of ICOTS 4 in Morocco in 1994. ICOTS 5 in Singapore, in 1998, was the first ICOTS Conference completely designed and organised by the IASE, so that the Forward Address in the Proceedings was signed by the IASE President in charge, M. G. Ottaviani.

HOW THE SCIENTIFIC CONTENTS OF THE ICOTS CONFERENCES HAVE EVOLVED

The Analysis and Its Method: A Proposal

The contents of any thematic conference are supposed to evolve and change over time. The ICOTS conference is now 20 years old and 6 meetings have taken place, thus the time has come to analyse if and how a line of thought exists which has developed from 1982 to 2002. To examine carefully the Proceedings of the previous five ICOTS is quite an impossible task as the number of published papers is 783, and ICOTS 6 is in preparation (Table 1). A possible solution to the problem posed comes from statistics itself, in particular from "*textual statistics*" (Lebart, Salem, Berry, 1998), a field of research helpful to describe, compare and classify sets of texts.

Table 1
Papers According to Author's (or First Author's) Nationality and ICOTS Series

Nationality (Geographic area)	ICOTS						Totals
	1	2	3	4	5	6*	
USA	13	27	37	47	59	76	259
Other N. America Countries	0	2	4	13	4	6	29
South and Central America	3	0	0	13	7	11	34
Europe	21	38	47	94	72	68	340
Asia	10	5	18	18	30	13	94
Africa	1	4	4	18	7	19	53
Oceania	0	4	49	29	60	31	173
Missing data	20	5	0	0	0	0	25
Totals	68	85	159	232	239	224	1007

* According to the International Programme Committee Web site on October 27, 2001

In this particular application of textual data analysis, the fundamental assumption is that the title of each ICOTS paper may be considered as the author/s' reply to the open question: "*What is your contribution to teaching statistics?*". Having made this premise, the text formed by assembling the papers' titles forms the *linguistic corpus* to be analysed, and each word (or *graphical form*) utilised in the papers' titles is a statistical unit. The set of all the different words utilised in the corpus is its *vocabulary.* On these bases a contingency table: titles *x* words can be held, named *lexical table*. If *i* is a title and *j* is a word of the vocabulary, the cell defined by row *i* and column *j* contains the number of times word *j* is used in title *i*. The simple correspondence analysis applied to the lexical table obtained from the corpus identify the words that are characteristics of the papers' titles, either through their presence or their absence. The two dimensional display generated by this correspondence analysis puts in evidence words having similar profiles and thus being closed. These words are usually employed together in a subset of the papers' titles. The software SPAD.T - utilised to implement this analysis - is able to show the subset of titles utilising the same word thanks to a procedure named *concordance*. This process allows one to see each word in a small fragment of the context in which it is used. As the comparisons of lexical profiles only become meaningful from a statistical point of view if the words appear with a certain minimum frequency, hapaxes or rare words are eliminated from the

correspondence analysis by choosing a suitable frequency threshold for words. By this and by a meaningful lemmatisation process, "the vocabulary size taken into account is reduced" and "the patterns obtained are more significant" (Lebart, Salem, Berry, 1998, p. 104).

ICOTS 1(1982): Mainly a Matter of Teachers and Teaching

The linguistic corpus of ICOTS 1 was obtained by assembling the 68 titles of the papers published in the Proceedings. To produce the lexical table to be submitted to the textual analysis procedure, lemmatisation was used and some words were grouped: "developing country/ies" were merged in DC and "elementary school/s" in elem. The articles and most of the prepositions were eliminated. The frequency threshold chosen was equal to 3 and it allowed one to eliminate those words appearing twice. Figure 1 shows the first principal plane (that is, the plane of the first two principal axes) of the correspondence analysis.

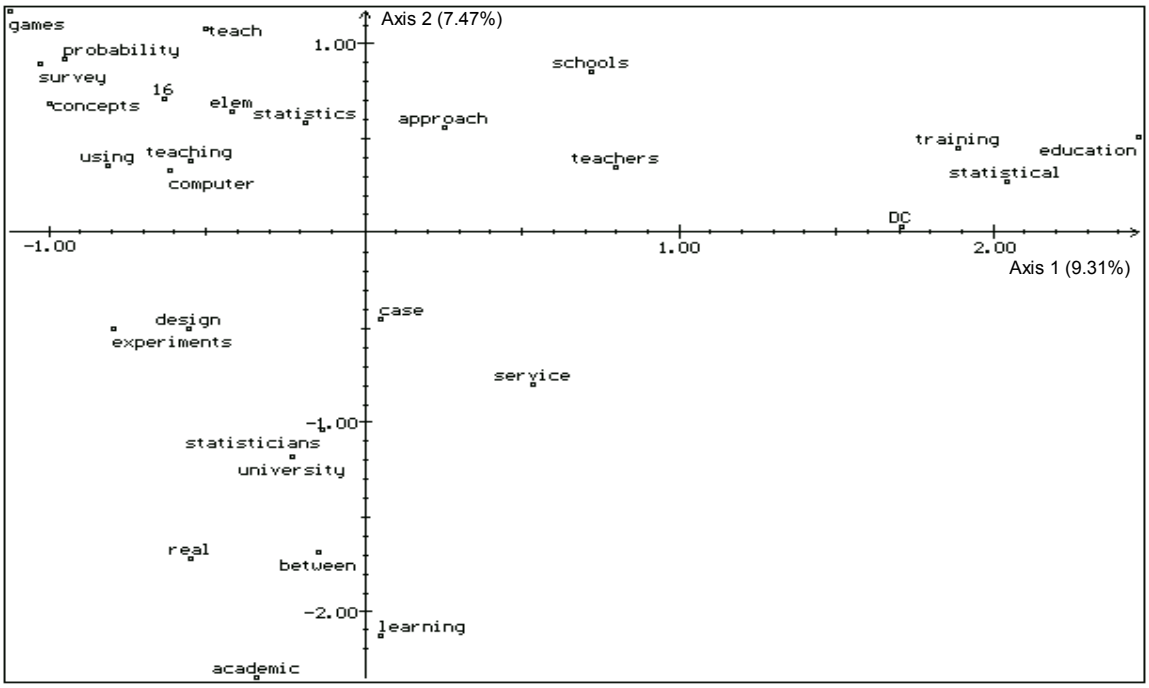

Figure 1. Correspondence analysis of ICOTS 1 corpus (first principal plane) (only a subset of words have been plotted)

At a quick glance the word missing in the plane is "students". This word appears only once in the corpus and is eliminated by the utilised threshold. This makes it evident that neither the students nor the learning problems were the focus of the conference. In fact ICOTS 1 was mainly a matter of teaching and teachers. *Teachers*, also in-*service* ones, were in evidence for their need for *training*. *Statistical education* and *training* were also a necessity in *developing countries*. *Teaching statistics* and *probability* were examined principally with reference to *elementary schools* and to *16*-18 year olds. *Computer, survey, games* were the proposed tools for teaching the two disciplines and particularly their *concepts*. In the u*niversity* the necessity was outlined for a new way of presenting the discipline allowing one to build bridges *between academic* and *real* worlds as well as to emphasise the co-operation between academic and practising, government, medical *statisticians*.

In ICOTS 1 the inclinations of the statistical educators towards applied statisticians and schools were evident, but also the necessity appeared to connect academic and real worlds. All this was in line with the objectives of the Education Committee, that were at the origin of the ICOTS Conferences.

Icots 2 (1986): Students in Evidence as Well as Teaching/Learning Problems

The linguistic corpus of ICOTS 2 was made by assembling the 85 titles of the papers published in the Proceedings. "Data analysis" was grouped as one word labelled DA, "developing country/ies" as DC. Lemmatisation was used. Articles and prepositions were eliminated. The threshold chosen was 3. Figure 2 shows the first principal plane of the correspondence analysis.

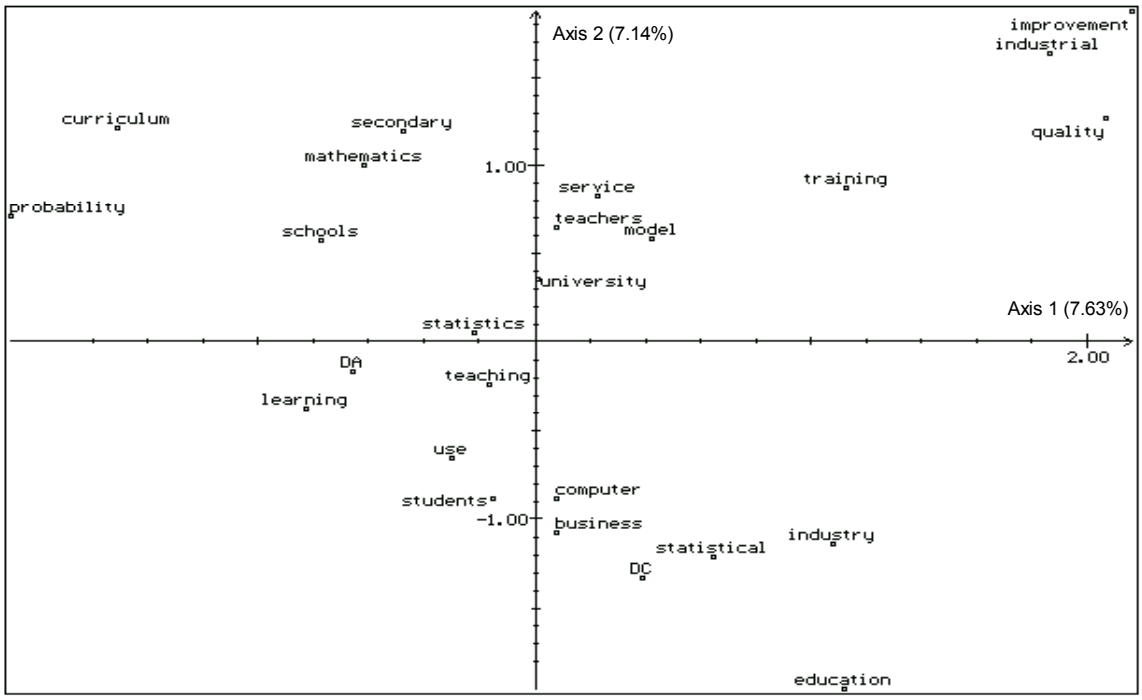

Figure 2. Correspondence analysis of ICOTS2 corpus (first principal plane) (only a subset of words have been plotted)

Statistics and *teaching* are near the origin of the axes as the focus of the Conference. In the third quadrant the word *students* is on the ground near *teaching* and *learning*. The teaching of statistics to students by the *use* of *DA (data analysis)* and *computer* was put in evidence as well as the problems of *learning* statistics and *probability*. This was particularly important for *schools*- mainly the *secondary* ones - where *mathematics curriculum* were reviewed with statistics and probability included. All of this required in-service *training* in statistics for mathematics *teachers* with appropriate instructional *models*. *Statistical education* had to advance in *developing countries*. The teaching of *business statistics* and the importance of statistics in *industry*, or rather, of putting statistical ideas across to industry, was strengthened, and the needs of *industrial training in quality improvement* was justified.

As the analysis shows, in ICOTS 2 the spirit of the Conference was changed as compared with ICOTS 1. This is evident also in the Proceedings Preface, where R. Davidson and J. Swift wrote: "For five days we met, listened and discussed our experiences" (Preface, xiii). It was no more a matter of academy, it was the direct, personal participation and observation as statistics teachers, that were involved.

ICOTS 3 (1990): The Students Right in the Centre

The linguistic corpus of ICOTS 3 was formed by assembling the 159 titles of the papers published in the Proceedings. Lemmatisation was utilised. Articles and prepositions were eliminated. The threshold used was 4, that is, words present three times were eliminated from the analysis. Figure 3 shows the first principal plane of the correspondence analysis.

In ICOTS 3 *students* were the centre of interest, while two main problems seemed to emerge clearly: where the teaching of statistics had to be examined, and in which way it had to develop. The teaching of statistics was considered at *school*, mainly in *secondary schools.* There, *teacher statistical education models* were needed and *quantitative* literacy *projects* and materials

had to be implemented. Also *curriculum* and *project curriculum* in *statistics* began to be discussed. In higher education, the statistics courses referred to were *introductory courses*, *applied courses,* eventually with *probability*. What did emerge was also the requirement of evaluating and assessing the students' *performance* in statistics courses. *Training* did not only concern junior staff in developing countries, but also students through *statistical consulting* and *data analysis.*

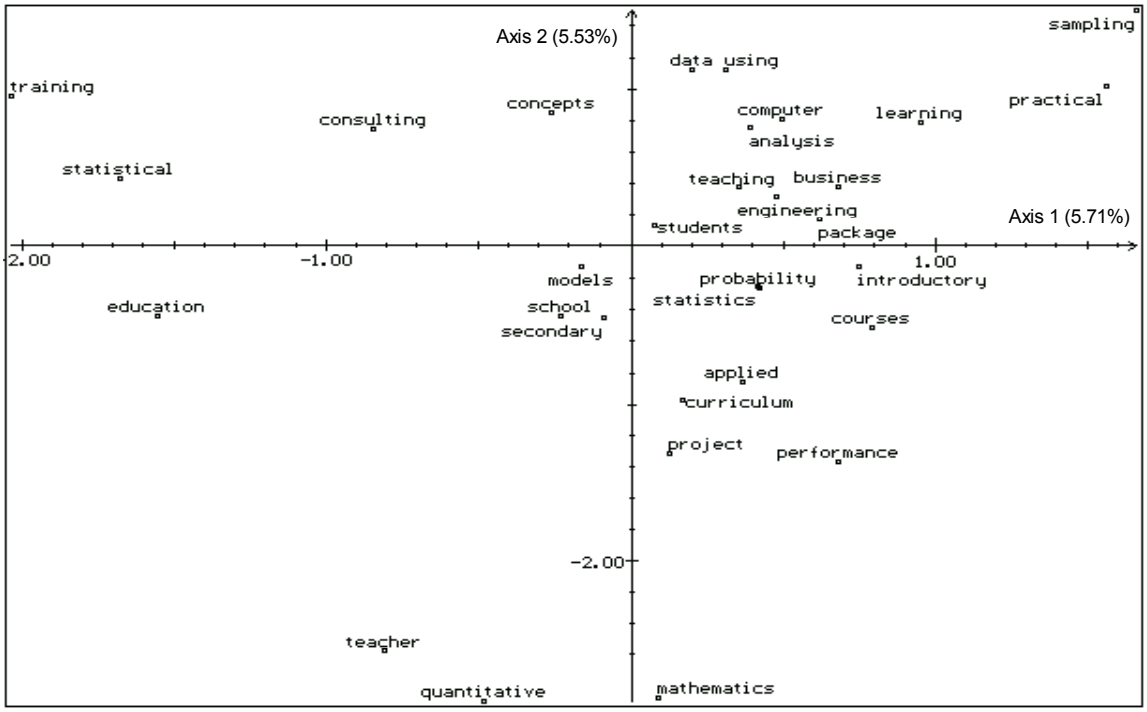

Figure 3. Correspondence analysis of ICOTS 3 corpus (first principal plane) (only a subset of words have been plotted)

Also the teaching of statistical *concepts* and the ways in which students organised them were considered important. The most frequent proposals concerning teaching and learning statistics involved the use of *data*, *computer* and statistical *packages. Sampling* emerged as a field to be better taught and learned *using* computers. *Business* students remained an important area to teach, while the industry sector was more clearly defined by its students: the *engineering* ones.

ICOTS 4 (1994): The Focus on the Teaching of Statistics per se

The linguistic corpus of ICOTS 4 was obtained by assembling the 232 titles of the papers published in the Proceedings. Lemmatisation was used. Articles and most of the prepositions were eliminated. The threshold chosen was 6. Figure 4 shows the first principal plane of the correspondence analysis.

ICOTS 4 seems to have focussed on the teaching of statistics at *school* - particularly *secondary* schools - and at *university*. At university what became evident was not only the teaching of *future statisticians,* but also of students of some *applied* disciplines (*medical, economic, social sciences)*, and obviously *engineering*. Instructional strategies in *statistics course* had to emphasise the use of *data analysis (DA)*. *Teaching statistics experience* looked worth divulging, and the use of *projects* in teaching statistics was seen as a promising path. *Students' learning* of *probability* and statistics had to be improved by *computer.* Students' *understanding* of computers, of *statistical concepts* and of computer-*based* simulation were considered as fundamental aids to comprehension of the discipline. The implications of statistical concepts for teachers' *training* were made evident.

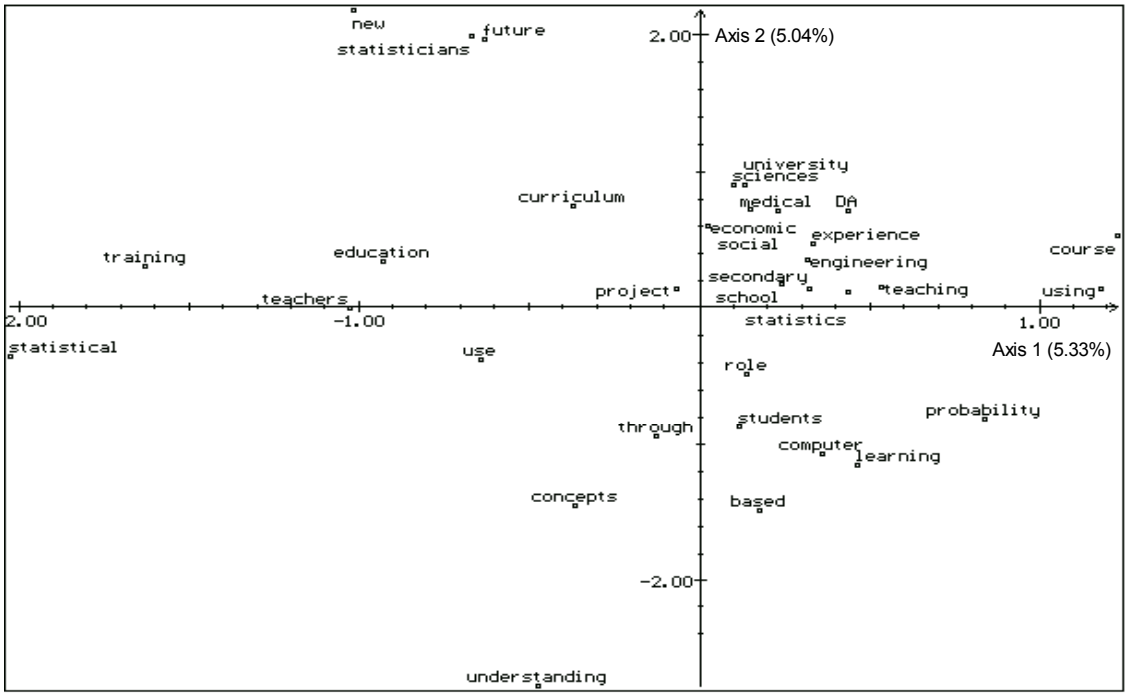

Figure 4. Correspondence analysis of ICOTS 4 corpus (first principal plane) (only a subset of words have been plotted)

ICOTS 5 (1998): About Statistics Students and Statistics Teachers

The linguistic corpus of ICOTS 5 was formed by the 239 titles of the papers published in the Proceedings. The lemmatisation process was utilised. "Data analysis" was grouped in DA and "developing countries" in DC. Articles and prepositions were eliminated. The threshold used was 7. Figure 5 shows the first principal plane of correspondence analysis.

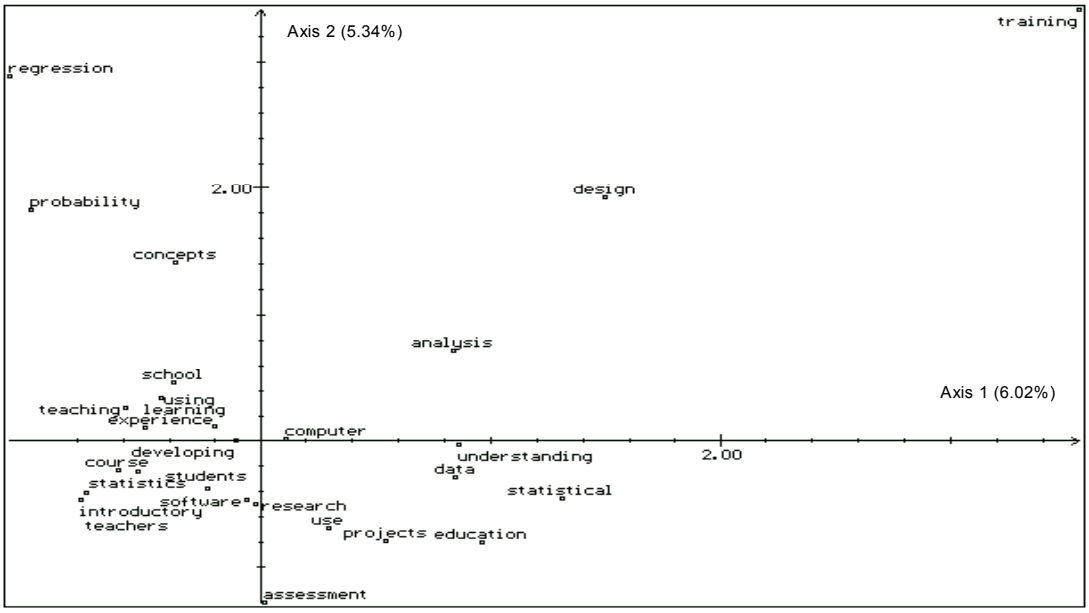

Figure 5. Correspondence analysis of ICOTS 5 corpus (first principal plane) (only a subset of words have been plotted)

Teachers and *students* appeared together in the third quadrant, in reality the idea was emphasised for the necessity to prepare teachers and students in *statistics* and to teach basic statistics to *student teachers*. The focus of the teaching was on *introductory statistics course*, where *computers* had an important role in *developing* and *teaching statistical concepts*. It was

emphasized that the utilisation of computers had a sense if real *data* were used and the impact of computer programs on the *learning* of statistics was assessed. *Assessment*, in general, concerned students' performance in statistics and statistical reasoning. By this also students' *understanding* of statistical data, concept*s, statistical design* had to be enhanced. The learning process concerned both statistics and *probability* concepts. Among the statistics arguments the teaching of *regression* emerged. The *training* had less importance if compared with the previous ICOTS, and concerned mainly official statisticians, with reference to sample surveys design and analysis. The word *research* appeared, but with different meanings, such as: teaching statistics by research, teaching statistics to research groups, research on statistical education. In particular, in the last case the predominant theme was the assessment of statistical education. This could in some way be due to the milestone book: *The Assessment Challenge in Statistics Education*, edited by J. Gal and J.B. Garfield, which had been published in 1997.

ICOTS 6 (2002): The Focus on Research

The corpus of ICOTS 6 was formed by assembling the titles of the 224 papers that appeared in the International Programme Committee Web site on October 27, 2001. Lemmatisation was utilised. Articles and prepositions were eliminated. The threshold used was 6. The first principal plane of correspondence analysis is shown in Figure 6.

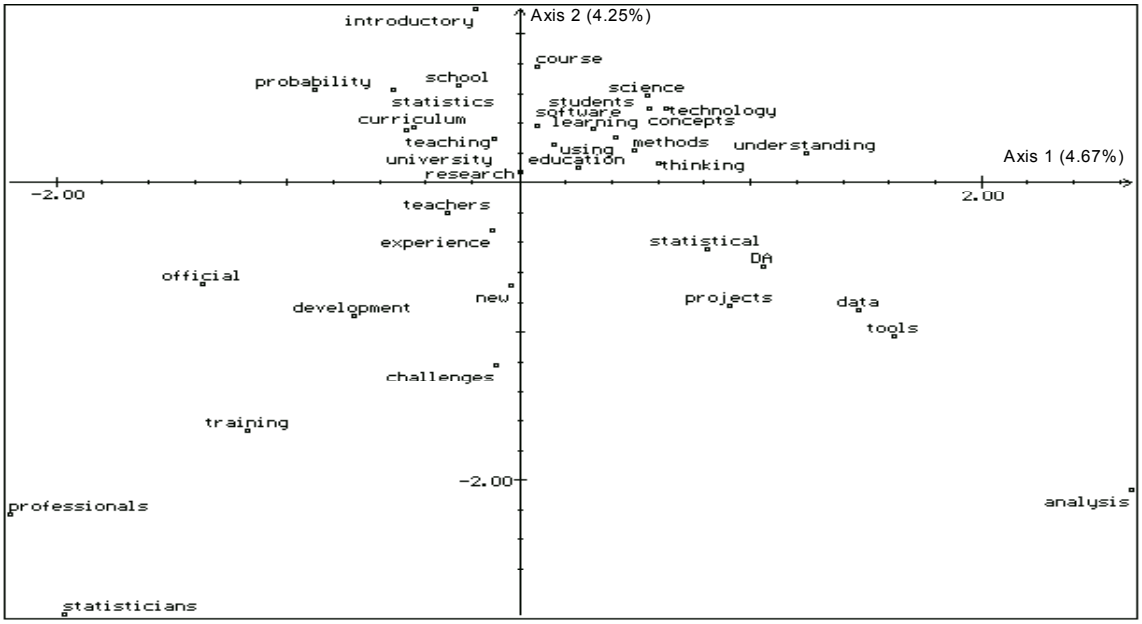

Figure 6. Correspondence analysis of ICOTS 6 corpus (first principal plane) (only a subset of words have been plotted)

The plane may be divided in two parts, the right one is identified by students and their activities: learning and understanding, the left one by teachers, schools, universities and by activities such as teaching and training. The word *research* mainly refers to research in *teaching* and *learning* statistics, but it also refers to a promising way to teach statistics to students of different research fields, that is: through research methods, and through experience in research. Another characteristic of ICOTS 6 seems to be the absence of computers and the presence of *technology*, a more complex and inclusive word. *Data analysis*, *data*, *project* are fundamental tools for learning statistics and for promoting statistical *thinking,* particularly among *school students*. *Training* concerns *professionals*, *statisticians, official statisticians* and *teachers*. It represents *new challenges* for *statistical development*. Also *new curriculum* are important; lessons have to be learned from their development in statistics, and models for teaching statistics across curriculum may be devised.

CONCLUSION

The textual analysis has proved to be a useful Ariadne's thread in the labyrinth of ICOTS Conferences corpora. It has allowed us to show that, from an academic *incipit* due to its strong link with university professors, the Conference moved from dealing with teaching/learning problems and came to enlighten students' problems and their performance. Teaching by real data, suitable computer packages, and research methods were emphasized to grasp the concepts of statistics and probability. In higher education the attention was on introductory courses and on courses of applied statistics for students of the experimental sciences. In more recent years, the proposal of instructional models suitable for statistics and probability, and the necessity to assess students have enhanced the interest of ICOTS attendants towards research. However, educational research in statistics and probability has not only a value *per se*, but is also finalised to help better disseminate and teach the discipline. Consequently the next step to develop might be reflecting on what the statisticians and the statistics education researchers have to offer and to learn from each other as part of a "sole" community. Statistics education researchers, for example, have shown the characteristics by which statistics may be considered a modern discipline, able to develop some skills needed by modern citizenship. Statisticians as the experts of quantitative research methods may offer their ability to deal with quantitative data observed during statistics education research. Education research in fact pertains to the social research field. As such, true randomised experiments cannot be applied when designing research. This implies that when examining the effects of an instructional strategy, attention must be paid to the internal as well to the external validity of the results (Campbell and Russo, 1999). This means that it is necessary to control what are, apart from the treatment, the other possible explanations of the effects obtained within an experiment, as well as to understand the extent to which the obtained effects can be generalised to other populations, settings, and treatment and measuring variables. In other words, statistical education researchers and statisticians have the possibility to work together trying to evaluate the replicability of the experiments done, their reliability, validity, generalisability (Kelly and Lesh, 2000). It is important for the statisticians, as statistics teachers, to understand what the working relationship between statistics education research and practice is. If statistics teachers have to learn from statistics education research, they need to know what the limits are - in time and place - to the practical generalisation of the results obtained by statistics education researchers, in order to put proposed instructional models and strategies in practice.

REFERENCES

Campbell, D.T., & Russo, M.J. (1999). *Social experimentation*. Thousands Oaks, USA: Sage Publications.

Gani, J (1979). The role of ISI in the field of statistical education: 1949-1979. *International Statistical Review, 47*, 3.

Kelly, A.E., & Lesh, R.A. (Eds.) (2000). *Handbook of research design in mathematics and science education*. Mahwah, USA: Lawrence Erlbaum Associates.

Lebart, L., Salem, A., & Berry, L. (1998). *Exploring textual data*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Grey, D.R., et al. (Eds). (1983). *Proceedings of the 1st International Conference on Teaching Statistics,* Volume I and II. Sheffield, UK.

Davison, R., & Swift, J. (Eds) (1986). *Proceedings of the 2nd International Conference on Teaching Statistic*. University of Victoria, Canada.

Vere-Jones, D. (Ed.) (1991). *Proceedings of the 3rd International Conference on Teaching Statistics.* Volume I and II. Dunedin, New Zealand.

Proceedings of the 4th International Conference on Teaching Statistics (Volume I-II) (1994). Morocco.

Pereira-Mendoza L., et al. (Eds) (1998). *Proceedings of the 5th International Conference on Teaching Statistic*. Volume 1-3. Singapore.

Vere-Jones, D. (1995). The coming of age of statistical education. *International Statistical Review*, *63*, 1.

Zarkovich, S.S. (1976). A reappraisal of the ISI statistical education programme. *International Statistical Review, 44*, 2.