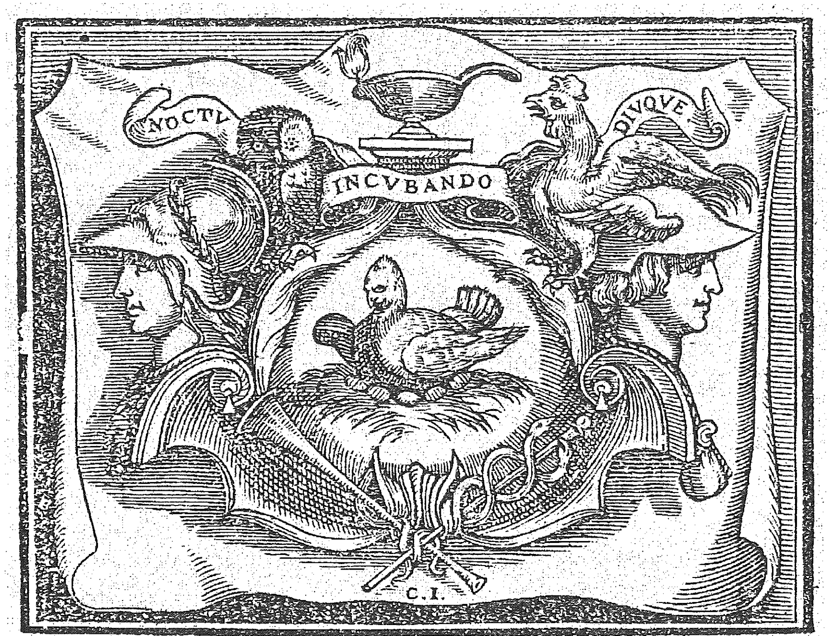


Emblème de l'imprimeur Meursius vers 1650, publicité de l'imprimeur Plantin, dû à P.P. Rubens, tiré de page de titre du Tractatus physico-Mathematicus de aestu maris de Th. Moretus publié en 1665.

Language engineering - The outcome of the intersection of Linguistics, Mathematics, Computer Sciences

MOLEA Amelia
Bucharest (Romania)

Abstract

The computer-based information and communication technologies are providing direct access to each and every ongoing field all over the world through a combination of written text, graphic images and sound. Language will, in some form, be at the core of the screen-based multimedia communications that are likely to become an intrinsic part of life both at the work place and at home.

As information technology keeps growing in different areas, equal opportunities for everyone to access information become a key issue. Clearly languages play an important role in this respect.

For Europe a dual challenge exists: to maintain its linguistic and cultural diversity and to ensure equal opportunities for business and citizens alike to participate fully in and share the benefits of the new information era.

Language engineering is the core of information technology and this will be the key industry of the 21st century. The information super-highways conceived today will soon carry infinite amounts of digital data, images, sounds, tables, figures, calculations and process protocols. If these data are to be intelligible, and make sense, they must be bound together by language. Without natural language processing, information remains incomprehensible.

Language, once a cultural asset only, has now become an economic commodity, too. The national language institutes have acquired new responsibilities. Now their task is to provide information from abroad available in the national language and for locally produced information to be distributed world-wide in major international languages. To train and employ more translators is not sufficient. We must also take care that the necessary language technology is being developed.

1 Current Issues in Language Engineering

Spell checkers were among the earliest successful language technology applications. They have been accepted as useful devices and are still being sold today in ever-improved versions. The majority of early machine translation systems, particularly the more sophisticated ones have not survived. SYSTRAN is still kept alive by the European Commission's translation services, but many others have disappeared without leaving a trace.

Spell-checkers don't need semantics. Even in the seventies they were based on little more than a list of the most frequent word forms, i.e. linguistic knowledge widely available or easy enough to generate from corpora.

Machine translation however needs semantics. Our understanding of word meanings or lexical semantics in the seventies was contained in dictionaries, and it was arranged in such a way that a user with some experience could understand using a great deal of implicit knowledge about the world and inductive reasoning. All assets computers do not possess, such as having the ability to draw analogies. Therefore, there is no surprise that early applications involving semantics were not very successful.

Language technology can work with two kinds of semantic information. One is rule based and presupposes a formal logical semantic analysis of the phenomenon under discussion. The other kind uses statistics and is not really semantic at all: it computes the five or ten words or word forms preceding or following the word in question and relates this information to the different translation equivalents found in parallel corpora. The German word *Schnecke* for example is translated into English either by *snail* or *slug*, where *snail* refers to the creature with a 'house' and *slug* to the one without. The rule based approach states just that and searches the German text for clues from which we can infer the correct translation equivalent. The statistic-based approach does not look at the meaning at all. It looks for words and other traces that frequently occur when *Schnecke* is translated as *snail* and for other patterns co-occurring with *Schnecke* being translated as *slug*. In the context of *slug*, we would probably find words like *vegetable* (garden), *lane*, *wet*, and various forms of *get rid of*; while in the case of *snail*, I would expect words like *table*, *course*, *wine* but also *vineyard* and *sunny*.

Today's successful applications involving semantics work with an amalgam of the rule and the statistic-based approach. The statistic approach has some attractive advantages: the data required can be generated from corpora with no or only little human intervention. As it is just an emulation of semantics, one does not have to be able to state explicitly what a lexical item means. Indeed it leaves out the entire question of meaning. Its inherent shortcoming is the rate of accuracy. Even a rate of 95 translation equivalents implies that every twentieth word in a text is mistranslated, practically, every sentence and certainly more than what most people would like to live with. On the other hand, the rule based approach can be a very expensive alternative. It presupposes something like a bilingual dictionary that would enable translation of a text correctly into an unknown foreign language. The reason why such dictionaries do not exist for human users or machines is that the explicit linguistic knowledge they would have to contain is not yet available and that this knowledge is extremely expensive to produce.

The new generation of language technology applications (monolingual and bilingual or multilingual ones) deals with semantic problems. They recognise the fact that computers cannot understand spoken or written texts in the way humans can. Therefore, these text processing systems can only emulate the human faculty of 'understanding' by a mixture of rules and probabilities. To find the right mix is less a question of theory and principles than of calibration and

learning by doing. The crucial point is the performance of an application under real life conditions. The application has to prove its cost efficiency, i.e. it must demonstrate that it can complete a task cheaper than a trained human. After two decades of experimental and pilot systems, the emphasis today is on robust applications for which there is a real market, i.e. one where users are willing to pay a fair price.

The linguistic formulated knowledge in existing grammars and dictionaries - is unsuitable for language technology tools for two reasons. First most of it is not corpus based; rather it reflects the individual linguist's (lexicographer's) competence based on a collection of data (citations); and however large their collection may be, it is permeated by a bias that cannot be avoided. Secondly, traditional grammars and dictionaries have been devised for human users, who differ substiantally from machines. Human beings use inductive reasoning and can draw analogies easily; faculties like these are taken for granted and are reflected in the traditional arrangement and presentation of processed linguistic data. Language technology tools cannot take recourse to common sense; as in this, all knowledge has to be spelled out in the form of rules, lists, and probabilities. But anyone who has gone to the sources has experienced the problem that when we start analysing language as it occurs in a corpus, we gain evidence that renders existing grammars and dictionaries as very unreliable repositories of linguistic knowledge. We discover that traditional linguistic knowledge gives us a very biased view of language, a view that has its roots in the contingency of over two thousand years of linguistic theorizing. We are so accustomed to this view that we take it as truth, as reality and not just for an interpretation of raw data. It is true that traditional grammars and dictionaries have helped us, fairly satisfactorily to overcome the linguistic problems we human beings have to deal with. But they will not be good enough for language technology applications.

That's why, however cumbersome and expensive it may be, language has to be described in a way that it is appropriate for language engineering. It has been demonstrated that monolingual and bilingual dictionaries are of no (or only little) use when it comes to automatically translating a word from one language into another in cases where there is more than one alternative. To reduce the cost of a corpus based language analysis from scratch, which is indispensable, corpus exploitation tools have to be developed. They will arrange the rough facts (including statistic driven devices for contextual analysis) and process them (with a great deal of human intervention for the semantic interpretation of data) into algorithmic linguistic knowledge and rules derived from objective data rather than individual competence. Perhaps this will result in the finding that traditional categories like noun, verbs and adjective do not, after all reflect categories useful for language processing.

2 Technological Processing of Romanian Language Data

Technologization of Romanian language is only at the beginning in spite of the fact that research has been carried out in this field for a long time. Most of the research was focussed on written language as well as speech. The first attempts to process written language started in the sixties. At that time a group of young mathematicians and engineers under the guidance of Erica Nistor Domokos started at the University of Timisoara a project on machine translation. Supported by the famous mathematician Grigore Moisil the group implemented in 1963 a prototype system of translation from French and English into Romanian. In spite of the outstanding results for the technology of that time presented in public demonstrations the lack of interest on behalf of authorities led to the abandoning of the official research in the field of machine translation.

After 1990, there was a tremendous increase of the possibilities of information and free circulations. The easier access to hardware and software advanced technologies has a positive impact on the technological field of language in general and technologization of Romanian in special. The international programmes of academic exchanges, the research stages, participations to important conferences allowed for adequate information of Romanian researchers on the advances and use of modern technologies for international and national priorities.

Linguistic engineering research focused on formalisms based on unification, on reversible methods of natural language processing. Modern linguistic theories, with computational relevance are used as instruments by more and more research groups consisting of computer sciences specialists and linguists.

We are glad to acknowledge introduction of computational linguistics and modern linguistic theory courses in the curricula of several higher education institutions in Romania (The Computer Sciences Faculty of Al. I. Cuza University of Iassy, Computer Sciences Faculty of Babes Bolyai University of Cluj-Napoca, Computer Sciences Faculty of the Department of Mathematics of the University of Bucharest, Foreign Language Faculty - Department of English of the University of Bucharest, Computer Faculty and Electronics Faculty of the Polytechnical University of Bucharest, Computer Faculty of the Technical University of Timisoara)

We have several national priorities set up in this filed regarding:

- development of computational linguistics.
- computer aided acquisition of concepts of national and European culture.
- language technologization (the setting up of computational linguistic resources.
- computer aided translation, foreign languages teaching, artistic languages presentation.
- development of multilingual computer aided services.
- development of authorial systems (systems capable to intelligently assist the utilizers to create and manage complex documents).

In an attempt to assess the most developed areas of language technology in Romania which could constitute a standing point for the later development of Romanian language technologization, the following areas have to be pin pointed:

- morpho-lexical and syntactic processing instruments (analysers and generators).
- vocal signal processing instruments (recognition, synthesis, prosody).
- spell-checkers.
- question/answer systems (friendly interfaces) on applications.
- almost complete descriptions of the Romanian morphology, both paradigmatic and derivational.
- partial syntactic descriptions of Romanian.
- vocabularies.
- dictionaries for computer-aided processing of written language.

- terminology thesauruses for various linguistic registers.
- specialises corpora.

3 Language Engineering Applications

In support of the ideas presented above I am going to provide some detailed examples of how language technology actually works in applications. I will start by introducing several facts concerning machine translation as this application is presented by Rajmund PIOTROWSKY.

The model designed to develop and implement a computer simulation of a psycho-linguistically realistic model of language behaviour is named 'linguistic automation' and consists of a hardware, an operating system computer program for natural language processing, a vastlinguistic database.

According to PIOTROWSKY, 'linguistic automation' may be conceived as a level based system. The first level is the 'linguistic information database' including linguistic and encyclopedic data with their probabilistic weights marked. This level comprises an automatic dictionary (structured as a lexicon providing information about word-forms, stems, phrases), lists of grammatical affixes, toponims, anthroponyms, abbreviations, etc. It also comprises organisation and updating programs. The second level comprises a set of functional modules (M) consisting of two subsets M_a (incorporating the analysing modules) and M_s (unifying generating modules).

$$M_a = (d, c, l_k, l, m, L_k, L, s_1, s_2, s_3)$$

where

d = module of graphemic text decoding
c = speller
l_k = module of lexical analysis of the key lexical items
l = module of lexical analysis of the lexical items in the text
m = module of autonomous morphological analysis of text words
L_k = module of lexical-morphological analysis of key lexical items
L = module of lexical-morphological analysis of text words
s = module of surface structure analysis
s_2 = module of deep topic-comment analysis of the sentence
s_3 = modules effecting semantico-pragmatic analysis of the text

$$M_s = (k, c, l', L', s_1', s_2', s_3')$$

k = module of encoding
c = speller
l' = module of lexical items generation in the text
L' = module of lexical-morphological generation of lexical items
s_1' = module of surface structure generation of the output sentence
s_2' = module of syntactic-semantic generation of the output sentence
s_3' = module of semantic-pragmatic generation of the target text

The third level comprises some concrete systems and subsystems of natural language processing such as: the systems and subsystems assigning 1) the analysed text to a certain language 2) word-by-word and phrase-by-phrase machine translation 3) rough lexical and morphological machine translation 4) semantic-syntactic machine translation 5) topic-comment machine translation of headlines and book titles 6) text fragmentation, compression and abstraction.

The forth level is carried out through incessant human being - computer interaction.

Taking into account that any natural language processing system has to deal with indeterminacy conditions under the form of a set of alternatives present in the database and in the algorithm blocks, language automation is endowed with an artificial brain module which has to select the appropriate decision. Similar to other control and management systems the decision making body of the linguistic automation can be described as a hierarchic structure with three levels: 1) self organisation 2) adaptation of linguistic automation to the given texts 3) selecting a suitable decision for a concrete task. The third level is very important for the development of linguistic automation and methods to work out the faults due to engineering and linguistic limitations appearing in the linguistic automation have already been put into algorithms. The methods may be interesting for specialists.

The existing polyfunctional natural language processing systems are not perfect linguistic automata. Man still plays an important part in maintaining feedback in man-machine dialogue and the interaction of linguistic automation modules. The role of man is greater in the higher linguistic automation levels, providing functional decision and synergistic organisation, than in lower, more primitive blocks. In the future, the main efforts of researchers are expected to concentrate on extending linguistic automation decision possibilities.

In order to implement various applications of natural language a lexicon is needed containing phonetic, morphologic, syntactic etc, information. The example discusses hereafter concentrates on morphological aspects of vocabulary as seen by S. COJOCARU who solves this problem appealing to two methods: static and dynamic. The standing point of the first method is the classification of Romanian language words into inflection groups. The author makes use of a 30,000 words dictionary, each word being attributed to the respective group number. To obtain a formal description of the inflection process in accordance to the above mentioned classification there have been formulated grammars with dispersed context containing rules of the form:

$$[/] * [\#][N_1]a_1 \urcorner b_1 a_2 ... a_{n-1} \urcorner b_{n-1} a_n \rightarrow a'_1 \urcorner b_1 a'_2 ... a'_{n-1} \urcorner b_{n-1} a'_n N_2$$

where a_i, a'_i are arbitrary words (taking into account the algebraic sense of the notion word), b_i is a non-empty word or a reserved symbol $*$, N_j indicates an inflection subparadigm. The rules are to be interpreted as follows: be w a lema word. Each sign "/" indicates a letter, which must be cut from the end of the word w. The word v, obtained from w after the letters at the end have been cut, is a stem (if N_1 exists), and N_1 is an index denoting the inflectional subparadigm containing the affixes which will be attached to the stem v to obtain the respective derivates. The substitution rule can be applied if $v = f_o a_1 f_1 a_2 f_2 ... a_{n-1} f_{n-1} a_n f_n$, where f_i is an arbitrary word (including an empty one) which does not contain the words "forbidden" b_i. If there is more than one representation of this kind of the word, v is selected the first (scanning from left to right, if sign $\#$ is not present, and in the opposite direction in the opposite case. $b_i = *$ does not impose restrictions on f_i. In the indicated context parallel substitution takes place: $a_1, a_2, ..., a_n \rightarrow a_1, a_2, ..., a_n$, a new stem v being obtained, to which the affixes in the subparadigm N_2 are attached.
Thus in order to generate the inflection of a word knowing its group number it is enough to

interpret the respective rules. One or more grammatical rules can correspond to a group number. To describe the inflection system of Romanian verbs, nouns, adjectives, articles, numerals and pronouns 866 rules have been found necessary making references to 320 subparadigms. Using in some cases scanning of the lema-word from right to left we can use simple rules, context free or having only substitution contexts (not the "forbidden" ones).

For the derivation of the words that are not in the inflection groups the dynamic method is used. The inflection programs initiate a dialogue to determine the part of speech, gender (for nouns) and other additional information for more difficult cases. To carry out morphological derivation it is necessary to find out: a) the vocalic and consonantic alternances, b) the application context of alternance rules, c) affix series. The tables containing affix series, alternances and their admissible co-occurrence are the basis for the functioning of inflection programs. As the irregular words are seldom, they are described a priori and processed in a special way. Obtaining the inflections of a word we can determine the number of the inflection group thus reducing the presentation of all the words in the lexicon to the static method.

The procedures described above have been used to set up a lexicon of approximately 65,000 words which served as basis for the realisation of a spell-checker for Romanian (ROMSP). ROMSP is implemented in Borland Pascal with Objects 7.01 for MS DOS having the following components: a) the spell-checker program with a friendly dialogue, very simple even for the unexperienced utilizer; b) the morphologic derivation program of the Romanian words, necessary for updating the database; c) database management program which carries out compactization of vocabulary, its integrity upgradation and control provides similar words to the given one.

In the implementation of the vocabulary two goals have been followed: a) diminution of the database volume; b) efficient access to the items in the base.

The vocabulary is divided into pages, the table of access to pages being a hash-function stored in RAM thus diminishing considerably the number of addresses to the hard disk. Each word is presented by three components: 1) the first two letters, stored separately (elements of a page are words which begin with the same two letters); 2) the rest of the stem (theme); 3) reference to the valid affix set for this stem.

The vocabulary occupies 1.2 Mb, the checking speed being of approximately 100 words per second on a very simple computer (IBM PC 286, 12 MHz).

Last but not least I would like to bring about some issues of current interest in the development of speech technology which can be considered as one of the most important fields using interdisciplinary skills.

Speech processing is a field encompassing a great variety of technologies and applications. Many of these applications such as automatic recognition and synthesis have become traditional as outcomes of several decades of intense research. Some others although are less known or more recent, but they are important and useful. In spite of the advances in this field the outcomes are still far from what they should be. The tasks initially set up proved to be extremely difficult in time. The causes are complexity of the vocal signal as well as the difficulties encountered in its processing linked either to the recognition of its informational content (the vocal signal strongly depending on the speaker and the context of a message) its production, the transmission of this signal at distance.

In the field of speech technology, automatic speech synthesis and especially text-to-speech synthesis has a special place since it can play an important part in the man-machine interface. A text-to-speech synthesis system can provide an important range of applications in many fields from electronic mail access to various databases through standard conventional communication networks to reading systems for the blind. I would like to present just a part of a larger research carried out by an interdisciplinary research team in the Military Technical Academy in Bucharest. This team worked on a text-to-speech synthesis based on the following general text-to-speech synthesis system which has two main parts: one dealing with linguistic processing which turns input text into a phonetic and prosodic representation. The other is one of acoustic processing which generates the speech signal using specific techniques for the type of synthesis and the acoustic units chosen.

A general schemata of a text-to-speech synthesis system comprises several levels aiming at providing speech similar to natural speech. The input of the system is the text and the output is the speech.

The input message can come from various sources, utilizing various writing protocols. Linguistic processing must convert the input letters into a form that can be appropriately processed. This stage of the synthesis must detect free spaces, sentence boundaries or sometimes the end of a text sample. Next the capital letters must be detected (both the ones at the beginning of a sentence and the ones that may come up later in the text) and the most usual abbreviations must be conveyed into normal writing. The protocol at this stage must also process numbers, integers, decimal numbers, hours and dates. At the same time punctuation must be interpreted which is to be used to establish prosodic features. A morphological analysis is carried out here to each word being assigned a grammatical category; this is a result of the fact that one word may be pronounced differently in accordance to the part of speech it represents. Accents may also be analysed at this stage. The output of this module will be a sequence of letters making up words with a grammatical description accompanied by accents and other prosodic items.

The next stage converts the text into phonemic symbols which describe the way the letter sequences are to be pronounced. The schemata uses phonemic units but obviously there can be used other units making up words - syllables for example.

A note should be made of the fact that most synthesis systems, irrespective of the basic unit chosen always makes use of a hybrid system to obtain such units: a number of rules are used which take into account pronunciation and also a dictionary comprising "exceptions", i.e. words for which rules cannot be established in the regular way.

The phonetic transcript will produce as output a row of characters representing all the allophones (the pronunciation variants of the phonemes) -in case we have used phonemes as basic units- accompanied by quantitative values (duration of the acoustic units, basic frequency and its variation) which are calculated by the prosody generation module providing the intonation and accent of the words spelled out.

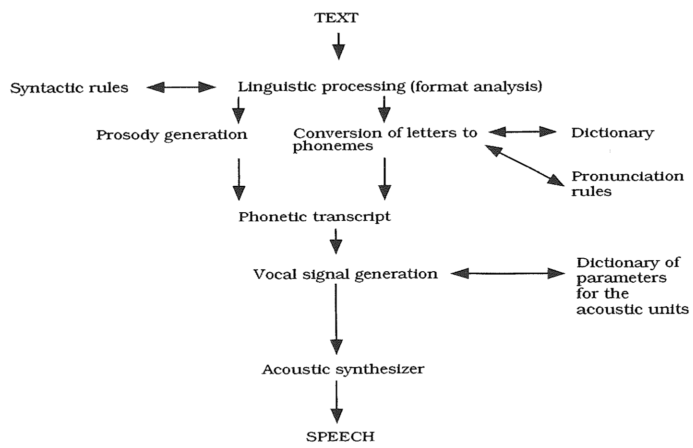
The last two modules accomplish synthesis of speech through the following operations: - information output from the previous module is decoded: - the dictionary provides the parameters corresponding to the acoustic units; - signal samples are generated using for instance LPC synthesis; - the segments thus obtained are concatenated: - the final wave form is synthesized by the acoustic synthesizer (D/A conversion, filtering, amplification).

This general system has been implemented on IBM PC using a linear prediction algorithm and here are its most important performances: compression rate 2.4kbs, real time synthesis (approximately 10% of the speech time of a statement for a 486 PC at 66MHz), good intelligibility and naturalism. The characteristics of the voice on which analysis has been carried out, previous to

undergoing synthesis, can be easily recognised.

We established syllables as the basic elements for synthesis, these segments in our opinion, offer a good compromise between the number of basic units and the number and complexity of concatenation rules for them. We still have to study prosody more thoroughly since problems are raised by the fact that, at least in Romanian, the accent is free and there are far too many exceptions to the rule. This is a problem that needs further research.

General schemata of a text-to-speech synthesis system



The conclusion of the above presented applications is that sciences such as linguistics, mathematics, computer sciences work together in achieving technological support for the future development of the global society.

The applications and facts about language engineering are pointing towards the "holistic-integrative" characteristic of modern science.

4 Some Epistemological Considerations

I would like to conclude this presentation with some considerations concerning the "epistemological" status of contemporary science, starting from some ideas offered by the work of Ilie PARVU.

The recent sciencific development fundamentally modified one of the parameters determining the setting up of epistemological methods. Generally speaking the classical epistemological systems proceeded by taking into account a single scientific subject whose presuppositions they tried to make explicit through universalization, and based on such grounds, they defined an abstract concept of "science". The simultaneous maturation of a great number of scientific subjects nowadays generated new centres of "methodological diffusion" and philosophical problematization within science. If the "philosophical part" played by a scientific subject essentially depends on its insertion in the theoretical configuration of a certain period, then nowadays we are witnessing a multiplication of the philosophical centres of problem defining of science.

In the past, mathematics and physics (and sometimes biology), represented the "paradigms" of knowledge, providing methods and reasoning models for all subjects, at the same time providing "the extraction field" of norms and principles defining the very essence of "scientificity". In our century various other branches of science have entered the "theoretical stage", the spiritual configuration of contemporary science looking more and more similar to a "complex constellation" of subjects with methods, techniques, tools and non-homogeneous conceptual systems.

Consequently, the generation of a general epistemological interpretation from the study of present day science must be preceded by the setting up of several "regional" epistemologies by defining the status of theoretical knowledge and the methodological specificity of various groups of subjects.

I am going to review several of the most significant theoretical achievements and methodological changes which occurred in several branches of science.

Starting from a taxonomy of science by C.F. von Weizsacker I would start this review with the structural sciences, a field of knowledge creating the abstract tools needed in all branches of science. Within the structural sciences there are included not only pure and applied mathematics, but also systemic analysis, theory of information, cybernetics, theory of games considered to represent the mathematics of temporal processes or "structural theories of temporal changes" whose "assisting instrument is the computer, whose theory is itself a structural science".

The modifications at the level of structural sciences have been influenced by the development of logic. It has essentially influenced "the spiritual configuration" of contemporary science contributing alongside with mathematics both to the setting up and expansion of a way of reasoning (structural-axiomatic) and, since it provides a methodological tool necessary to the philosophical analysis of science, to the foundation of a conception on science comparable as rigour to the highest theoretical achievements of science itself. In addition to that, through its own internal elements logic has become one of the most active centres of current knowledge, generating new philosophical problems, participating in the new reconceptualisations of epistemology.

Summarising the contributions made by contemporary mathematics to the setting up of a modern epistemological conscience the following facts should be mentioned: 1) the setting up and /or operationalisation of several "ways of reasoning" with great applicability, necessary to achieve theoretical domination of complex processes and systems (structural-axiomatic, statistic, strategic, interdisciplinary reasoning); 2) through its self-reflexivity mathematics offers tools for the self-knowledge of science, for the setting up of a theory of knowledge; 3) new developments in mathematics allowed a substantial progress in the understanding of the structure and objectives of scientific theoretization; 4) by its internal evolution mathematics led to outstanding outcomes. They require re-thinking of relationships between formalism and intuitive constructions, empirical knowledge and a priori knowledge, or analytic and synthetic, making important suggestions concerning a general epistemological view on knowledge.

In the field of physics, the central subject of natural sciences, recent research has been developed, against the conceptual background set up by the theories. This influenced the "reasoning style" of contemporary physics, the theory of relativity and quantic mechanics. They have brought about a deep revolution of physic knowledge producing new sub- or supra-mathematical structures (the theory of restricted and generalised relativity), new logical onthological formalisms (quantic mechanics), or even a new "methodologic order" in theoretic construction (cosmology), proposing a new form for the scientific law, new types of theories meeting new completion standards and new criteria of "physic reality".

Out of the recent developments in psychology two research directions are of special interest for epistemology. They offer direct data for the interpretation of several basic aspects of the process of knowledge. They are the psychology of cognition and the studies on perception and representation. The theoretical models using the tools provided by the theory of information and computer simulation ("artificial intelligence"), the setting up of "universal grammars" which should define "linguistic universals", "the cognitive faculty of language" which is specific to humans, the modelling of the learning processes, they all have provided new elements for a reconceptualisation of psychology and for a deeper understanding of knowledge structure and mechanisms and for the construction of an integrated systemic model of the human being.

Important progress has been achieved lately by linguistics, a field undergoing modification of methods and concepts and setting up new cognitive objectives. As different from classical linguistics, a taxonomic empirical science, whose task was to set up the language corpus. The new trends initiated by CHOMSKY aim at going into the internal structure and the functioning mechanism of natural languages, formulating fundamental explicative models. R. MONTAGUE's theory in Universal Grammar carries out for the first time an integral semantic-syntactic analysis of natural languages with the methods of contemporary logic; this theory has had influences on several contemporary themes of epistemology such as: the logic - language relation, a priori knowledge, logical truth and analycity, etc.

However, I didn't attempt a global analysis of current science, that's why I think I can conclude at this point that the epistemological characteristics of global science is given by its "constellation of ways of reasoning" (structural-axiomatic, synthetic- integrative, evolutionist, historic, statistic, organisational, architectural, etc) which correspond to the diversification and maximum methodological-instrumental and thematic-conceptual expansion of knowledge. Contemporary science seems to have detached itself from the "methodological monism" and conceptual reductionism of other epochs admitting plurality of "methodological centres" of knowledge and the diversity of types of laws and theoretical explanations. The only unification seems to be related to the tendency of cognitive submission of "complex totalities" (highly complex phenomena and systems) present in the majority of fields of current research. The "holistic" tendencies of current research attach a great importance to the synthetic-integrative way of reasoning which became manifest in current science.

This integrative holistic tendency brought about the expansion of mathematics over all branches of research as a "reaction to a too sharp differentiation of subjects". This refers not only to the use of mathematical language or modelling techniques but also to the expansion to all fields of knowledge of the "mathematical way of reasoning" in its present aspects of functional, analogic, axiomatic, recursive, strategic, organisational, architectural, etc. ways of reasoning.

I would like to end this paper by saying that the world today is facing various problems making up a dichotomy whose terms can be formulated, as CHOMSKY puts it, as "Plato's problem" (pure knowledge) and "Orwell's problem" (social existence). The first tries to explain why we know so much in spite of the fact that we have such limited data. The second tries to explain why we know so little in spite of the multitude of available data. This is the modern paradox of knowledge as CHOMSKY suggests:

> Plato's problem, as compared to Orwell's, looks to me more profound and inciting from an intellectual point of view. But if we won't be able to understand Orwell's problem and recognise its significance in our own cultural and social life, to overcome it, the human species will have few chances to survive long enough to discover the answer to Plato's problem or to other problems inciting intellect and imagination. (Noam CHOMSKY, Knowledge of Language, 1985)

References

BURILEANU C., OANCEA E., RADU M., BURILEANU D., ARHIP J., *Sinteza pornind de la text pentru vocabularul limbii romane*, in "Limbaj si tehnologie", Ed. Dan Tufis, Editura Academiei Romane, 1996.

CHOMSKY N., *Knowledge of Language*, Praeger, New York, 1986.

COJOCARU S., *Lexicon roman*, in "Language and Technology", Ed. Dan Tufis, Editura Academiei Romane, 1996.

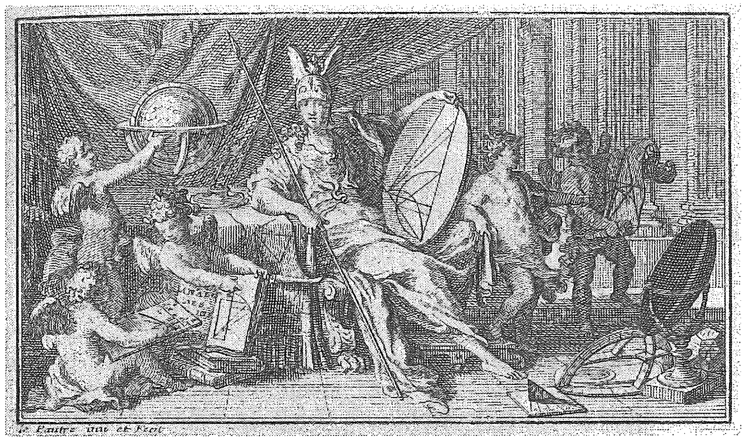
CORNILESCU A., *Concept of Modern Grammar*, Editura Universitatii Bucuresti, 1995.

MONTAGUE R., *Universal Grammar*, Theoria, 3, 1970.

PARVU I., *Introducere in epistemologie*, Polirom, 1998.

PIOTROWSKI R., *Machine Translation*, in "Limbaj si tehnologie", Ed. Dan Tufis, Editura Academiei Romane, 1996.

VON O. QUINE W., *Philosohy of Logic*, Prentice Hall, Englewood, 1970.



Exploring Fregean perspectives in mathematics education

MOREIRA Cândida, Oporto University
GARDINER Tony, University of Birmingham (UK)

> I shall persevere until I find something that is certain - or, at least, until I find for certain that nothing is certain.
> (Descartes)

Abstract

Over the past two decades there has been a major upsurge of interest in the ideas of Gotlob FREGE (1848-1925). Two important and related themes in FREGE's writings are the logical foundations of mathematics and the importance of an appropriate conceptual notation in deriving mathematics from logic. While his thesis that mathematics is a branch of logic and his conceptual notation were deemed to fail, FREGE's perspectives concerning logic were to have a profound impact in both logical and mathematical developments in the 20th century. The goal of this lecture is to review a few of FREGE's ideas, not so much in terms of their contribution to either logic or mathematics, but rather, as the title of this lecture suggests, to try and explore Fregean perspectives for mathematics education. In particular, we explore three aspects of the link between FREGE's work and the way mathematics is taught in schools and universities : 1) the medium through which mathematics is communicated, 2) the nature of mathematical entities; and 3) the distinctiveness of the methodology of mathematics - based as it is, not on everyday ideas, but on abstract objects, exact calculation and proof.