

Part B: Institute-level Reports

Humanities

701 Institute for Bulgarian Language

Introduction

The Institute was founded in 1942 and passed through all periods of Bulgarian post-war history maintaining a national profile.

The Institute is one of the bigger BAS research units with presently 115 positions, of which 111 are occupied. Of these, 75 are researchers (incl. leaders) and 36 supporting staff (PhD-, MA-, and BA- students). During the period under review, there were a total of 34 PhD-students, 10 of whom defended their theses and three were nearly ready. 14 are continuing their studies.

The Institute currently consists of 12 (14) units: 11 thematic departments and 1 Information Centre.

- NSC Dept of Modern Bulgarian language (11 occupied positions);
- NSC Dept for Bulgarian Lexicology and Lexicography (21 ops);
- NSC Dept of Bulgarian Terminology and Terminography (5 ops);
- NSC Dept of Computational Linguistics (6 ops);
- NSC Dept for Bulgarian Dialectology and Linguistic Geography (15 ops);
- Dept of History of the Bulgarian Language (13 ops);
- Dept of Bulgarian Etymology (4 ops);
- Dept of Applied Onomastics (4 ops);
- Dept of Contrastive Research of Bulgarian and Other Languages (4 ops);
- Dept of Ethnolinguistics (11 ops);
- Dept of General and Applied Linguistics (5 ops)

The Information Centre (Library and Administration not counted separately) has eight occupied positions. Five of the 11 departments are “national scientific centres” (NSC), while the remaining ones are called “expert scientific centres”.

The main task of IBL is the research on and description of the Bulgarian language and its dialects and sociolects and of the standard language and its styles, both synchronically and diachronically. One of the tasks is also the promotion of language culture.

Evaluation summary

IBL is the biggest institution of its kind in Bulgaria and elsewhere. It carries out research into various aspects of the Bulgarian language. In terms of personnel, output, and professional recognition this Institute is still very important on the national and international scene. The Institute’s journals are internationally distributed. The Institute holds a key position as regards one of the main objectives of BAS: “to strengthen the national identity, historical, cultural and spiritual values originated by Bulgarian people on its land”. Since it is an institute with mainly national ambitions its importance internationally depends to some extent on the situation for the Bulgarian language and Bulgaristics (and Slavistics) abroad.

Large long-term projects are a cornerstone in the Institute's work as they should be in an academy research institute: such work concerns the description (but also prescription) of the Bulgarian language synchronically and diachronically, dialect descriptions and dialect atlases, dictionary and lexicological work, large-sized electronic corpora, digitization of archives etc. In some areas there is somewhat of a mismatch between these large-scale projects on the one hand, and, on the other hand, the small size of some of the respective departments.

For Bulgarian society, for the study of the Bulgarian language and Slavistics and linguistics in general, the existence of such an Institute is important, especially during the transition period which brings about radical cultural change: the Institute should continue receiving the necessary support to proceed with this work, possibly in cooperation with other BAS institutes, university institutions and institutions abroad. For any long-term planning, it must be understood that such work as the Institute does would not and could not be done anywhere else or by others.

An overall score on **quality and productivity** is made difficult due to the rather divergent fields of research across departments, of which many have a strong emphasis on applied research. By awarding a "B" the Panel indicates that the "*work is internationally visible*" and that "*the Institute has made valuable international contributions to the field*". Computational linguistics moves closer to A, even in an international context, other departments have high visibility for the Bulgarian language community due to their specialised work; however, most of it is published in Bulgarian which restricts its visibility. In terms of **relevance**, the critical mass provided by the Institute makes it still an important player in setting the national language policy, but competition is growing in the universities. As yet, for relevance the overall score is "A" ("*high*"). The Institute has identified relevant themes and partners for the future, and seems to be able to attract young researchers. In terms of **prospects**, the overall score is "A" ("*high*").

Evaluation Report

a) Quality and Productivity

Quality (international recognition and innovative potential)

The main focus of the Institute is on Bulgarian as the national language of Bulgaria. The Institute carries out fundamental and applied research on diverse aspects concerning Bulgarian and has made it its priority to work towards the preservation of linguistic diversity and the preservation of the national language's richness. The Panel noticed a distinct nationalistic undertone, and wondered about the status of minority languages inside Bulgaria.

The Institute has a leading role also due to its critical mass of expertise as a whole – even though in total numbers the more specialised departments may be smaller than individual university departments (e.g.: the section for Contemporary

Bulgarian Language has 11 researchers, all with higher research degrees, whereas the department for the Bulgarian language at the University of Sofia has 27 members of staff of all categories).

A detailed plan of strategies, policies and priorities for future development has been developed and presented along with an identification of the main actors within and outside the Institute, accompanied by a clear sense of what would be the main instruments, partners and outcomes. The list of areas of fundamental and applied research is to some extent a continuation of current work, albeit more focused or elaborated. There is an emphasis on seeking a multiplicity of approaches (including from abroad), alongside rather strong rhetoric on national needs. Either some focusing or a better match of long-term large-scale projects and staffing are necessary.

A priority area is “to provide for complete and consistent language electronic resources, as well as powerful computer programs for their processing”. The site visit and the report on the project Bulgarian WordNet confirmed the impression that research in computational linguistics and corpus linguistics is on an international level. Consequently, 30% of all published contributions in conferences abroad are related to computational linguistics.

Scholars from the Institute participate or have participated in a number of important international projects, securing additional funding from a variety of sources. The scholars of the Institute have made good use of the BAS' bilateral agreements with other academies (visits abroad, mostly in connection with bilateral projects; mostly to Serbia, Czech Republic, Poland, Croatia, Slovakia, Romania, but also to another 21 countries), with some of the researchers staying also for longer periods abroad. And the Institute has been visited by over 100 foreign scientists mostly in connection with joint projects. The Institute collaborates with other academy institutes abroad through bilateral agreements; the cooperation with the Czech academy has been especially fruitful.

Seven researchers sit on editorial boards of foreign journals (or book series). This is a comparatively unsatisfactory representation given the very strong number of active researchers based at the Institute.

Around 18 % of the Institute's scientific production has been published abroad (mostly articles in journals and in proceedings). The ratio is best for papers presented at congresses etc., where however the incidence of computational linguistics is most felt. Four books were published abroad, three of which are conference proceedings. Overall, the computational linguists are leading in international publications in English.

A large part of the Institute's work is “traditional” (though no less necessary and important for that reason): dictionaries, orthographic manuals, grammars, dialect atlases, etymological dictionaries, etc. The Institute has gathered immense archives that are now being digitized. Most of this valuable fundamental work is published in Bulgarian and thus accessible only to Bulgarians (and, to some extent, Slavists). It would be useful to make larger parts of this work accessible also for the wider community of researchers in linguistics. The co-workers of the Institute have also

compiled huge electronic corpora for Bulgarian. Some of them, like the National Corpus (400 000 000 words) are available online. Usage statistics of this and other resources put online show a good interest of the community (single users; though data provided does not differentiate between origin of users of the online material). Almost every year the Institute convenes one or two (summer) schools, most with international participation.

Productivity (scientific output and international standing)

Projects

For the period 2004-2008, IBL reports a total of 98 projects. 16 of these are in collaboration with other BAS institutes, and 32 with international partners. Most projects are funded from the BAS subsidies, but IBL has also tapped into other additional financial sources: the Bulgarian National Science Fund, contracts with ministries etc, contracts and programs in EU, NATO, UNESCO and other international organizations. The Institute has made good use of the Academy's bilateral agreements in some projects. On average, 20% funding from sources other than BAS subsidies, but there are very strong variations from year to year.

17 of the reported scientific projects are single-researcher projects, 23 have two participants, and in the rest of the projects the number of participants varies between 3 and 20. The project "Academic explanatory volume Dictionary of the Bulgarian Language" has the largest number of participants (20). As this project is, and ought to be, among the priorities of BAS, the number of participants is not astonishing.

Teaching

Like other institutes, the Institute's researchers taught in the period many courses (9.579 hours), practicals and seminars (2.814 hours), post-graduate training and specialization courses (350 hours) at Bulgarian universities and in other institutions; they also organized schools and conferences with international participation and participated in lecture courses abroad. They supervised theses of 47 Bachelor and Master students at Bulgarian universities. 35 of the researchers at the Institute took part in these activities although the main burden of education is carried by a smaller group of scholars. In the SER the connection to the universities is commented upon and placed in the context of Bulgarian labour law (educational activities should not exceed 20 % of the researchers' time). Conversely, the Institute argues, research activities should be up to 20% of a regular university lecturers' activities.

Contacts between IBL and the universities are to a certain extent unilateral. Agreements with some universities for "collaboration and joint work" are one way out of the situation, but the best way is probably through joint research projects. An example for such a research project is "Language and idiolect" with research partners in Sofia University, University of Veliko Tarnovo and University of Shumen.

In the five years under review, 34 PhD-students were educated at the Institute. For the same period, 15 new PhD-students were accepted – 11 regular and 4 correspondent, and 10 PhD-students successfully defended their dissertations. Three PhD-students are ready for a defence, and 14 continue their studies. This corresponds to a dropout rate of around 20% in the period, a figure that is somewhat unsatisfactory. Those PhD students that stay on are encouraged to publish: the documentation points out that during the period 2004-2008 PhD students published 71 scientific articles and notes and presented 53 reports at scientific fora.

Collaboration

IBL collaborates with other institutes at BAS mostly in areas of computational linguistics and corpus linguistics.

An area where collaboration could be enhanced is the study of the history of Bulgarian and the study of Old Bulgarian / Old Church Slavic/Old Slavic (language, text and manuscript studies), an area that the Department of History of Bulgarian Language covers at least partly. Their researchers have among other things published a two-volume Old Bulgarian dictionary. This is a popular research area, where interests intersect from Balkan Studies, History, Literature and, of course, Cyrillo-Methodian Studies. The project “Legal texts from the Middle Ages” is a step towards more cooperation in this area (CMRC; Institute of History; Department of History of Bulgarian language).

The Institute is also a partner in important European projects: European Linguistic Atlas, Slavic Linguistic Atlas, projects connected to the Fifth, Sixth and Seventh Framework Programmes etc. Under FP6, the Institute was involved in “KT DIGICULT: Knowledge Transfer for digitization of Cultural and Scientific Heritage in Bulgaria”, and under FP7 as task leader in “Discovering our Past World with Digitized Maps (DIGMAP)” and notably as consortium partner in the important research infrastructure project “CLARIN – Common Language resources and Technologies Infrastructure”.

Publications

In the period under review the Institute published 80 books, and 906 studies and articles, as well as 15 textbooks/teaching aids and 177 more popular articles. Around 18% of the scientific production is published abroad. The amount of published books and articles is respectable considering that most of the researchers work in long-term projects and that many researchers also teach at universities. Due to the subject and the scientific audience, most of the texts are published in Bulgarian and/or for a Bulgarian-reading public.

Most of the reported five top scientific achievements and main applied results are fruit of teamwork, which is important and shows the degree of cooperation inside the Institute.

IBL publishes two important journals: “*Balgarski ezik*” / The Bulgarian language (in Bulgarian), which for many years has been the most important Bulgaristic journal,

and “*Balkansko ezikoznanie. Linguistic Balkanique*” (in foreign languages). Both titles have been submitted to ERIH.

An overall score on **quality and productivity** is made difficult due to the very divergent fields of research across departments. By awarding a “**B**” the Panel indicates that the “*work is internationally visible*” and that “*the Institute has made valuable international contributions to the field*”. Computational linguistics moves closer to A, even in an international context, other departments have high visibility for the Bulgarian language community due to their specialised work; however, most of it is published in Bulgarian which restricts its visibility beyond that relatively narrow circle of readers.

b) Relevance (socio-economic impact)

Politically, the Institute is important since BAS is a central authority in shaping national discourses. The Institute has influence on the national language policy, and takes part in many councils, commissions and other expert bodies. The Institute gives expert advice and provides services to ministries, the National Assembly, the Association of Bulgarian Teachers, but also to different units of the Bulgarian Court system. It reaches wider publics by publishing orthographic manuals, grammars and textbooks. Its Bureau for Language Consultations is important in this area, and numerous more popular appearances in radio and television on questions on language contribute to this task. Increasingly, such roles are being taken forward through European consortia: the first example is the collective long-term project “National Languages in the European Communities”, membership in the European Federation of National Institutions for Language, and participation in the FP 7 project proposal “A Common Language for Europe or Multilingualism”.

Institute researchers teach at universities across the country and also have other forms of collaboration with them. Within the scientific community, the institute has a leading position due to its size and the fact that it has experts in all the areas covered by the Institute. For some of these areas there does not seem to be any competition from the universities, but in some areas, as for example modern Bulgarian, the universities are relatively well-equipped with qualified personnel.

IBL collaborates with other institutes at BAS: interdisciplinary research is conducted with other humanities and social sciences (folklore, literature, sociology, philosophy, pedagogy, and history) in 17 projects during the last five years, but also with natural sciences in the computer processing of natural language: mathematics and informatics (15 projects during the last five years).

In terms of **relevance**, the critical mass provided by the Institute makes it an important player in setting the national language policy, but competition is growing in the universities. As yet, for relevance the overall score is “**A**” (“*high*”).

c) Prospects (vitality and feasibility, management & leadership; future potential and ability of the Institute to tackle new scientific challenges)

The Institute went through a restructuring at the beginning of 2008 which left, out of the previous seventeen units, a mere eleven plus one Information Centre. The effectiveness of this reform cannot yet be gauged.

The age structure does not seem to be a great problem at present, although extrapolations produce a less rosy picture for the near future. Currently there are 10 staff in the age group 26-30, 26 in the age group 30-40, 18 in the age group 40-50, 38 in the age group 50-60, 14 in the age group 60-70, and one over 70. The age group 40-50 is most worrying, and since there is not one single professor or a senior research fellow I degree among them. The Institute risks having a lack of supervisors for PhD-students, leaders of research teams etc. Some consider it may become necessary to recruit people from outside BAS for this purpose, even to leading positions or at least solve these problems combining efforts with the universities. The Institute traces this difficulty to the last decade of the 20th century, when it was decided to freeze the PhD program at the Institute. The current solution is to engage university lecturers on supplementary contracts as supervisors of PhD-students (currently the case for six).

On the other hand, the Institute seems to have resolved the recruitment of young scholars – 30 were accepted during the period under review as researchers and investigators: however, two defended their dissertations abroad, another three have become PhDs at Bulgarian universities. De facto this means that the production of PhDs is not sufficient to satisfy the Institute's own needs.

The Institute is trying to make an academic career more attractive by including young scholars in important research projects: two of new projects are starting in 2009 (funding under the last call of the Structural Funds in the section for “Advance of PhD students, post-doctorants and young scientists”). One is a joint project with the Faculty of Mathematics and Informatics at Sofia University (*Mathematical logic and computational linguistics: development and mutual penetration* –worth 330 000 BGN of which about half for the Institute, with participation of 11 PhD students, post-doctoral researchers and young scientists from the Institute). The other is a joint project with the Faculty of Slavic Studies at Sofia University (*Computational and interactive means for historical language research* – worth 400 000 BGN, of which about $\frac{1}{4}$ supports the participation of 7 PhD students, post-doctoral researchers and young scientists based at the Institute). These are good developments in the right direction: support for young talent, cross-institutional successful competitive applications, interdisciplinary, research-infrastructure based, ambitious research projects with a potential appeal also for international audiences.

The Institute has identified relevant themes and partners for the future, and seems to be able to attract young researchers. In terms of **prospects**, the overall score is “A” (“high”).

Overall Strengths and Weaknesses

Strengths

- Some sections (computational linguistics) are on a good internationally level, on a par with other departments of computational linguistics;
- Important position in national language policy;
- Good portfolio of strategic priority themes and projects (research infrastructures);
- Good international network.

Weaknesses

- Problematic age structure (both at junior level and at mid-career level);
- Uneven quality of research departments;
- As yet insufficient international visibility of most departments.

Recommendations

An overall score on **quality and productivity** is made difficult due to the very divergent fields of the research across departments. By awarding a “B” the Panel indicates that the “*work is internationally visible*” and that “*the Institute has made valuable international contributions to the field*”. Computational linguistics moves closer to A, even in an international context, other departments have high visibility for the Bulgarian language community due to their specialised work; however, most of it is published in Bulgarian which restricts its visibility. In terms of **relevance**, the critical mass provided by the Institute makes it still an important player in setting the national language policy, but competition is growing in the universities. As yet, for relevance the overall score is “A” (“*high*”). The Institute has identified relevant themes and partners for the future, and seems to be able to attract young researchers. In terms of **prospects**, the overall score is “A” (“*high*”).

Against this background, the following recommendations aim at suggesting some general and specific ways forward:

- maintain the balance between exploration of new research fields (computational linguistics) and “traditional” areas of research (for example etymology, (applied) onomastics, dialectology etc.) as a typical mix of tasks for a national institute for language;
- support long-term research projects, and develop a strategy for including them in research infrastructure plans;
- make best use of the BAS network of institutes through targeted cooperations as well as with university institutions (as in the newly created network “Language resources in the Humanities and Social Sciences”);
- reflect on the appropriate staffing of long-term projects;
- consider a possible new project: network for discussing the theoretical and practical basis for a new Academy Grammar of Bulgarian.