

1. Summary of the research plan

The goal of the proposed project is to study connectors in the Russian language. The study will be corpus-based. One of its core points is that Russian connectors are approached through an original classification of logical-semantic relations. The classification is based on semantic principles and formed by combining such parameters as:

- Type of the underlying semantic operation for each relation (implication, comparison, location on the chronological scale, correlation between an element and a set);
- Linguistic level on which this relation is established (proposition, speech acts, text structure).

The proposed classification is superior to the existing ones as it shows a semantic mechanism common for several types of relations, thus locating each relation in the system. Also, it allows for demonstrating what is similar and different about contiguous semantic categories. Three groups of connectors will be tested in order to see how functional the classification is. These are the connectors expressing: 1) a *concomitance relation* that has been ignored so far by Russian, French, and Anglo-Saxon linguists; 2) a *relation of enunciative addition* that has not been given any description in contemporary linguistics either; 3) a *reformulation relation* (also dubbed as periphrastic or explicative relation in other linguistic traditions).

The project will be focused on a little-investigated group of multiword connectors in order to locate them in the system of the Russian language. This becomes increasingly important with further development of corpus linguistics and new information resources, namely annotated corpora and supracorpora databases. The connectors will be described employing an original technique of unidirectional contrastive analysis. It makes use of the Russian-French corpus, which will add a statistical dimension to the semantic analysis and help to reveal language-specific connectors in the compared languages.

Within the project, it is expected to build a supracorpora database of Russian connectors containing at least 5,000 *monoequivalences*, a monoequivalence being a pair in which a Russian connector is aligned with its French equivalent. The data will be extracted from the Russian-French parallel corpus of at least 2,000,000 words. After the completion of the project, the database will be available on the Internet.

Keywords: corpus linguistics, contrastive linguistics, semantics, text linguistics, logical-semantic relations, connectors, parallel corpora, supracorpora databases, Russian, French.

2. Research plan

2.1. Current state of research in the field

Discourse markers (including connectors) have been investigated since the early 1970s, and the bulk of scientific knowledge is continuously growing. Yet, despite numerous studies of different markers in various languages, this research field remains open, with no terminology or methodology settled. It is noteworthy that the scholars have not come to any consensus about either what has to be included in the class of discourse markers or what place is taken by connectors in it. Very often, studies of discourse markers are heavily dependent on some particular linguistic tradition. This is evidenced by recent special thematic issues of the journals *Revue française de linguistique appliquée* (XVI-2, 2011¹) and *Langages* (184, 2011²). Together with two volumes on discourse markers (Fischer 2006, Aijmer & Simon-Vandenberg 2006), they suggest that the data from the Russian language are very seldom, if at all, taken into account.

¹ D. Flament Boistrancourt & A. Trévisé (éds), *Les connecteurs : descriptions, traduction, apprentissage*.

² A. Rodríguez Somolinos (éd.), *Les marqueurs du discours : approches contrastives*.

Moreover, the number of works on this issue varies greatly with the language considered. Anglo-American linguists productively explore connectors and relations expressed by them. This is also the case of the German (cf. specifically, the fundamental project on German connectors that resulted in publishing two volumes Pasch & *al.* 2003, Breindl & *al.* 2014). However, there are much less Italian and French studies of the kind. In Russian linguistics, the interest in discourse words arises in the mid-1980s. The most important works concerned with the phenomenon are: Nikolaeva (1985, 2004), Ljapon (1986), Sannikov (1989), Čeremisina & Kolosova (1987), Baranov & *al.* (1993), Paillard & *al.* (1998), Paillard & Kiseleva (2003), three volumes resulted from a French-Russian project on discourse words and giving an analysis of slightly more than twenty linguistic items (*Les particules énonciatives en russe contemporain* 1986-1987). So far, the theory of Russian grammar has been predominantly focused on one specific subset of connectors, that of coordinate conjunctions (see works by T. Nikolaeva, I. Fougeron, E. Uryson, V. Sannikov). With the appearance of the above works, one could see a sharp decrease of interest in the topic, while a number of questions have remained unsolved, which holds not only for Russian but for all Slavic languages. This is proved by the appropriate section in the authoritative publication summarizing research on Slavic studies *Die Slavischen Sprachen / The Slavic Languages* edited by Kempgen & *al.* (2009). There, connectors and logical-semantic relations were given a total of 10 of the more than 1,000 pages.

In contemporary linguistics, there are two main directions of research in the domain of connectors:

- Description of semantic, syntactic, and pragmatic properties of connectors;
- Description of logical-semantic relations expressed by connectors and their classification.

As regards the second direction, today's specialists in text linguistics, semantics, and pragmatics still hotly debate the question of how to classify logical-semantic relations between utterances and larger text units, given the importance of these relations for text coherence. Here, a researcher would possibly have two solutions proposed. The first one implies the classification to be generally based on the linguistic markers of these relations either in one language, mostly in English (e.g., Warner 1979, Halliday [1985] 1994, Blakmore 2002, Fraser 2006, 2009), or in several languages (e.g., Kortmann 1997), with each relation being assigned to a corresponding prototypical marker. Followers of the cognitive and pragmatic approaches assume, on the contrary, that a relation can be established between two text segments but it may also remain unexpressed by any special marker.

Whatever the approach adopted, the relations are presented as a list that remains open and changes, in the first case, from language to language, and in the second, from study to study. With different analysis techniques, scholars have more or less details, and, theoretically, the list can be infinitely expanded because the basis of such classifications is not clearly evident. Thus, in Kibrik & Podlesskaja (2000), to the initial list of "rhetorical relations", borrowed from Mann & Thompson (1988) and containing 15 basic/12 additive relations, another 36 were added.

Subsequently, in order to limit the number of possible relations that can be established between parts of a text during its interpretation, cognitive researchers proposed to classify them according to basic cognitive operations. While in Knott & Sanders (1998) there are only two such operations – "causal" and "additive", in Charolles (1995) three of them can be found – association by similarity, contiguity in space and time, and by causality (borrowed from the treatise of 1748 "Research on Human Understanding" by D. Hume).

On the one hand, it is more than obvious, however, and it was noted by Hume himself, that these basic operations do not cover all the different logical-semantic relations. On the other hand, many of the proposed classifications do not differentiate between the levels, where a relation is established. Even when this differentiation is made (in both cognitive and semantic studies), it is represented as a binary opposition: the level of proposition *vs.* the level of speech act (*semantic vs. pragmatic*), e.g., in the above-mentioned work by Knott & Sanders (1998), or proposition *vs.* text structure, e.g., in Breindl & *al.* (2014). It is also assumed that if a relation is established on one of these levels, it cannot

be established on another, which clearly disagrees with the linguistic facts. For instance, an oppositional relation can be established on the levels both of proposition and of speech act (Inkova-Manzotti 2001). Similarly, a temporal relation or a relation of generalization manifests itself at two levels – those of proposition and of text structure (Inkova & Manzotti in press)³.

In Russian linguistics a classification of logical-semantic relations was devised as early as in the middle of the last century, with mainly syntactic criteria taken into account (*Russkaja grammatika* 1980). The validity of some relations (e.g., scalar) is questionable, while certain relations (e.g., those of enunciative addition [prisoedinitelnye otnošenija], explanatory [pojasnitelnye otnošenija], etc.) require a more precise definition.

The description of connectors, in its turn, is also approached in two different ways. It can be either onomasiological or semasiological. Onomasiologists focus on a particular type of relations to analyze its possible linguistic markers (cf. the analysis of a concessive relation in Russian (Apresjan 2006); the contrastive Russian-French study of an oppositional relation (Inkova-Manzotti 2001); works on a causal relation (Boguslavskaja & Levontina 2002, 2004). In doing so, researchers do not generally seek to locate the relation in the whole system of logical-semantic relations.

Semasiologists describe how a specific connector or a group of connectors functions so as to determine the type of relation that could be expressed by them. A major disadvantage of such studies is that they often do not question whether the analyzed item can be universally used as a connector, whatever the circumstances are. Yet, it is well known that many linguistic items, being able to function as connectors, are multifunctional. Therefore, they cannot always be classified as connectors. Where to draw the boundaries of the functional class of connectors? What are the criteria to be used to decide whether a discourse marker is a functional connector? These are the questions that remain unsettled, and the attempts to answer them are quite different (see, e.g., Rossari 2002, Fraser 2009, De Cesare & Borreguero Zuloaga 2014). Nevertheless, the questions are still of crucial importance, as this class is constantly growing. In this respect, it is necessary to set a border between the functional class of connectors and similar phenomena, e.g., anaphoric adverbials with sentential scope (cf. *в этом отношении/в etom otnoženii* [in this respect] vs *по этому поводу/по etomu povodu* [because of this/due to this/then/so]). Both types of linguistic items refer to the left-hand context, but in the right-hand context their functions differ. While the former sets up the “world”, or, more precisely in this case, a perspective in which the statement is true, the latter can be classified as a connector denoting a causal relation, cf. (1) and (2):

- (1) Немаловажной для выпускников является оценка организации своего досуга. *В этом отношении* родной город значительно уступает городу предполагаемого отъезда.
[Evaluating the organization of leisure time is not unimportant for graduates. *In this respect*, the hometown is much less attractive than another possible place to live in.]
- (2) Работаю, сынулька мой иногда болеет, *по этому поводу* мы бываем на больничном.

[I’m working, and sometimes my little sonny happens to fall ill, *then* we take a sick leave.]

Also remain under-investigated such phenomena as combinations of connectors (Oates 2001, Fraser 2013) and multiword connectors, including the problem of differentiating between the former and the latter. For example, English construction *but, on the other hand* is considered to be a combination of connectors (Fraser 2013) whereas this conclusion is not so obvious when it comes to French *et puis* [and also] or Russian *а может быть/а может быт’* [but maybe]. In Russian linguistics, this problem was raised as early as in 1977 by Prijatkina (1977) but until now it has not received its due attention. Prijatkina maintains that Russian is notable for its abundance of “connector combinations”,

³ On the need to differentiate between three semantic levels (which the authors refer to as *des relations de discours sémantiques, textuelles et praxéologiques*) see Roulet, Fillettaz & Grobet (2001).

i.e., combinations of a coordinate conjunction (*u/i* [and], *a/a* [but/and], *no/no* [but], *da/da* [and], *ili/ili* [or]) with particles, modals, adverbs or “conjunction alternatives”. One should note that elements of such combinations may be positioned either adjacently or distantly (*но однако/но однако* [but], *не то чтобы... a/ne to čtoby... a* [not that... but], *или может быть/ili možeť byt'* [or maybe], etc.). There is no doubt that combinations such as these are quite stable, but should each of them be considered a distinct connector that may be opposed to all the other Russian connectors? This question has not been properly investigated yet. In contemporary studies of the Russian language such combinations are usually considered idioms (see, e.g., Mustajoki & Kopotev 2004) but this approach is hardly viable. For example, the combination *не только ... но (u)/ne tol'ko ... no (i)* [not only... but (also)] is traditionally regarded as a distinct conjunction. On closer investigation though, one soon discovers that its form is highly variable: *не только/не tol'ko* [not only] may be combined not only with *но/no* [but] but also with *a/a* [but]; the analogy particle *u/i* [also] may be preceded by *еще/еще* [then] or *также/takže* [as well], or be omitted. Consider the following example of usage of *a/a* [but] instead of *но/no* [but] in this combination:

- (3) Оборонная безопасность должна обеспечиваться *не только* военной силой, *а* всеми силами и средствами государства по следующим основным направлениям.
 [The defense capability must be *not only* due to the military power *but* to all the powers and means of the state related to the following essential activities.]

Furthermore, an utterance with *не только/не tol'ko* [not only] may be used without any conclusion, or, on the contrary, it itself may follow conjunction *a/a* [but/and] (*а не только/а не tol'ko* [and not only]). Faced with such extreme variability, developers of the annotated Russian corpus HANCO (Helsinki Annotated Corpus of Russian Texts, <http://www.helsinki.fi/venaja/english/e-material/hanco/>) were forced to include in their list of multiword grammatical items 10 synonymous (!) “phrasemes” with *не только/не tol'ko* [not only]. Moreover, the variant *а не только/а не tol'ko* [and not only] appears in the list as totally unrelated to these “phrasemes” while the item *но u/no i* [but also] is qualified as an independent multiword conjunction. We believe that it is more reasonable to consider these items as free collocations even though they may be quite frequent and predictable (cf. similar approach to French *non seulement... mais* [not only... but] (Montchaud 2015)). However, this solution does not apply to all connector combinations; e.g., cf. collocation *как... так u/kak ... tak i* [just as... so] in the following example (4). This item should be qualified as a multiword connector that establishes the relation of analogy (Inkova 2014) and not as a combination of connectors.

- (4) Как хороший каллиграф не имеет своего почерка, а имеет все, *так и* актер не должен иметь на сцене своей походки, а иметь всякие.
 [Just as a good scribe does not have a handwriting of his own but has them all, *so* a good actor should not have his own demeanor on the stage but all kinds.]

Developments in traditional corpus linguistics, on the one hand, and emergence of new kinds of linguistic information resources (annotated corpora, parallel corpora, supracorpora databases), on the other hand, have created a reliable theoretical and empirical basis for solving these urgent problems.

2.2. Current state of your own research and partnership aspect

Problems related to description of logical-semantic relations and their linguistic markers have remained the focal point of O. INKOVA-MANZOTTI's scientific interests for over 15 years. Her doctoral thesis was centered on this subject (Inkova-Manzotti 2001) and included comparative analysis of contrastive connectors in French and Russian. Her subsequent research focused on anaphoric connectors in French (Inkova-Manzotti 2002) and later in Russian (Inkova-Manzotti 2005). Some of this research was performed as part of the project on semantic classification of French connectors that was supported by The Swiss National Science Foundation (No. 610-062821). Various aspects of their

functioning were examined, including interplay between anaphoric connectors and scalarity (Inkova 2009a, 2010c) or expressiveness (Inkova 2009b, 2010b); functioning of anaphoric pronouns in correlative structures (Inkova 2008, 2010a, 2011a, 2011b, 2013a).

In Russian linguistics, connection means are divided into conjunctions and relative pronouns based on the dual function of the latter (connective and syntactic). Inkova's research has shown, in particular, that this division often results in a contradictory classification of connection means and complex sentences. For example, such linguistic items as *как/kak* [how; as] и *когда/kogda* [when] are considered to be homonyms in their correlative vs. non-correlative use. Some connection markers (*насколько/naskol'ko* [as; as much as], *настолько/nastol'ko* [so]) are qualified as conjunctions solely based on their initial position in a sentence (and as pronouns in other cases).

Inkova's analysis demonstrated that consistent and universal description of such items is possible if we take into account their scope (i.e., whether they modify a word or a sentence). Results of this research were published in *La corrélation en russe: structures et interprétations* (Inkova 2014a); the publication was financed by The Swiss National Science Foundation (B-0010_154261). Also covered by Inkova's latest research are:

- Issues of differentiation between language levels on which relations are realized by connectors, which is required to trace the meaning common to different patterns of their use (Inkova 2013b);
- Issues of classification of logical-semantic relations (Inkova 2013c);
- Propositions for defining new relations – those of analogy (Inkova 2014b) and concomitance (Inkova in press1);
- Propositions on how to define criteria that would allow to classify a linguistic item as belonging to the functional class of connectors (Inkova in press2).

Anna ZALIZNIAK is the author of several books on Russian lexical and grammatical semantics (Zalizniak 1992, 2006, 2013; Zalizniak, Levontina & Šmelev 2005, 2012). She is also the author of a number of articles in which a unique methodology of unidirectional contrastive corpus analysis was proposed and implemented. These articles also describe linguistic interfaces for the database of Russian verbal forms with French translation equivalents and for the database of Russian language-specific words (Buntman & al. 2014, KruzHKov & al. 2014, Zalizniak 2015). Furthermore, she has developed the methodology for investigating semantics of language-specific connectors and proposed a description of one of such Russian items – the conjunction *a/a* [but] (Zalizniak & Mikaelian 2005, Mikaelian & Zalizniak 2005).

Irina LEVONTINA is the author of numerous works on grammatical words, including causal relations (Boguslavskaja & Levontina 2002, 2004), causal prepositions, discourse words, particles, and connectors (Levontina 2011, 2013, 2014). In cooperation with D. O. Dovrovol'skij, she has conducted corpus-based contrastive studies of discourse words (e.g., Dovrovol'skij & Levontina 2012).

Nadezhda BUNTMAN and Vitaly NURIEV conducted a number of contrastive linguistics studies, including contrastive analysis of verbal forms in Russian and French and of Russian markers of indefiniteness. They carried out an expert evaluation of several thousand correspondences between Russian items (verbal forms, language-specific items) and their French translations (KruzHKov & al. 2014; Zatsman & al. 2014; Buntman & al. 2014). Nadezhda Buntman is an expert in stylistics and translation studies; she has an extensive experience in teaching these disciplines and French in Lomonosov Moscow State University (MSU) and in the University of Lausanne (Switzerland). She has translated more than 20 works of French literature into Russian. In cooperation with her French colleagues, she has proposed a new approach to creation of a corpus-based typology of translation difficulties (Buntman *et al.* 2010). Vitaly Nuriev is an expert in translation studies. He has experience of teaching French and English at the Department of foreign languages of the Institute of Linguistics (Russian Academy of Sciences). He has translated numerous works

of French and English literature into Russian (including works by Françoise Sagan, Régis de Sá Moreira, Neil Gaiman). His scientific interests include the problem of literary translation's aging. He was twice (in 2010 and 2014) elected a member of the Maurice Wachsmacher prize jury (prize for the best translation of French literature into Russian awarded annually by the Embassy of France in the Russian Federation).

In a series of papers, Igor ZATSMAN described the methodology for developing databases designed to support contrastive analysis of grammatical forms and their translations in parallel texts (Kruzhkov & al. 2014; Zatsman & al. 2014; Buntman & al. 2014). Such analysis can bring to light tacit knowledge of professional translators that has never been reflected in existing grammars. The methodology proposed by Zatsman does not rely on specific language pairs, constructions types or number of available translations. It was used in contrastive analysis of Russian verbal forms and their French translations and proved to be entirely feasible. By using this methodology, Russian members of the proposed project team were able to extract and annotate approximately 10,000 Russian verbal forms and their translation correspondences and statistically process these data. For example, experimental analysis of 2,000 Russian present tense verbal forms allowed to discover at least eight new models for translating such forms into French, in addition to nine models that had already been described in contrastive grammars. The methodology allowed to calculate frequencies for the 17 revealed models and to investigate how each of them correlates with various combinations of context properties of Russian verbal forms.

Mikhail KRUZHKOVA developed methods, software, and information technologies for creation and use of *supracorpora databases* – a new type of linguistic information resources. He also developed a new method of multilingual lexical/grammatical search in cross-linguistic databases. This method supports three types of search functions: those based on lexis, those based on grammatical properties and the combination of the above (Kruzhkov & al. 2014; Zatsman & al. 2014). The proposed method proved its feasibility after being successfully implemented in the Database of Russian and French Verbal Forms (http://a179.ipi.ac.ru/corpora_dynasty/main.aspx).

In 2014 Anastacia KOCHETKOVA POQUERUS defended her graduate thesis on Russian connectors of enunciative addition (*Les connecteurs d'ajout énonciatif en russe*). At present she works on her post-graduate thesis "*Les adverbiaux cadratifs anaphoriques du russe : sémantique et modalités du fonctionnement*" [Russian anaphoric adverbials with sentential scope: semantics and modes of operation] at the Faculty of Arts [*Faculté des Lettres*] of the University of Geneva, under supervision of O. INKOVA-MANZOTTI. She plans to define the specified class of linguistic items and to describe their semantic and syntactic features.

The younger Russian members of the project team, Olga PETRUSHKINA and Natalia POPKOVA, have a considerable experience in building mono-equivalences, i.e., aligned correspondences between Russian linguistic items and their French functionally equivalent fragments. Their skills will be extremely useful when it comes to the task of filling up the supracorpora database of connectors, the main information resource of the proposed project.

Areas of scientific expertise of Russian and Swiss partners complement and reinforce each other. The methods of semantic and contrastive analysis have been developed by A. Zalizniak and O. Inkova-Manzotti, the leaders of the project. The methodological approach and information tools have been jointly tested by the Swiss and Russian members of the proposed project while the first version of the supracorpora database of Russian connectors and their French translations was being developed. This database is currently used for study of Russian connectors. Initial results of this research will be jointly reported at the International Conference on Contrastive Text Linguistics "*Langues slaves en contraste*" in Bergamo, May 2015 (Kruzhkov & Inkova 2015), while another joint report was submitted to the International Conference "Corpus linguistics 2015" that will take place in Saint-Petersburg, June 2015 (Zalizniak & al. 2015, in print).

2.3. Detailed research plan

Based on the above (§ 2.1.), the current research is aimed to:

- 1) Develop theoretical principles of semantic description for Russian connectors. Connectors expressing relations of enunciative addition, reformulation, and concomitance will be taken as an example;
- 2) Fill the supracorpora database with Russian connectors and their French translation equivalents: make at least 5,000 monoequivalences (i.e., a pair in which a source text fragment is aligned with one of its French translations, hereinafter ME), with the text data collected from two-way human translations (Russian-French and French-Russian), as well as from Russian-French machine translations;
- 3) Develop theoretical principles of linguistic description and presentation in a supracorpora database of simple and multiword connectors;
- 4) Statistically process the total number of ME built and identify relevant criteria (features and characteristics) for Russian connectors and their translation equivalents;
- 5) Describe the types of interlanguage correspondence in human and machine translation, including types of the most frequent errors of the latter;
- 6) Give a linguistic interpretation of statistical data that would help to evaluate the language-specific semantics of Russian connectors.

2.3.1. Developing theoretical principles for semantic description of Russian connectors

An original classification of logical-semantic relations based on semantic principles is to serve as the ground for description of connectors. First of all, it is proposed to separate the conceptual and linguistic aspects in classifying logical-semantic relations. There is no one-to-one correspondence between types of logical-semantic relations and their markers (e.g., adversative relations may be expressed with conjunction *no/no* [but] or *a* [but/and], yet on the other hand, *a* [but/and] can express not only adversative, but also contrastive relations, as well as relations of enunciative addition (*prisoedinitelnye otnošenija*)).

We deem reasonable to proceed from the four semantic operations (implication, comparison, location on the chronological scale, correlation between an element and a set) that underlie the whole of logical-semantic relations. This approach will allow us to produce a universal classification applicable to all languages, given that the differences between languages, as shown with the example of relation of opposition (Inkova-Manzotti 2001), manifest themselves either on the level of particular types of a certain relation or in its discourse representations. The advantage of this classification over the existing ones (see. § 2.1.) is that it enables us to show the semantic mechanism, common to several types of relations, and thus to locate a particular relation within the system.

The principal difference of the proposed classification from the classifications based on cognitive principles is that cognitive linguistics attempts to identify the basic cognitive mechanisms. These mechanisms underlie text coherence and assist speakers in decoding a text. The basic criteria of such classifications are primarily interpretive and rely on the hypothesis about speaker's communicative intention. This approach makes the issue of linguistic means that help to mark relations rather insignificant. The proposed classification, on the contrary, encompasses the linguistic means used by speakers to shape a text and to express a particular relation between the components of utterances.

Our classification is presented as a two-input table (with two attributes): 1) type of relation, with the indication of its underlying semantic operation (e.g., "implication" is the underlying operation for relations of cause, consequence, condition, purpose, concession); 2) linguistic level on which the relation is established (proposition, speech act, text structure). The advantage of this representation is that it allows us to see that several relations distinguished in the linguistic literature are in fact one and the same relation, but established on different linguistic levels (cf. relations of

cause and motivation, of junction and enunciative addition). The proposed classification will lay groundwork for developing semantic description of connectors expressing relations of enunciative addition, reformulation, and concomitance.

Limiting the scope of research in the proposed project to three semantic classes of connectors (those expressing relations of enunciative addition, reformulation, and concomitance) is motivated primarily by the fact that they are the least studied in the Russian language, in comparison to relations of cause, concession or opposition (see § 2.1. above). In addition, what all three classes of relations have in common is that they are based on a comparison operation. It establishes a similarity between text fragments linked together by connectors. However, connectors establish these relations on different linguistic levels. Finally, many of the existing studies rely on an *a priori* classification, without questioning the validity of such analysis (cf., e.g., the study of connectors expressing a causal relation in Čžon 2003).

The choice of the above connectors is also justified by the fact that they allow us to comprehensively examine many unexplored aspects. Thus, the semantic description of connectors classified in the Russian grammar as those of enunciative addition and explicative (here referred to as connectors of reformulation) is to give a more precise definition of logical-semantic relations themselves in consistence with the current linguistic theory. To study connectors expressing relations of concomitance, it is necessary to reveal the criteria for assignment of a particular linguistic item to the functional class of connectors (cf. *при этом/при etom* [while] that may be the circumstantial modifier without a connective function. Nevertheless, due to its anaphoric nature, it alludes to the left context: *Ящики с картинами раскрыли в среду. При этом присутствовал милиционер* [The paintings were unboxed on Wednesday, while a policeman was present]). A number of connectors from these semantic classes are multiword; cf., e.g., connectors formed with the conjunction *да/да* [and] (*да u/да i* [and indeed], *да еще/да ešče* [and also], *да еще u//да ešče i* [and moreover], *еще u/ešče i* [as well] and *да u нрпом/да i притом* [and then]). Finally, many of them are also language-specific for the language pair of Russian-French (see §§ 2.3.3. and 2.3.4-6. below).

2.3.2. Operating the supracorpora database of Russian connectors and their French translation equivalents

The study proposed within the framework of the project is of contrastive nature. Its distinguishing characteristic as opposed to existing comparative studies of discourse markers (see, in particular, the French-German project *Les Invariables difficiles*, Métrich & al. 2002), lies in applied methods and computer tools. Our study utilizes an original unidirectional method of contrastive analysis (see Buntman & al. 2014, Zalizniak 2015 (in press)). Making use of parallel corpora, this method is based on considering translation of lexical items of the Russian language as a tool for semantic analysis, namely as quasi-interpretation, that provides a source of additional information about the semantics of studied items in Russian. Thus, comparing two languages is not an aim in itself, but rather a tool for analyzing one of them.

On the basis of the unidirectional principle, the project team has developed an original set of methods and information technologies for contrastive corpus analysis that have been used for corpus-base processing of both Russian-French and French-Russian parallel texts. In the case of Russian-French translation, each occurrence of the linguistic item under consideration is assigned to its functionally equivalent fragment. In doing so, the Russian text is analyzed to identify some relevant characteristics that could have influenced the choice of linguistic means in the target language. In the case of French-Russian translation, the semantics of the Russian item in question is defined by different characteristics of the foreign language text, since they determine the appearance of this particular item in translation. In both cases, a translation, created by a professional translator, provides a source of valuable information on the semantics of the item under consideration both in the source language and in the target one.

The project team has created special methodical and computer tools that enable text alignment process to be executed on the level of the linguistic items and their functionally equivalent fragments (Buntman & al. 2014). These tools were employed to build a database of Russian connectors that contains aligned texts of Russian literary works and their French translations (up to four translations for each source text). Such databases were labeled supracorpora databases (SCDB). Besides text alignment at the level of the source linguistic items and their functionally equivalent fragments from the target texts, the SCDB allows us to register additional characteristics of the source linguistic item and its functional translation equivalent as well as those of the context of their use. In the SCDB, the data on connectors are presented in the form of ME, i.e., pairs in which a Russian connector is aligned with its French equivalent (see Table 1 below).

The first column gives the number of the ME. The second column shows the minimal right-hand and left-hand context in which the connector set in bold type occurs (a broader context is also accessible in the database). The third column (type and additional linguistic characteristics of the source item) lists the characteristics relevant for describing of how connectors function. At the current stage of the database development, these characteristics are limited to: type of logical-semantic relations, syntactic type of a text fragment marked by a connector, position of the connector in this fragment, order in which the components linked by the connector come. The next column lists French functionally equivalent fragments and their additional characteristics. The connector may as well be absent in the translation. In this case it is denoted by **zero**. The denomination is introduced for the cases when a Russian connector does not have a lexical equivalent in the French translation. Characteristics for other translation means (a certain verbal form, subordinate clause, participle, etc.) are to be developed as part of the proposed project. The last column shows the special features, if any, of a particular ME. For instance, since the same linguistic items can have different functions, the connective one is marked with a special attribute (Cnt).

№ МЭ	ЛГФ оригинала	Вид и доп. признаки ЛГФ оригинала	ЛГФ перевода	Вид и доп. признаки ЛГФ перевода	Доп. признаки МЭ
187	Петр Игнатьевич, [...] рассказывает длинно, [...] причем говорит не просто Пти, а непременно Жан Жак Пти.	причем < с предикацией > < начальная > < p CNT q >	Ignatiévitch, [...] fait un exposé long, [...] de plus il ne dit jamais Petit tout court, mais infailliblement Jean-Jacques Petit.	de plus < с предикацией > < начальная > < p CNT q >	<input checked="" type="checkbox"/> Exp <input type="checkbox"/> Ext_Up <input type="checkbox"/> Ext_Down <input type="checkbox"/> Up_Down <input type="checkbox"/> NB <input checked="" type="checkbox"/> Cnt

By the end of 2015, the SCDB will have included at least 1,000 connector-ME. This number is planned to be increased to 5,000 during the implementation of the project. These ME will be supplemented by machine translations. The SCDB's interface is going to be in Russian, French, and English; the database will be made available online and users will be able to update it with new linguistic data of the Russian, French, and other languages.

2.3.3. Developing theoretical principles for linguistic description and presentation of simple and multiword connectors in a supracorpora database

Filling a database with data inevitably brings up a question of how multiword connectors should be presented there. In the above-mentioned example of ME (§2.3.2), the utterance with the connector *причем/pričem* [besides] expresses not only a relation of concomitance, but also another type of logical-semantic relations. This is a relation of correction

expressed by the negative particle *не/ne* [not] and the conjunction *а/а* [but/and]. These items set up a relation of correction within a complex sentence. Hence, the following problem arises, that of how to annotate the indicators of this relation: whether we should consider *не ... а* [but/and] to be a multiword connector, which is common in Russian linguistics, or whether we should consider only the conjunction *а* [but/and] to be a connector. (cf. in French, *mais* [but] has two meanings: “adversatif” [adversative] vs. “rectificatif” [corrective], and the latter conveys a relation of correction). It seems very likely that in order to answer the question of how to locate multiword connectors within a language system, one has to:

- Differentiate between the conceptual level of analysis (type of logical-semantic relations) and the linguistic level (discourse representations of this relation, also by means of connectors);
- Adopt a concept of “construction” (Croft 2007, Goldberg 2003) that allows for taking into account the peculiarities of the context in which a linguistic item occurs.

The Russian language has indeed a large number of constructions that consist of a coordinative conjunction *и* [and] and an element of a different morphological origin, i.e., particles, adverbs, modal words (cf. the following connectors: *и притом/и притом, но притом/но притом, да притом/да притом, притом же/притом же* [moreover, while, on the contrary, at the same time]; see Examples (5)-(9) below in §2.3.4-6.). However, it is still unclear what their place is in the Russian language system. This question sharpens at the lemmatization stage, when building annotated corpora and databases. To find out whether these items are multiword connectors or combinations of connectors, developers of the HANCO corpus (see §2.1. above) rely on the data provided by dictionaries of grammatical words. However, such linguistic data may vary from dictionary to dictionary and are rather contradictory (Inkova 2015). This results from the fact that a particular combination is assigned to “the word’s equivalent” (Mustajoki & Kopotev 2004) by “a deliberate decision” of some linguists (this problem is reflected in summary tables of grammatical words in Bogdanov & Ryžova 1997; cf. also how multiword connectors are presented in Métrich & al. 2000). The database of connectors, which is being created, and the unidirectional method of contrastive analysis used within the framework of this research will provide a solution to this problem based on a large amount of linguistic and statistical data.

2.3.4.-6. Statistical processing of data and their linguistic interpretation. Description of language-specific connectors in Russian compared to French

The method of unidirectional analysis has already been successfully applied to the analysis of language-specific words that do not have any direct translation equivalent but many variants of translation or cannot be translated at all. The Russian conjunction *а/а* [but/and] gives an example of the kind. It does not have an exact equivalent in a number of languages. Nor do the connectors with the preposition *при/при* (*притом/притом* [besides], *причем/причем* [while/and], *при этом/при этом* [besides/at the same time], *при том что/при том что* [and with all this]). They do not have any direct French or Italian equivalents but have them in German. Translators offer different interpretations that reflect a semantic pattern of the expressed relation. See examples of *притом* [besides/to that] from the SCDB of connectors:

- (5) ...но всё же несколько оскорбленного в своем достоинстве человека, *и притом* твердо решившегося потребовать объяснений.
 [...nevertheless, his dignity offended *and, at the same time*, firmly determined to obtain explanations.]
 mais, néanmoins, un peu offensé dans sa dignité et, *en même temps*, fermement décidé à exiger des explications.
- (6) Я говорила тогда, что лифчик надо длиннее кроить *и притом* в два полотнища.
 [I have said then that you should have cut bras longer *and, besides that*, used two pieces of cloth.]

J'avais bien dit qu'il fallait couper le corsage plus long et, *en outre*, prendre deux hauteurs.

- (7) он начнет рассказывать так бойко, так живо, что рассмешит и ее, *притом* он такой понятливый!

[He tells stories so glibly, so lively that it makes her laugh. *On the top of that*, he is so quick-witted!] [...] qui narre ses aventures avec tant de fougue et de vivacité que très vite elle se déride. *Et puis* il est si intelligent!

- (8) Я отвык совсем ездить; с непривычки, да еще зимой, признаюсь, мне бы трудно было, не хотелось бы... *Притом* же в деревне одному очень скучно.

[I've completely lost the habit of travel. To tell the truth, with this lack of habit, in winter it would be difficult for me, I would not like to... *Besides* it is so boring to live alone in the countryside.]

Mais j'ai perdu l'habitude de voyager, et de plus j'avoue qu'en hiver... *Et enfin*, vivre à la campagne tout seul, c'est tellement ennuyeux!

- (9) Это, кажется, уже решенный вопрос, что противоположные крайности если не служат поводом к симпатии, как думали прежде, то никак не препятствуют ей. *Притом* их связывало детство и школа – две сильные пружины.

[It seems to be a proven truth that even if opposite extremes are not a sufficient condition for sympathy, as we once thought, they do not preclude it either. *Moreover*, they were bound by memories of their childhood and school – those two powerful ties.]

La réponse est que les affinités sont rares, et qu'elles ne contribuent pas, comme on le croit trop souvent, à créer la sympathie. Les deux hommes étaient unis par les souvenirs de l'école et de l'enfance, ces deux liens puissants.

Thus, in Example (5), *i pritom* is translated as *et, en même temps* [and, at the same time]. This item makes explicit a “simultaneity of two situations”, which pertains to a relation of concomitance. In Examples (6) and (7), *pritom* is translated as *et, en outre* and *et puis* [moreover, besides]. Their main function is to demonstrate both that the statement is added to the previous one and that it contains some additional information as to what has already been said. This is the characteristic function of Russian connectors of enunciative addition. More specifically, it allows us to assume that *pritom* can also express this type of relations. In Example (8), *pritom že* is translated as *et enfin* [and eventually], a discourse structure marker signalling that the item it introduces comes last in the enumeration (this meaning is absent in Russian *pritom*). The choice of this translation is nevertheless justified in the context: there are two arguments against the countryside life and *pritom* introduces the last one in the enumeration. Finally, in Example (9), the translator ignores the Russian connector *pritom*.

In addition, electronic parallel corpora and databases provide statistical data on how different the ways of translating a language-specific item could be, which adds a further dimension to the semantic analysis.

2.4. Work division, schedule and milestones

Year 1 (2016)

1. Draft a preliminary list of Russian connectors that express relations of concomitance and enunciative addition; group the connectors in clusters, based on formal criteria (O. Inkova-Manzotti, A. Zalizniak, I. Levontina).
2. Develop methodical principles for filling the SCDB with data in which Russian connectors will be aligned with their translation equivalents; adapt techniques of contrastive corpus analysis to the description of connector semantics (O. Inkova-Manzotti, A. Zalizniak, I. Zatsman).
3. Test the proposed approach on the Russian conjunction *a* [but/and] and connectors with preposition *pri*; verify and adjust the previously obtained results of connector analysis (A. Zalizniak, O. Inkova-Manzotti).
4. Draft a preliminary list of language-specific Russian connectors that express relations of concomitance and enunciative addition (A. Zalizniak, I. Levontina).
5. Research language-specific Russian connectors of the specified semantic classes (A. Zalizniak, I. Levontina).

6. Develop a second version of the SCDB of connectors; elaborate the methodology and technology of filling the database with Russian-French machine translations of text fragments containing connectors (I. Zatsman, M. KruzHKov).
7. Build 1,000 ME in addition to 1,000 ME that already are in the SCDB by the start of the project (N. Popkova, O. Petrushkina).
8. Add French machine translations of the sentences containing Russian connectors to both 1,000 ME that already are in the SCDB and to 1,000 newly built ME (M. KruzHKov).
9. Compare the outcomes of human and machine translations (N. Buntman, V. Nuriev).
10. Generate statistical data reflecting correlations between human and machine translations (M. KruzHKov, N. Buntman, V. Nuriev).
11. Perform an expert linguistic evaluation of the correlations between human and machine translations (M. KruzHKov, N. Buntman, V. Nuriev).
12. Hold a joint research workshop in Moscow; it will consider correlations between human and machine translations (I. Zatsman, O. Inkova-Manzotti, A. Zalizniak,).
13. Produce a description of connectors with the preposition *pri*: *pri tom*, *pri čem*, *pri etom*, *pri tom što* [as to, besides, on condition that]; based on this description, define relations of concomitance and enunciative addition (O. Inkova-Manzotti).
14. Elaborate criteria for delimiting the functional class of connectors from similar types of linguistic items, based on the experimental data collected from the SCDB (A. Kochetkova-Poquerus).

Year 2 (2017)

1. Finish describing connectors of enunciative addition; begin describing connectors of reformulation (O. Inkova-Manzotti, A. Zalizniak, I. Levontina). Continue to develop and verify criteria that allow us to classify multifunctional items as connectors, based on the experimental data collected from the SCDB (A. Kochetkova-Poquerus).
2. Develop principles for description and structured representation of multiword connectors in the SCDB (O. Inkova-Manzotti, A. Zalizniak, I. Zatsman).
3. Adapt the unidirectional method to the analysis of language-specific connectors, based on French-Russian translations (A. Zalizniak, N. Buntman, V. Nuriev).
4. Build 1,500 ME in addition to 2,000 ones built before the start of year 2 (N. Popkova, O. Petrushkina). Add machine translations of corresponding connectors to 1,500 newly built ME (M. KruzHKov).
5. Compare outcomes of human and machine translations; describe the outcomes in the SCDB; generate statistical data on the degree of correlation between human and machine translations, based on 1,500 new ME (A. Zalizniak, O. Inkova-Manzotti).
6. Describe language-specific Russian connectors, based on 3,500 ME built within two years (A. Zalizniak, I. Levontina).
7. Hold an international conference in Geneva that will bring together scholars working in the field of contrastive description of connectors and discourse words (O. Inkova-Manzotti, A. Zalizniak, I. Zatsman, A. Kochetkova-Poquerus, I. Levontina, V. Nuriev, N. Buntman, M. KruzHKov, N. Popkova, O. Petrushkina).
8. Generate statistical data on multiword Russian connectors, based on 3,500 ME built within two years (M. KruzHKov, N. Buntman, V. Nuriev).
9. Develop a French interface of the SCDB (O. Inkova-Manzotti, A. Zalizniak, M. KruzHKov, N. Buntman).

Year 3 (2018)

1. Build 1,500 ME in addition to 3,500 ones built before the start of project year 3 (N. Popkova, O. Petrushkina).
2. Supply 1,500 newly built ME with machine translations of connectors related to the above semantic classes (M. KruzHKov).
3. Compare human and machine translations, give a statistical analysis of the data and their linguistic interpretation (O. Inkova-Manzotti, A. Zalizniak, N. Buntman, V. Nuriev).
4. Finish describing Russian language-specific connectors that express relations of enunciative addition, reformulation, and concomitance (A. Zalizniak, I. Levontina).
5. Develop an English interface of the SCDB (A. Kochetkova-Poquerus, A. Zalizniak, M. KruzHKov, I. Zatsman).
6. Hold the final research workshop in Moscow; it will bring together scholars working in the field of contrastive description of connectors and discourse words; the project results will be presented at the workshop. (O. Inkova-Manzotti, A. Zalizniak, I. Zatsman).
7. Finish the study; prepare a monograph on Russian connectors (O. Inkova-Manzotti).

List the planned visits between the Swiss and Russian research groups (visiting scientist, hosting scientist, purpose of visit, date and duration of visit).

2016

March, Moscow, 5 days, O. Inkova-Manzotti (visiting scientist), I. Zatsman, (hosting scientist), a joint workshop on clustering connectors and describing their characteristics in the SCDB.

October, Geneva, 5 days, I. Zatsman and M. KruzHKov (visiting scientists), O. Inkova-Manzotti (hosting scientist), a joint workshop on the methodology and technologies of operating the SCDB.

November, Moscow, 5 days, O. Inkova-Manzotti (visiting scientist), I. Zatsman (hosting scientist), a joint workshop on the correlations between human and machine translations.

2017

February, Moscow, 5 days, O. Inkova-Manzotti (visiting scientist), I. Zatsman, (hosting scientist), a joint workshop on using the SCDB in contrastive studies of connectors.

March, Geneva, 5 days. A. Zalizniak (visiting scientist), O. Inkova-Manzotti (hosting scientist), a joint workshop on describing language-specific connectors in the SCDB.

October, Moscow, 5 days, A. Kochetkova Poquerus (visiting scientist), I. Zatsman (hosting scientist), a joint workshop on the results of the first half of the project.

November, Geneva, 5 days, N. Buntman (visiting scientist), O. Inkova-Manzotti (hosting scientist), a joint workshop on the expert evaluation of the ME built for connectors in the SCDB.

2018

February, Moscow, 5 days, O. Inkova-Manzotti (visiting scientist), I. Zatsman (hosting scientist), a joint workshop on using the SCDB in describing a classification of connectors.

October, Geneva, 5 days, M. KruzHKov and V. Nuriev (visiting scientists), O. Inkova-Manzotti (hosting scientist), a joint workshop on the statistical assessment of connector-ME.

November, Moscow, 5 days, O. Inkova-Manzotti (visiting scientist), I. Zatsman (hosting scientist), the final workshop on the project.

2.5 Importance, impact and results

Scientific importance and impact

- We will develop the first classification of logical-semantic relations based on identifying both the underlying semantic operation for each relation and the linguistic level on which it is established. This classification will be universal and applicable to other languages, given that differences between languages manifest themselves either on the level of particular types of a certain relation or in its discourse representations.
- Within the project, it is expected to construct a unique supracorpora database of connectors. It will allow for a contrastive analysis of Russian and French connectors, which will help to reveal new linguistic insights on how the text is organized. The project will significantly contribute to the further development of Russian grammar. Some considerations will be specified, namely the ideas about how connectors function, what semantic features they have, what logical-semantic relations they express.
- For the first time, the project will provide statistical data on correlations between Russian connectors and their French translation equivalents. To that end, we will use a representative sample of 5,000 connector-ME. In order to build these ME, a number of parallel Russian-French texts will be analyzed, which amounts to at least 2,000,000 words.

Expected results and dissemination plan

After the completion of the project, the following results are expected:

- 1) Integral semantic description of Russian connectors expressing relations of enunciative addition, reformulation, and concomitance and more precise definition of these relations;
- 2) Principles for description of multiword connectors and identifying their place within the system of the Russian language;
- 3) Principles for presentation of multiword connectors in annotated corpora and supracorpora databases;
- 4) Criteria for delimiting functions of multifunctional linguistic items (those that may function as a connector);
- 5) SCBD of Russian connectors and their French functional equivalents, available online and containing at least 5,000 ME. Its interface will be in Russian, French, and English. After the completion of the project, it will be possible to fill the SCBD with data in Russian, French, and other languages.
- 6) Statistical data on relevant operating parameters of Russian connectors, translation equivalents of these connectors, on types of cross-language equivalents in different translations, as well as statistical data on types of regular errors in machine translation;
- 7) Description of Russian connectors expressing relations of enunciative addition, reformulation, and concomitance in terms of their linguistic specificity.

The results of the present research will be made public in:

- Articles in Russian and international high-ranking journals;
- Reports at Russian and international conferences, including the conference in Geneva planned as part of the project;
- Database of connectors, available online after the completion of the project;
- Monograph on multiword Russian connectors.

Broader impact

The results of this joint research would be used in socially important fields such as:

- *Vocational and higher education sector*: the results would have a genuine impact on teaching general linguistics, theory of translation, the Russian language, contrastive grammar of Russian and French; in developing higher education programs on philology, language teaching, translation and interpretation in Switzerland, Russia, and other countries;
- *Lexicography*: the project results would help in creating monolingual and bilingual dictionaries;
- *Linguistic information resources*: after the completion of the project, the database of connectors will be available on the Internet, the users will be able to fill it with data;
- *Machine translation*: the expected results would have a meaningful impact on improving and enhancing the quality of machine translation, when processing linguistic items that organize the text;
- *Professional translation*: this study would provide rich information sources for literary, technical and scientific professional translation, namely a bilingual database and a monograph on multiword Russian connectors and possibilities of their translation into French.

In a wider context, the expected results could apply to the economy, as a better quality of machine translation would facilitate the global flow of goods and services via the Internet.

3. References

- Aijmer K., Simon-Vandenberg A.-M. (2006), *Pragmatic Markers in Contrast* (Studies in Pragmatics), Amsterdam, Elsevier.
- Apresjan V. (2006), « Ustupitel'nost' v jazyke », in Apresjan Ju. (ed.), *Jazykovaja kartina mira i sistemnaja leksikografija*; Moskva, Jazyki slavjanskoj kul'tury, pp. 615-712.
- Baranov A. & al. (1993), *Putevoditel' po diskursivnym slovam russkogo jazyka*, Moskva, Pomovskij i partnery.
- Blakmore D. (2002), *Relevance and Linguistics Meaning: The Semantics and Pragmatics of Discourse Markers*, Cambridge, Cambridge University Press.
- Bogdanov S. & Ryžova Ju. (1997), *Russkaja služebnaja leksika. Svodnye tablicy*, Sankt-Peterburg, Izdatel'stvo Sankt-Peterburgskogo universiteta.
- Boguslavskaja O. & Levontina I. (2002), « Podvedenije itogov v russkom jazyke », *Logičeskij analiz jazyka. Semantika načala i konca*, Moskva, Indrik.
- Boguslavskaja O. & Levontina I. (2004), « Smysly 'pričina' i 'tsel' v estestvennom jazyke », *Voprosy jazykoznanija* 2.
- Breindl E. & al. (2014), *Handbuch der deutschen Konnektoren 2. Semantik der Deutschen Satzverknüpfen*, Berlin, New York, Walter de Gruyter.
- Buntman & al. (2014), « Informacionnye tehnologii korpusnyh issledovanij : principy postroenija krosslingvističeskikh baz dannyh, *Informatika i ee primenenija* 8, 2, pp. 98-110.
- Čeremisina M. & Kolosova T. (1987), *Očerki po teorii složnogo predloženiya*, Novosibirsk, Nauka.
- Charolles M. (1995), « Cohésion, cohérence et pertinence du discours », *Travaux de linguistiques* 29, pp. 125-151.
- Croft W. (2007), « Construction Grammar », in D. Geeraerts & H. Cuyckens (eds), *Handbook of Cognitive Linguistics*, Oxford, Oxford University Press, pp. 463-508.
- Čžon H.H. (2003), *Prisoedinitel'nye skrepy v sovremennom russkom jazyke : sintaksis i semantika*, Moskva, Diss. kand. nauk, MGU.
- De Cesare A.-M. & Borreguero Zuloaga M. (2014), « The contribution of the Basel model to the description of polyfunctional discourse markers », in S. Pons Bordería (ed.), *Discourse Segmentation in Romance Languages*, Amsterdam, John Benjamins, pp. 54-94.
- Dobrovol'skij D. & Levontina I. (2012), « Russian NET vs. German NEIN 'NO': a semiotic approach », *Russian Linguistics* 36, pp. 213-219.
- Fischer K. (éd.) (2006), *Approaches to Discourse Particles* (Studies in Pragmatics, vol. 1), Amsterdam, Elsevier.
- Fraser B. (2006), « A theory of discourse markers », in Fischer K. (ed.), *Approaches to Discourse Particles* (Studies in Pragmatics, vol. 1), Amsterdam, Elsevier, pp. 189-204.
- Fraser B. (2009), « An Account of Discourse Markers », *International Review of Pragmatics* 1, pp. 1-28.
- Fraser B. (2013), « Combinations of Contrastive Discourse Markers in English », *International Review of Pragmatics* 5, pp. 318-340.
- Goldberg A. (2003), « Constructions: a new theoretical approach to language », *Trends in Cognitive Sciences* 7 (5), pp. 219-224.

- Halliday M.A.K. ([1985] 1994), *An Introduction to Functional Grammar*, London/Melbourne, Edward Arnold.
- Inkova O. (2008), «Problemy perevoda sootnositel'nyh predloženij : rusko-francuzskie sootvetstvija», in *Russkij jazyk i kul'tura v zerkale perevoda : materialy meždunarodnoj naučno-praktičeskoj konfencii, Saloniki 14-18 maja 2008 goda*, Moskva, MGU, pp. 208-227.
- Inkova O. (2009a), « Scalarité endophrastique vs scalarité exophrastique : quels types d'échelle ? », *Orientalia Parthenopea IX*, pp. 43-72.
- Inkova O. (2009b), « De l'identité à l'intensité : la sémantique des locutions russes v takoj stepeni vs do takoj stepeni », *Studi Linguistici e Filologici Online 7.1.*, pp. 19-44, <http://www.humnet.unipi.it/slifo>.
- Inkova O. (2010a), « Sootnositel'nye predloženija russkogo i francuzskogo jazykov : shodstva i različija », in A. Bonola & O. Inkova (eds), *Categorie verbali e problemi dell'organizzazione testuale = L'Analisi linguistica e letteraria I*, 2010, pp. 141-160.
- Inkova O. (2010b), « L'expressivité par l'anaphore. Le cas du russe », in L. Gautier & P. Monneret (eds), *La fonction expressive*, vol. 2, *Recherches en linguistique*, Besançon, Presses Universitaires de Franche-Comté, pp. 69-79.
- Inkova O. (2010c), « De l'anaphore à la scalarité. De quelques expressions anaphoriques en russe », in P. Hadermann & O. Inkova (éds), *Approches de la scalarité*, Genève, Droz, pp. 67-112.
- Inkova O. (2011a), « Les connecteurs anaphoriques du russe : entre subordination et corrélation », *Revue française de linguistique appliquée XVI-2*, pp. 25-40.
- Inkova O. (2011b), « Kol' et tol' et leurs dérivés en russe d'aujourd'hui : une traduction laborieuse », *Cahiers du CIRHiLL (Centre interdisciplinaire de recherches en histoire, lettres et langues)*, Paris, L'Harmattan, pp. 255-286.
- Inkova O. (2013a), « Predloženija proporcional'nogo sravnenija v russkom jazyke : nekotorye utočnenija », *Russian Linguistics 37* (2/2013), pp. 103-124.
- Inkova O. (2013b), « "A v ostal'nom, prekrasnaja markiza" : o semantike v ostal'nom, v pročem i vpročem », *Russkij jazyk v naučnom osveščeenii 1* (25), pp. 21-40.
- Inkova O. (2013c), « O semantike tak nazываемых gradacionnyh sojuzov ne stol'ko... skol'ko i skoree...čem », *Voprosy jazykoznanija 1*, pp. 38-52.
- Inkova O. (2014a), *La corrélation en russe : structures et interprétations*, Bern, Peter Lang.
- Inkova O. (2014b), « Come i kak: opyt sopostavitel'nogo issledovanija sravnitel'nyh konstrukcij v italianskom i v russkom jazykah », in O. Inkova & al. (eds), *Architettura del testo*, Alessandria, Edizioni dell'Orso, pp. 91-122.
- Inkova O. (2015), « K voprosu o lemmatizacii mnogokomponentnyh edinic », *Korpusnaja lingvistika-2015* (in print).
- Inkova O. (in press1), « Du syntagme prépositionnel au connecteur : la préposition russe pri et la conjonction pri tom čto », in A. Bertin & H. Bat-Zeev Shyldkrot, *Prépositions, conjonctions, connecteurs*, Paris, H. Champion.
- Inkova O. (in press2), « O granicah funkcional'nogo klassa konnektorov », in E. Quero Gervilla & al. (eds), *Trends in Slavic Studies*, Moskva, Editorial URSS.
- Inkova O. & Manzotti E. (in press), « Sur la notion de généralisation », in P. Hadermann & D. Van Raemdonck, *La complexité grammaticale*, Bern, Peter Lang.
- Inkova-Manzotti O. (2001), *Konnektory protivopostavlenija vo francuzskom i russkom jazykah. Sopostavitel'noe issledovanie*, Moskva, MGU-Informèlektro.
- Inkova-Manzotti O. (2002), « Les connecteurs accommodants : le cas de autrement », *Cahiers de linguistique française 24*, pp. 109-141.
- Inkova-Manzotti O. (2005), « Encore sur la conjonction russe a to », *Revue des études slaves*, tome LXXVI, fascicule 4, pp. 485-496.
- Kempgen S. et al. eds (2009), *Die slavischen Sprachen /The Slavic Languages. Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung /An International Handbook of their Structure, their History and Their Investigation*, Band 1 / Volume 1, Berlin / New York, Walter de Gruyter.
- Kibrik A. & Podlesskaja V. eds (2000), *Rasskazy o snovidenijah. Korpusnoe issledovanie ustnogo russkogo diskursa*, Moskva, Jazyki slavjanskih kul'tur.
- Knott A. & Sanders T. (1998), "The classification of coherence relations and their linguistic markers: An exploration of two languages", *Journal of Pragmatics 30* (2), pp. 135-175.
- Kortmann B. (1997), *Adverbial Subordination. A Typology and History of Adverbial Subordinators Based on European Languages*, Berlin – New York, Mouton de Gruyter.
- Kruzhkov M. & Inkova O. (2015), "Parallel'nye francuzsko-russkie basy dannyh glagol'nyh form i konnektorov", IVe Congrès international de linguistique textuelle contrastives langues slaves – langues romanes *Langues slaves en contraste*, 14-16 May 2015, University of Bergamo, Italy.
- Kruzhkov M. & al. (2014), "A database of Russian verbal forms and their French translation equivalents", *Komp'juternaja lingvistika i intelektual'nye tehnologii: Trudy meždunarodnoj konferencii «Dialog'2014» [Computational Linguistics and Intellectual Technologies. Proceedings of International Conference «Dialog'2014»]*, Vol. 13(20), pp. 284-297.
- Les particules énonciatives... 1986-1987 – Les particules énonciatives en russe contemporain*, Paris, Institut d'études slaves.
- Levontina I. (2011), « Časticy peresprosa i pripominanija », in *Slovo i jazyk. Sbornik statej k 80-letiju akademika Ju. Apresjana*, Moskva, Jazyki slavjanskih kul'tur, pp. 269-278.

- Levontina I. (2013), «O pričinom značenii sojuza *a to*», in *Komp'juternaja lingvistika i intelektual'nye tehnologii. Trudy Meždunarodnoj konferencii Dialog'2013*, Moskva, RGGU, pp. 434-446.
- Levontina I. (2014), « Diskursivnye slova v voprositel'nyh predložženijah », *Die Welt der Slaven* LIX, 2, pp. 201-218.
- Ljapon M. (1986), *Smyslovaja struktura složnogo predložženija i tekst. K tipologii vnutritektovnyh omošenij*, Moskva, Nauka.
- Mann W.C. & Thompson S.A. (1988), "Rhetorical Structure Theory: Toward a functional theory of text organization", *Text* 8, pp. 39-78.
- Métrich R. & al. (2002), *Les invariables difficiles en 4 volumes: Tome 1, Aber-Ausserdem; Tome 2, Bald-Geradezu; Tome 3, Ern-Nur; Tome 4, Obendrein-Zwar*, Paris – Nancy, Association des Nouveaux cahiers d'Allemand.
- Mikaelian I & Zalizniak A. (2005), « The Russian conjunction *a* revisited », in *Vostok – Zapad : Vtoraja meždunarodnaja konferencija po modeli « Smysl ⇔ Tekst »*, Moskva, pp. 304-315.
- Montchaud P. (2015), « Reciprocal Connection in French », in A. Bondaruk & al. eds, *Young Minds vs Old Questions in Linguistics: Proceedings of the Fourth Central European Conference for Postgraduate Students*, Lublin, The Institute of East-Central Europe and the John Paul II Catholic University of Lublin, pp. 139-155.
- Mustajoki A. & Kopotev M. (2004), « K voprosu o statusu èkvivalentov slova tipa *potomu čto*, v *savisimosti ot, k sožaleniju* », *Voprosy jazykoznanija* 3, pp. 88-107.
- Nikolaeva T. (1985), *Funkcii častic v vyskazyvanii. Na materiale slavjanskih jazykov*, Moskva, Nauka.
- Nikolaeva T. ed. (2004), *Verbal'naja i neverbal'naja opory prostranstva mežfrazovyh svjazej*, Moskva_ Jazyki slavjanskoj kul'tury.
- Oates S. (2001), *Multiple Discourse Markers Occurrence: Creating Hierarchies for Natural Language Generation*. Unpublished MA Thesis, University of Brighton.
- Paillard D. & al. eds (1998), *Diskursivnye slova. Opyt kontekstno-semantičeskogo opisanija*, Moskva, Metatekst.
- Paillard D. & Kiseleva K. eds (2003), *Diskursivnye slova russkogo jazyka : kontekstnoe var'irovanie i semantičeskoe edinstvo*, Moskva, Azbukovnik.
- Pasch R. et al. (2003), *Handbuch der deutschen Konnektoren. Linguistische Grundlagen der Beschreibung und syntaktische Merkmale der deutschen Satzverknüpfen (Konjunktionen, Satzadverbien und Partikeln)*, Berlin, New York, Walter de Gruyter.
- Prijatkina A. (1977), « Ob otličii sojuza ot drugih svjazujuščih slov », *Russkij jazyk v škole* 4, pp. 102-106.
- Rossari C. (2002), « Les adverbies connecteurs : vers une identification de la classe et des sous-classes », *Cahiers de linguistique française* 24, pp. 11-43.
- Roulet E., L. Filletaz & A. Grobet (2001), *Un modèle et un instrument d'analyse de l'organisation du discours*, Bern, Berlin, etc. *Russkaja grammatika* / pod red. N. Švedovoj, Moskva, Nauka, 1980.
- Sannikov V (1989), *Russkie sočiniteľ'nye konstrukcii: Semantika, Pragmatika, Sintaksis*, Moskva, Nauka.
- Warner R. G. (1979), *Discourse connectives in English*, Thesis PhD, Ohio State University.
- Zalizniak A. (1992), *Issledovanija po semantike predikatov vnutrennego sostojanija*, München, Otto Sagner Verlag, Slavistische Beiträge, B. 298.
- Zalizniak A. (2006), *Mnogoznačnosť v jazyke i sposoby ee predstavlenija*, Moskva, Jazyki slavjanskih kul'tur.
- Zalizniak A. (2013), *Russkaja semantika v tipologičeskoj perspektive*, Moskva, Jazyki slavjanskih kul'tur.
- Zalizniak A. (2015), *Lingvospesificnye edinicy russkogo jazyka v svete kontrstivnogo korpusnogo analiza*, in *Komp'juternaja lingvistika i intelektual'nye tehnologii. Trudy Meždunarodnoj konferencii Dialog'2015*, in press.
- Zalizniak A. & al. (2015), « Nadkorpusnye bazy dannyh kak lingvističeskij resurs », *Korpusnaja lingvistika-2015* (in print).
- Zalizniak A., Levontina I. & Šmelev A. (2005), *Ključevye idei russkoj jazykovoj kartiny mira*, Moskva, Jazyki slavjanskih kul'tur.
- Zalizniak A., Levontina I. & Šmelev A. (2012), *Konstanty i peremennye russkoj jazykovoj kartiny mira*, Moskva, Jazyki slavjanskih kul'tur.
- Zalizniak A. & Mikaelian I. (2005), «Russkij sojuz *a* kak lingvispecifičnoe slovo », in *Komp'juternaja lingvistika i intelektual'nye tehnologii. Trudy Meždunarodnoj konferencii Dialog'2005*, Moskva, pp. 153-159.
- Zatsman I. & al. (2014), "Conceptual framework for development of computer technology supporting cross-linguistic knowledge discovery", in *Proc. 15th European Conference on Knowledge Management* (Santarem, 2014), Academic Publishing International, Reading, pp. 1063-1071.