RUSSIAN IN CONTRAST: FORM, MEANING AND PARALLEL CORPORA

ATLE GRØNN AND IRENA MARIJANOVIC

Department of Literature, Area Studies and European Languages
University of Oslo

[1] RUSSIAN IN CONTRAST – WHO ARE WE?

As an initial taste of what is to come, we offer our readers this introduction as a kind of smorgasbord where we present most of the contributions to Russian in Contrast (Oslo, September 2009), namely, those papers which made their way to these proceedings. We also take this opportunity to discuss more general challenges currently facing Slavic linguistics and share some of our views on the role of parallel corpora in contrastive linguistics.

Let’s start with the conclusion: This is a valuable and highly stimulating collection of articles, thirty one in total. The proceedings comprise three parts, and in the spirit of the conference, with parallel publications in English and Russian. These introductory remarks relate to the first two parts which contain the articles in English.

The papers from Russian in Contrast reflect the diversity, strength, and, perhaps, shortcomings of our field. Indeed, we have succeeded in gathering some of the best minds in the field, but we hear ourselves exclaim almost immediately what field? The field we are referring to is quite big, but apparently not considered big enough – so far – to have its own channel for dissemination of research. With this at the forefront of our minds, it became evident that the goal of our September conference in Oslo was to create such a channel or meeting place.

We are slavicists and/or linguists with interest in the semantics of Russian. This, we argue, is a legitimate occupation – the readers can judge for themselves by taking a closer look at the papers – but our impression is that this union is better described as rather an uncomfortable marriage. Looking at it from a historical perspective, that is, either in modern times or in the past, and irrespective of the geographical location, whether here in Norway or in Russia, the matrimony between the linguist and Russian semantics has not always proved to be a particularly good match. Of course, exceptions are always to be found, and luckily, several of these rarities have contributed to the present volumes.

For obvious reasons, some of us are Slavicists whereas others are linguists. But irrespective of our affiliation, all of us are subject to and bound by the same
rigorous standards imposed by research: if we as Slavicists want to say something about meaning, we must be as good as the best semanticists;\(^1\) if we as linguists want to say something about Russian, we must be as good as the best Slavicists (and native speakers).\(^2\)

[1.1] **Contrastive semantics**

The title *Russian in Contrast* was purposely curtailed – what is it that Russian stands in contrast to? In fact, “the contrasts” and “tensions” are numerous, as we will see below. Plainly, we are analysing Russian in contrast to other languages, but our conference was also an attempt to gather divergent approaches to Russian semantics.

Let’s start with the theory. If asked, most of us would probably agree that what we are doing is contrastive semantics. However, can we conceive of contrastive semantics as a theoretical approach in any meaningful fashion? Unless subscribing to a crude form of, what could be dubbed as, linguistic essentialism whereby the existence of impersonal constructions, say in Russian, is inextricably linked to some inherent national characteristic, such as “submissiveness”, semantics is surely universal? So, what we probably mean by contrastive semantics then is the study of contrastive forms (syntax, morpho-syntax, constructions).\(^3\)

But even if by “contrastive semantics” we actually mean “contrastive forms”, semantics is obviously our main preoccupation. For instance, we want to find out which phenomena, encoded in the syntax or morphology, pertain to semantics in one language and to “pragmatics” in another. (A simple example would be the explicit aspectual encoding in Russian vs. the lack of aspectual morphology in Norwegian – although both languages can with equal success describe events in the world).

One of the main ideas of *Russian in Contrast* is thus to use data from different languages to illuminate and give us insight into corresponding phenomena in Russian. A fitting example, in the spirit of the Norwegian-Russian RuN project that organised the conference, is Barbara Partee’s paper, where she uses Scandinavian (Danish) data in order to understand differences between copular sentences in Russian and English. We hope, however, that all of the papers presented here contribute to this common research agenda.

Finally, although the label “contrastive semantics” is somewhat fuzzy, the way it is informally understood makes it compatible with various theoretical approaches: with both cognitive grammar (which is popular in Slavistics) and for-

---

\(^1\) This is what makes it easier for, say, a Norwegian interested in Russian language to call herself a Slavicist rather than a linguist.

\(^2\) This is what makes it easier for, say, a Norwegian interested in Russian language to call herself a linguist rather than a Slavicist.

\(^3\) Indeed, in the lexical domain the realia may also differ from language community to language community – see the papers of Rachilina and Ivaniščeva in Part 2 and Melent’eva in Part 3.
Malsemantics (which is not really “popular” anywhere, but is occasionally appreciated for its beauty). Hence, contrastive semantics is a field where all linguists and their pet theories can in fact meet.

[1.2] Semantics vs. syntax

As linguists, we cannot do without form, of course. The relationship between form and meaning – syntax and semantics – should be at the heart of our preoccupations. That is why one would expect at least to come across some interesting claims about syntax in this volume, but such theorising is almost nowhere to be found.

That no thought-provoking results on syntax proper came to the fore during these proceedings may impel many a Slavicist to shout joyfully: So much the better! However, unlike most of our colleagues in Russian semantics for whom Chomskyan syntax is an obscenity of the most vulgar sort, we are somewhat saddened by such state of affairs, all the more since Center for Advanced Study in Theoretical Linguistics (CASTL) in Tromsø, in the North of Norway, has played an instrumental role in introducing Chomsky to Russians, mainly by opening its doors to Russian PhD students. All the same, we did not succeed in attracting generativists to our conference.  

Whatever views one has on syntax, we cannot do serious semantics without appealing to form (syntax); and it is rather unclear what kind of syntax-semantics interface we do assume in “our field”. In this respect, there is still much work to be done, for many generations to come. In our opinion, the whole field is in desperate need of some ten to twenty different advanced textbooks on the syntax of Russian, so that we can continue doing our semantics, occasionally casting downward glances from our ivory tower at the syntax laid out in a given textbook.

These problems are not restricted to Slavic linguistics, of course. Romance linguistics to pick one, or Norwegian linguistics to pick another, are also characterised by a certain immaturity of attitude in that there very often seems to be an almost complete absence of consensus regarding a minimum number of commonly shared theoretical assumptions.

Indeed, modern theoretical linguistics is a very callow science that started with Chomsky and Montague (Partee 2009). As often in science, these two giants didn’t see eye to eye on most things, and consequently had little liking for each other’s work. In Slavic linguistics, Montague is probably the lesser evil – and has not become a commonplace profanity like Chomsky – though that may in part be due to the fact that Montague is simply relatively obscure there.

[4] Perhaps we didn’t really try, or perhaps this is as it ought to be. Generativists working on Russian have had their own venues for a long time: regular conferences such as Formal Description of Slavic Languages (FDSL) in Europe and Formal Approaches to Slavic Linguistics (FASL), and the journal Journal of Slavic Linguistics (JSL) in the United States. Although semanticists are occasionally accepted into the bosom of the above-mentioned generative family, we never really feel completely at home in these surroundings.
To get a more balanced view on these matters, we can only recommend Barbara Partee’s special pre-conference talk (Partee 2009) on her two influential teachers, available in the video format on the RuN home page. We fully share Partee’s position where she reconciles the Chomskyan and Montagovian approaches (without subscribing to all details of their agendas).

At another level, a micro-level, from a representative of the new generation, we were inspired by some passages in Natalja Sljusar’s recent monograph Na styke teorij ... (Sljusar’ 2009). Sljusar’ comes from a long tradition of Russian functionalism in St. Petersburg but wrote her PhD thesis in generative syntax in Utrecht. She published an extended version of her thesis as a monograph in Russian in which she observes: “С полученными в рамках генеративной теории выводами можно, и как нам кажется, зачастую нужно спорить. Но их уже нельзя игнорировать.” (Sljusar’ 2009, 130). In our view, a similar statement could be made with regard to formal semantics: Should we, however, keep on ignoring the sheer beauty and elegance of generalised quantifier theory? It was interesting to note the reaction from Tat’jana Černigovskaja to this publication⁵: “... Эта книга — примета нового времени, свежая и сильная, один из первых успешных синтезов отечественных лингвистических традиций и достижений генеративного подхода. Она показывает, как должен быть подготовлен современный ученый, чтобы уверенно себя чувствовать в лингвистике XXI века”.

Our intention here is not to make much ado about this particular book or even about generative grammar (a field which we, regrettably, cannot say with any confidence to know too well), but this example illustrates some of the challenges in our “field”. Sljusar’’s open-minded and eclectic approach, combining insights from both Russian and Western linguistics, is unquestionably something that ought to be encouraged in the hope that more researchers will be drawn to employing such an integrative approach in their own work. So, if Partee can reconcile Montague with Chomsky and Sljusar’ a functional approach with a generative one, it may not be too fanciful to believe that (theoretical) (Slavic) linguistics can, some day, abandon its sectarian tendencies and emerge as a unified clan after all.

When all is said and done, what is imperative and of central importance is, of course, the notion of quality. Any work in linguistics, regardless of its theoretical orientation, should be informed by a rigorous application of scientific method, draw our attention to hitherto unnoticed and unobserved connections between specific facts of a given language, or indeed languages, and use sound and valid argumentation with respect to the phenomena it essays to explicate. For that

[⁵] Černigovskaja is, of course, a major figure in Russian linguistics, oriented towards cognitive grammar but with an extremely broad research agenda, and not the least, she is also the co-author of a contribution to this volume with Elena Tkačenko, now affiliated to Oslo University College, but a former student of Černigovskaja.
reason, our field ought to adhere to the following principle: any set of results, whether obtained in the West or in the East, should only be rejected because new research provides a deeper and more comprehensive understanding of the matter at hand, and not, as it seems to be today’s practice, on the basis of ignorance or ideological prejudice.

[1.3] Contrastive speakers

Given the diversity of the field, our goal was to attract the best Russian semanticists if not from all possible worlds, then at least from our world. To help us in this endeavour, a handful of guest speakers were invited in order to crystallise what our objective is even further. It was for this reason that an invitation simply had to be extended to both Barbara Partee and Elena Viktorovna Padučeva as they both share all the desired qualities: both linguists stand above the crowd and continue to enjoy a wide international readership. Padučeva is famous for her informal but extremely lucid observations which easily lend themselves to further formalisation (if the reader desires so).

Partee is evidently more formal, as the mother of Western style formal semantics, but due to the close connections she has had with Russia for many decades, she is well versed in both Western and Eastern, that is, formal and non-formal approaches to semantics. Partee is one of the few linguists in the field who can relate to everyone with an astounding ease. (This is probably a sign that “the field” is not big enough to have its own yearly conference after all, otherwise there would have been more people like Partee around.) In other words, Partee is unequivocally one of the leading glitterati in the linguistic community at large and in Slavic Linguistics for sure.

There is a general impression that the cognitive community has the upper hand in Slavistics, and we were therefore delighted to invite Laura Janda and Tore Nesset from Tromsø. This is an extraordinarily forceful duo in both science and life, and Janda’s arrival to Tromsø has been a real boost for the community of Slavic linguistics in Norway.

It was also a great pleasure to invite Ekaterina Rachilina, a most natural choice given our joint interest in corpus linguistics, a linguist who is also known for her rather enthusiastic lecturing style as well as for making a significant contribution in the field of lexical semantics.

All the invited guest speakers are known to give excellent lectures; their enthusiasm, sincerity and the sheer passion for their chosen subjects made this conference into an unforgettable event. A warm, welcoming and constructive atmosphere was punctuated by tongue-in-cheek yet friendly comments, like Rachilina’s opening remark to her own presentation right after von Stechow’s λ-talk on sequence of tense: “... and now to something interesting”. On behalf of von Stechow and his co-author, we can assure the reader that we were most amused.
Among other distinguished speakers were notably two German professors emeriti, one linguist and one Slavicist, Armin von Stechow and Hans Robert Mehlig. Von Stechow started popularising Montagovian formal semantics (and to some extent also Chomskyan syntax) in Europe some forty years ago, and has been one of the most influential semanticists ever since (outside Slavic linguistics).

Mehlig has been one of the leading aspectologists for decades inside Slavic linguistics. It was clear that Padučeva and Mehlig spoke in a mutually intelligible tongue (both contributions are in Russian – more on Russian in the preface to Part 3), and the same may be observed of Partee and von Stechow. Evidently, however, there is still some way to go and a few mountain faces to scale. It is unfortunate for our field that the nature of things is such that understanding is not a transitive relation — if you understand x who understands z, then you don’t necessarily understand z.

From the younger generations, we were happy to see the contributions from our not so green Russian neophytes who assimilated into Western style formal linguistics, namely, Daniel Altshuler, Katja Jasinskaja and Svetlana Krasikova. They probably do not consider themselves Slavicists and they all did their PhDs in linguistics in the West.

Another strong research community hails from Tromsø featuring several insightful talks from the Russians working in Janda and Nesset’s research group, namely, from Julia Kuznecova, Olga Ljaševskaja and Svetlana Sokolova.

One more community which contributed to the present volume with several papers is our partner in the RuN project from Murmansk whose work is mostly focused on Scandinavian languages, notably Norwegian (see the introduction to Part 3).

[2] RUSSIAN IN PARALLEL AND CONTRAST TO OTHER LANGUAGES

So, our field is not without its challenges, although the reader may question the way in which this state of affairs was presented above. However, everyone can agree on the necessity of providing good and reliable data.6 Our conference invited a certain focus on data: going beyond Russian itself, new data was presented from parallel corpora, juxtaposing Russian with English, Scandinavian languages and Romance.

[6] This does not mean that everybody actually uses good data. Traditionally, there has been an extremely unfortunate demarcation between data-oriented and theoretically oriented linguistics. It suffices to remind the reader of the helpless toy sentences used for decades in Chomskyan and Montagovian linguistics. Again Partee stands out – famous for her original and good examples – she always has an example up her sleeve for any puzzle at the syntax-semantics interface, as von Stechow pointed out in his introduction to Partee’s final talk closing the conference.
Corpus linguistics is an ideal meeting point for both formal and cognitive linguists. Indeed, the cognitive camp seems to hold sway in this area.

Janda and Tromsø have taken cognitive grammar a step further by implementing statistical approaches and thereby taking a quantitative turn in Slavic linguistics. In this respect, it will be exciting to follow their new project *Neat theories, messy realities: How to apply absolute definitions to gradient phenomena* (Janda 2009).

We have several examples in these volumes using monolingual corpora, mostly the Russian National Corpus. Several authors compare their findings for Russian with the data from other monolingual corpora for other languages.

Given the right statistical tools, this approach can give rise to very interesting results and it further has the advantage of having access to large data sets for each language (that is, monolingual corpora such as the Russian National Corpus, Google, Yandex, and so forth).

One important aspect of the conference was to elicit papers based on the use of parallel corpora. Perhaps we only partly succeeded in this respect. Indeed, so far— including these proceedings— sophisticated quantitative and statistical analyses have mostly been done with monolingual corpora. Three reasons come to mind why this is so: first, there are not many parallel corpora available; second, the existing parallel corpora are too small for our research tasks; and third, parallel corpora, generally speaking, are not suited for the kind of research questions linguists are wont to pursue. (While acquiescing to the validity of the first two points, we raise strong objections about the validity of the third.)

Although the use of parallel corpora raises some methodological issues (Johansson 2007), the advantages it offers are numerous: it banishes the tedium more often than not engendered by more traditional methods, something which makes doing research more fun, as well as proving to be more entertaining for the reader.

Parallel corpora places at our disposal the technology which gives us ready access to authentic language from high-quality authors, as well as to highly accomplished, and often outstanding, translations of the texts in question. This is especially the case when the study of language is approached from the perspective of second language learning (and in contrastive studies, most researchers are confronted with second language learning unless a particular linguist is one of the lucky bilinguals).

Although languages in general, and translations in parallel corpora in particular, differ in the most unexpected ways, we should be neither discouraged nor daunted by this fact. First, if the parallel corpus is big enough, then the most devi-
ating and “noisy” translations will disappear as statistically insignificant. Second, the complexity of data from parallel corpora will force us to face and pose really hard questions since many theoretical puzzles are bound to arise as systematic patterns are observed in individual languages as well as cross-linguistically.

This methodology – the use of parallel corpora – will give us both cross-linguistic and language internal frequency distributions. The data obtained in this manner will incontestably open up for a whole range of investigations involving statistical models and a number of statistical methods relevant for linguistic applications are waiting to be used. As noted by (Janda 2009) and others, a corpus-based analysis of authentic language invariably yields observations that are difficult to reconcile with theories of language. We can expect the “clash” between traditional theoretical linguistics and statistical methods to produce some extremely interesting theorising.

In contrastive linguistics today, the limited access to good parallel corpora means that the use of parallel corpora must typically be supplemented with other traditional sources (monolingual corpora, Google, and so on) and tools (such as informants or introspection). However, we would like to advocate the principle that studies in contrastive linguistics should start with a survey of the data available in existing parallel corpora. If the data is scarce, this may be an interesting finding in itself.

Ideally, by compiling sufficiently large parallel corpora, we could create a new set of standards for contrastive linguistics. Several phenomena – also treated at this conference – occur so frequently that we can already, at this stage, empirically test some of the claims made in the literature.

As a closing remark on the issue of the use of corpora in linguistics, we would like to draw our readers’ attention to a point frequently made by one of the founding fathers of the Russian National Corpus, namely, Vladimir Plungjan. He remarks that it is rather puzzling why linguists are so keen on constructing ungrammatical toy sentences when corpora is bursting with so many facts yet to be discovered, especially when linguists are not in the position to boast of having a full understanding of grammatical sentences either.

[2.3] The RuN Corpus

At the University of Oslo – notably at our department of European languages – we have a strong tradition going back to the late Stig Johansson in building parallel corpora, that is, searchable collections of texts (both the originals and their translations) aligned at the sentence level. Such is, for instance, the OMC – the Oslo

[7] This is a novel perspective in theoretical syntax and semantics/pragmatics, but we are convinced that the availability of different kinds of reliable frequency data will lead to a range of new research questions. If construction x occurs 10 times more frequently than construction y, we – as linguists – should be able to explain the difference in distribution.
Multilingual Corpora – which involves Norwegian, English, German and French texts. In the RuN project, we started developing the RuN corpus comprising Norwegian and Russian originals and their respective translations.

We are currently expanding the RuN corpus to include more texts and languages, notably English, Bulgarian, Serbian, Polish, Italian, French, Swedish and German. The new corpus is labelled the RuN-Euro Corpus.

The RuN-Euro parallel corpus and the OMC share a common web search interface using the corpus explorer tool, Glossa, developed by the Text Laboratory at the University of Oslo.

Glossa is a graphical interface built on top of the IMS Corpus Workbench query system. The Glossa system allows post-processing of the search results. Collocations can be viewed and counted in a number of ways, thus facilitating statistical analysis of the data. The corpora are also linked to grammatical tags word-by-word. Tagging has been done using and adapting existing tagger models (Nygaard et al. 2008).

The RuN-corpus is of course not the only parallel corpus for Russian. As part of the Russian National Corpus, there exists a big Russian-English/English-Russian parallel corpus, and a smaller Russian-German/German-Russian corpus.

Other Russian bilingual parallel corpora include the Russian-Slovak Parallel Corpus and the Russian-Finnish Parallel Corpus. The latter is part of Michail Michajlov’s dissertation (Mihailov 2003), which also contains a survey of parallel corpora in general and parallel corpora with Russian as one of the languages in particular.

In spirit, the RuN-Euro Corpus is closer to the Parasol project. Both corpora are truly multilingual. While the main languages in the RuN-Euro Corpus remain Russian, Norwegian and English, the focus of Parasol is primarily on the whole family of Slavic languages in addition to German.

The RuN-Euro Corpus, as well as the OMC, is heavily biased towards modern fiction texts and thus is not a representative corpus. Though the priority is given to contemporary prose texts, there is no harm in including older texts or other genres in the corpus since the search interface (Glossa) allows the search to be restricted according to parameters such as date of publication, author and genre.

The main limitation of existing parallel corpora is certainly their modest size compared to monolingual corpora, some of which contain hundreds of millions of words. The reason for this discrepancy is obvious: it is much harder, in all possible respects, to compile a parallel corpus than a monolingual corpus. Accordingly, researchers in contrastive linguistics will always feel that their parallel (multilingual) corpus is too small. Our answer is that this limitation should not be a reason to give up – we must continue the work, building larger and larger parallel (multilingual) corpora!
Some applications at Russian in Contrast

Let’s confront some real data. The first two illustrations come from the research carried out as part of the RuN project relative to complement tense and converbs. Both phenomena occur frequently in real texts and display a highly interesting cross-linguistic variation. A corpus-based approach will thus immediately establish which constructions actually occur (and at which frequency) in natural languages.

In contrastive, cross-linguistic studies of theoretical syntax, semantics and pragmatics, this use of parallel corpora represents a significant innovation.

Complement tense

Russian is known to be a so-called non-sequence-of-tense language, that is, a language without the kind of tense agreement between the matrix tense and subordinate tense as is found in English or Norwegian.

Apart from the canonical and well-studied verba dicendi, such as “say”, “ask”, and so forth, the literature on embedded complement tense raises several questions. What happens in complements of past tense perception verbs, like “see” and “hear”, and factives, like “know” and “understand”? Do we have a relative present tense in the Russian complement or an independent past? In fact, some of these questions were raised by the curious and insightful members of the audience at our conference Russian in Contrast. A puzzling situation arose since although some of the brightest minds in Russian semantics were present at the conference and were actively engaged in the debate around this issue, the latter proved to be rather slippery and seemingly resistant to any attempt at clarification.

A good start is to look at what the parallel corpora have to offer. Indeed, both the present and the past tenses are attested not only in these environments but also, surprisingly enough, in the same sentence:

(1) R Она глядела, как сизые кольца от сигары Азазелло уплывали в кamin и как кот ловит их на конец шпаги.
    (Michail Bulgakov, “Master i Margarita”)

E She watched as Azarello blew (Rus: past!) smoke-rings at the fireplace and the cat spiked (Rus: present!) them on the end of his sword.

N Hun så på Azazello, som sendte ringer av sigarrøk inn i peisen, og på katten, som fanget dem opp med kordespissen.

The discussion at the conference led us to a more systematic examination of this issue in Grønn and von Stechow (this volume). Interestingly, the authentic data abound, for instance, there are many examples of factive matrix verbs and embedded past tense in Russian with a simultaneous interpretation:
Of course, the corpus data is just the beginning. Very little has been said so far, and, we can assure the reader, nothing further will be said in these introductory remarks. If you are really interested – you could look up the analysis at your own peril by scrolling down to the article itself.

**Converbs in Russian/English/Norwegian**

In her contribution to this volume, Maria Filiouchkina Krave looks at systematic differences in translation of **converb** constructions (gerunds, *deepričastija*) in Russian. As “expected”, the translations of perfective converbs of, say, the delimitative Aktionsart explicitly encode the temporal precedence relation with PPs headed by *after*:

(4)  
R  — Но скажите, пожалуйста, я никогда не могла понять, — сказала Анна, — помолчав несколько времени [...]
(Lev Tolstoj, “Anna Karenina”)  
E  “But do tell me, please, I never could make it out,” said Anna, *after being silent* for some time [...]
N  “Men si meg en gang, er De snill, jeg har aldri kunnet forstå,” sa Anna *etter å ha sittet taud* en stund [...]

However, the translation data from the RuN corpus reveal certain peculiar properties of perfective converbs of the semelfactive kind:

(5)  
R  Сергей сидел в углу, *закинув* ногу на ногу, и курил [...]
(Viktor Pelevin, “Generation P”)  
E  Sergei sat in the corner *with his legs crossed*, smoking [...]
N  Sergej satt *med korslagte bein* i et hjørne og røykte [...]

OSLa volume 2(1), 2010
The use of the ‘with/med + DP’ construction in English/Norwegian suggests that the converb event is interpreted – at least by the translators – as simultaneous with the matrix event, which is unexpected given the standard semantics of temporal precedence associated with perfective converbs in Russian.

Krave argues in her paper that the reason for this interpretation is that certain semelfactive verbs lexically denote the “target state” of the event – in this case the state of the legs’ being crossed. It is precisely this state, and not the preceding event/activity of crossing the legs itself, which is referred to by the construction chosen by the translators. Hence, the complex situation is interpreted as simultaneity although the “semelfactive event” in a narrow sense may still precede the matrix.

The lesson to be drawn from examples like (5) is that translations in corpora may reveal facts which have been largely ignored in the literature based on grammaticality judgments alone. Systematic patterns arise which challenge the conventional wisdom.8

We are convinced that authentic cross-linguistic data will serve as an impetus to rethink a whole range of theoretical issues. Although our main focus here will be to show that parallel corpora can be a unique source and an indispensable tool in the study of grammatical categories across languages as well as in the lexicon, it should be emphasized that parallel corpora may also be of considerable interest and value to researchers and practitioners in other language-related disciplines such as computational linguistics, translation studies, and, last but not the least, second-language teaching.9

3 The Contents

As mentioned earlier, the papers from the proceedings are split into three parts, that is, into three separate books: Grammar (1); Lexicon (2); and Русский язык в контрастивном аспекте (3).

This division, as well as the titles of individual books, is indeed somewhat arbitrary. The second part includes some articles which could just as well have been added to Part 1.

Part 3 is truly “Russian” in that all papers are written in Russian, starting with an engaging paper by Elena Padučeva. The article, supported by the evidence afforded by the corpus data, endeavours to answer the question formulated some thirty–forty years ago by Vendler and Kartunnen: What is the presuppositional

8 We remember our old Russian teacher, the late professor Geir Kjetsaa (1937-2008), also a brilliant translator, who told us on more than one occasion to automatically use the preposition “etter” (after) when translating a perfective converb. Well, now in hindsight, we see that things are slightly more complicated.

9 More examples from parallel corpora will be given in the next section. See also the introduction to Part 3, notably Klonova on impersonal sentences in Russian and Norwegian, and Janygina on implicit information in the Norwegian novel “Naiv. Super” and its Russian translation (from the RuN corpus).
status of indirect questions embedded under factive verbs? This kind of fruitful, although admittedly slow, dialogue between Russian and Western linguistics is what lies at the heart of our conference: what we envisage our conference to be about and what we hope to see more of in future works in the field of Russian linguistics.

Below we focus on Parts 1 and 2, that is, on the papers written in English.\(^\text{10}\)

[\[3.1\] Grammar

As already mentioned, (formal) semantics (and syntax) and Slavistics have traditionally not been a very good match. For some – certainly a minority – of us, “grammar” more or less means \textit{compositional semantics}.\(^\text{11}\) Thus, from our perspective, the main obstacle to a long, fruitful and harmonious marriage concerns \textit{compositionality}. It is very much our hope that this contrast, at a theoretical level, will be overcome by future generations.\(^\text{12}\)

It should, however, be admitted that it is hard to provide a fully compositional explanation for many linguistic phenomena – a fact that has a surprisingly positive aspect to it: there is work enough for everyone.

\textbf{Partee}

The beginning of our conference was ushered in by Barbara Partee’s pre-conference talk – whose video may be seen on the \textit{Russian in Contrast} web page (\textcite{partee2009}) – and brought to a close with her paper on \textit{Specificational copular sentences in Russian and English}. The latter paper kicks off the first part of the proceedings, a natural choice in view of the previous paragraph since Partee has been active from Amherst to Moscow in popularising the Fregean principle of compositionality since before we were born!

In much of her work, Partee addresses phenomena which pose a considerable challenge to a straightforward compositional analysis. This is also the case on this occasion: what is presented is a contrastive analysis of copular sentences in Russian and English, a topic, it seems, of such elusive and slippery quality that even individual linguists find it hard to agree with themselves\(^\text{13}\) on apparently “simple” questions such as what constitutes the subject of a sentence. The theo-

\[\text{\[10\] See the introduction to Part 3 for further discussion of the Russian papers.}

\[\text{\[11\] A personal note: compositionality is the only thing we truly believe in relative to the study of form and meaning; the rest is darkness. If even a single piece of indubitable evidence could show that compositionality is an illusion, then we'll be the first ones in line to seek early retirement and endeavour to find comfort in other gods, naturally outside linguistics.}

\[\text{\[12\] We noted with satisfaction that Nesset’s opening talk of \textit{Russian in Contrast} placed some emphasis on the fact that cognitive grammar is not incommensurable with the concept of compositionality. Nesset even included a photo of Frege in his powerpoint presentation, something which cannot be described otherwise than an excellent start to our conference.}

\[\text{\[13\] Partee traces the history of the debate in her article and shows how researchers, including herself, are forced to shift positions on this highly intricate topic.}]}
retical complexity of these simple matters is such that we dare not look at what the parallel corpus may spout at us.

The younger generation

Next, we turn to the younger generation. Sveta Krasikova, a former student of Arnim von Stechow in Tübingen, contributes with one of the most complex papers in these proceedings. Behind the complexity, we find phenomena which should interest any Slavicist. As the topic of the paper is gradable adjectives, let’s take a look at one example from the parallel corpus:

(6) E And then, just as the beasts had burst out of the earth, there burst out from the shoulders of Fledge wings that spread and grew, larger than eagles’, larger than swans’, larger than angels’ wings in church windows. (Staples Lewis, “The Chronicles of Narnia. The Magician’s Nephew” – Russian National Corpus)

R А потом — точь-в-точь как звери, появлявшиеся из земли — на спине у Стрелы прорезались крылья, которые росли и расправлялись, стали большие орлиных, шире лебединых, громаднее, чем крылья ангелов на церковных витражах.

The reader may find it obvious why the translator chose the morphologically marked comparison forms šire, gromadnee instead of bolee širokie/gromadnye, and so forth. On reading Krasikova however, the reader will soon realise that nothing is trivial in the domain of gradable adjectives – and that the reader’s favourite theory on this topic can hardly be called a theory.

Daniel Altshuler presents his recent work – parts of his PhD from 2010 – on the evergreen issue of Russian aspect, the general-factual interpretation of the imperfective. Instead of illustrating the phenomenon with Kto čital “Vojnu i mir”? or a chrestomatijnyj, but constructed, example from Rassudova, we take the liberty to present an example from the RuN-Euro Corpus. The author is all the same chrestomatijnyj.

(7) R Варенька сказала, что Анна Павловна присылала сказать, что вы не поедете. (Lev Tolstoj, “Anna Karenina”)

E Varenka said that Anna Pavlovna sent word you were not going.

N Varenka sa at Anna Pavlovna hadde send bud at det ikke ble noe av turen.

The data in (7) display several interesting tense-aspect combinations, notably related to sequence-of-tense phenomena. In tribute to Altshuler, we note that the relative past in (7R) is expressed with an imperfective past, prisylala (“sent”), in-
stead of the perfective aspect. This should be an *obščefaktičeskoe značenie* as good as any.\textsuperscript{14}

**TAM in contrast**

Next follow three more papers on tense, aspect and mood, namely the papers by Grønn & von Stechow, Filiouchkina Krave (see above) and Dobrušina. Nina Dobrušina compares some interesting data from Russian and French relative clauses: the alternation between the subjunctive and the indicative. Her data are mostly from the monolingual Russian National Corpus, but she notes certain interesting parallels with the French data, which were elicited from informants:

(8) \textbf{R} Надо разработать такую социальную систему, которая максимально защищала бы [OK: защитит] детей из малообеспеченных семей ‘We should create a social infrastructure that would protect children who come from low income families.’

\textbf{F} Il faut mettre au point un système qui protège (SUBJ) [OK: protégera (FUT)] les enfants de familles pauvres.

It is rather obvious that it would be extremely interesting to back up these patterns with authentic data from parallel corpora.\textsuperscript{15}

**In search of nothing**

Lilli Parrott’s paper is a comparison of the Russian direct address form usage – the vocative – with that in other languages, especially English, Czech and Polish. Since our corpora are not tagged for the vocative case, the search is therefore not straightforward. Nevertheless, it was possible to find some curiosities, for instance, the English translation in (9) – Mash! – possibly a mistake on the translator’s part engendered in all likelihood by his ignorance of the new Ø-vocative form with Russian -a nouns:

(9) \textbf{R} Ой, *Маш*, ты, что ли? (Ljudmila Ulitskaja, “Medeja i ee deti”)

\textbf{E} Oy, *Mash*, is that you?

\textbf{N} Oj, *Masja*, er det deg, da?

\textsuperscript{14} The example is not chosen accidentally: Altshuler’s previous work on sequence-of-tense has inspired our own work; and our own work on the general-factual imperfective plays a certain role in Altshuler’s paper in this volume. That said, no consensus has been reached on either of these matters, as far as we can see.

\textsuperscript{15} We are currently adding Romance texts (French and Italian) to the RuN-Euro Corpus, but there is a long way to go before a really good Slavic-Romance corpus becomes available.
If you are interested in ellipsis, searchable corpora may not at first sight appear to be the right tool to use. However, inspired by Jekaterina Mažara, we managed in the end to retrieve some data\textsuperscript{16}, such as the following example:

(10) R — Ну, как тебе ∅ мой грим, Валер?
     (Ljudmila Ulitskaja, “Medeja i ee deti”)
E Well, what do you \textit{think} of my makeup, Valerii?
N Nå, hva \textit{synes} du om sminken min, da, Valerij?

\[3.2\] \textbf{The Lexicon}

Albeit not so unambiguously clear-cut, the focus in Part 1 is on traditional grammatical categories and their interpretation at the sentence level, whereas Part 2 represents more innovative areas of research, as the lexicon assumes the focal point. In the lexical domain, depending on the research topic, parallel corpora, if sufficiently large, can provide new insights.

\textit{Morphology in contrast}

We start Part 2, that is, \textit{OSLa}, volume 2(2), 2010, with an original study by Laura Janda, in which she discusses the role of metonymy in suffixal word formation in three of her favourite languages: Russian, Czech and Norwegian.

In the Glossa search interface, if the suffix is known, it can be searched for by a \textit{truncated} search. An example would be the Russian suffix “-*nica”. One kind of metonymy associated with this suffix is \textit{contained for container} and standard examples include words such as \textit{pepel’nica} and \textit{sacharnica}, where the names for containers are derived from nouns that refer to a substance that is \textit{contained} in them.

An example of such metonymy is given in (11):

(11) R Посреди комнаты — тяжелый, как гробница, стол, накрытый белой скатертью [...] (Michail Bulgakov, “Sobač’e serdce”)
N Midt i værelset et bord, tungt som et mausoleum og dekket med en hvit duk [...] 
E In the middle of the room was a table, heavy as a gravestone and covered with a white tablecloth [...] 

The Norwegian translator correctly comprehends the metonymy involved as exemplified by his choice of the word “mausoleum” (contained for container). The English translator, on the other hand, either has not understood the Russian origi-

\begin{footnotesize}
\textsuperscript{16} Looking for data in a parallel corpus is, in game theoretical terms, a fight between two often rather non-cooperative players: \textit{Y}(ou) and \textit{C(orpus)}. \textit{Y} makes a move (search). \textit{C} answers by finding a way to escape \textit{Y}’s intensions. \textit{Y} realises his naive mistake in the first round and makes a new move making it harder for \textit{C} to escape \textit{Y}’s intensions. In the end, either \textit{Y} or \textit{C} wins by default.
\end{footnotesize}
nal or he takes the liberty of choosing another metaphor ("heavy as a gravestone", gravestone = \textit{nadgrobný kamen}).

We know that metaphor (theory) is complicated, and so is metonymy. As demonstrated by Janda, the presence of a suffix like \textit{*nica} does not in and of itself determine the metonymic designation. The next example fittingly illustrates Janda’s point as the Russian suffix in (12), \textit{-nica}, this time signals the metonymy action for location:

(12) R Окурки на пол не бросать — в сотый раз прошу. Чтобы я более не слышал ни одного ругательного слова в квартире! Не плевать! Вот плевательница. (Michail Bulgakov, “Sobač’e serdce”)
N Ikke spytte. Der har De spyttebakken.
E And don’t spit everywhere! The spittoon’s over there.

The interesting word \textit{plevatel’nica} is obviously derived from the activity verb \textit{plevat’} according to the productive model given by Janda. Neither of these metonymic designations are attested for Norwegian, which has overall a more restricted system of affixal word-formation. In Norwegian, we typically find compounds like “spyttebakke” [spit + ground, that is, spittoon].

This brings us to Janda’s husband and colleague, Tore Nesset, whose project is close in spirit to hers, but narrower in focus. Nesset presents a contrastive case study of word formation in Russian and Norwegian in the domain of negative characterizations of people, such as in the pair below:

(13) R — А под вашею полною достоинства личиною, — отнесся артист к Дунчилю, — скрывается жадный паук и поразительный охмурало и врун. (Michail Bulgakov, “Master i Margarita”)
N Og bak Deres maske av lutter verdighet, sa artisten henvendt til Dunchill, skjuler det seg en grådig ederkopp, en helt fantastisk svindler og løghals.
E “And under your so very dignified mask,” the artiste adverted to Dunchil, “is concealed a greedy spider and an astonishing bamboozler and liar.”

Following Nesset we note, as we did with Janda’s examples above, that Norwegian frequently uses a compound noun [“løghals” = lie + throat].

Our corpus can, of course, be used to search for such components with a truncated search, e.g. \textit{*fant} [= tramp, vagabond in English]:

(14) N Greven har jevnlig figurert i avisenes sladrespalter. Han er en Don Juan, en \textit{fyllefant} og en bråkmaker.
(Nikolaj Frobenius, “Latours katalog”)
The Count had featured regularly in the gossip columns; he was a Don Juan, a drunkard, a troublemaker.

Den Klatschspalten zufolge ist er ein Don Juan, ein Säufer und Randalierer.

Son nom défraie les ragots des gazettes. Don Juan, sac à vin, faiseur de scandales.

Other Russian translations of the Norwegian “fyllefant” in the RuN-corpus include: p’janica and p’jančužka. Nesset shows convincingly that common gender nouns in -a in Russian, such as in (14) above, systematically correspond to a compound noun in Norwegian.

Nesset himself gives only one example from the RuN Corpus thus tacitly pointing to its limitations. (Since Nesset finished his study, the corpus has been significantly expanded, but still, of course, his initial point is duly taken: the corpus should be at least ten times its present size. To rectify this rather wretched state of affairs, all that’s needed is lots of ready dosh, which, once in our coffers, will allow us to do the rest. A Maecenas, anonymous or otherwise, would be most welcome.)

The last morphology paper by Tkačenko and Černigovskaj features the same languages as above, namely, Russian and Norwegian, but from a rather different perspective. Their paper is an interesting comparison of the acquisition of verbal morphology by Norwegian adult learners of L2 Russian with the children acquiring Russian as L1. For once, we will not try to squeeze any further examples out of our parallel corpora.

Lexicon in contrast

It is Ekaterina Rachilina’s earnest plea for the often ignored and dismissed part of natural language, namely, the lexicon, and for lexical studies in general, that is the first paper to appear in this section (translated from Russian to English by Irena Marijanović). It is envisaged as a first tentative step towards a future, fully comprehensive study of verbs of sound. In this paper, Rachilina focuses on verbs denoting animal sounds, in particular where the latter are metaphorically used in relation to various reactions and actions performed by human beings.

A parallel corpus would certainly prove to be an indispensable tool in a study of this nature since it ultimately aims to develop a topology of verbs of sound. The corpus would, of course, have to be significantly larger than what can presently be offered by the RuN-Euro Corpus. However, even in its present, relatively circumscribed edition, some data may be obtained. The following examples illustrate
the extension of the Russian verbs fyrkat’, ryčat’ and šipet’, and their corresponding verbs in Norwegian and English, from the animal into the human domain.

(15)  
R  Пристает! — продолжал Ляпунов, сердито фыркая.  
E  “He won’t leave off,” Lyapunov went on, snorting angrily.  
N  — Han gir seg ikke! fortsatte Ljapunov og fnyste forarget.

(16)  
R  [О]н страшно конфузился и, когда уходили гости, хватал ее за руку и шипел сердито: [...] (Anton Čechov, “Dušečka”)  
E  He was dreadfully embarrassed, and when the guests had gone, he would seize her by the hand and hiss angrily: [...]  
N  [H]an ble fryktelig brydd, og når gjestene var gått, kunne han gripe henne i armen og frese sint: [...]  

(17)  
R  Он не хотел уходить с окна, рычал, и вздрагивал, и порывался спрыгнуть вниз. (Michail Bulgakov, “Master i Margarita”)  
E  Here refused to leave the window, growled and twitched, and kept trying to jump out.  
N  Den ville ikke ned fra vinduet, men knurret og rykket og ville hoppe utfor.

Next, we turn to Olga Ivaniščeva, Murmansk, a former student of the late Valerij Berkov (1929-2010), an expert on Nordic languages, in particular Norwegian, and a world-class lexicographer. Not surprisingly, Ivaniščeva presents a paper concerned with lexicographical issues from a Scando-Russian bilingual perspective. This is precisely the kind of data where our parallel corpus would have been extremely useful had it been merely a hundred times bigger. Such a corpus would be capable of revolutionising bilingual lexicography, including the work of Berkov (for whom the availability of such a Russian-Norwegian corpus tools was, perhaps, beyond imagination.) With the RuN Corpus, we can at least imagine what a sufficiently large corpus would be like and the myriad of ways in which it could be applied.

All the same, our modest corpus may prove useful already at this stage. Let’s take the Russian concept of njanja.

The Norwegian translator of Anna Karenina, Erik Egeberg (professor emeritus, University of Tromsø), also uses the more old-fashioned word dadda:

(18)  
R  Несчастный ребенок! — сказала няня, шикая на ребенка.  
(Lev Tolstoj, “Anna Karenina”)  
N  “Stakkars unge!” sa barnepiken og hysjet på barnet.  
E  “Luckless child!” said the nurse, hushing the baby.
In other contexts, *njanja* has a rather different meaning (still probably “nurse” in English), corresponding to the Norwegian *pleierske*:

(20) R Его вынесли няньки и отдали подлецу, я сунула няне трешку. (Ljudmila Petruševskaja, “Vremja noč’”)  
N Pleierskene kom bærende på ham og rakte ham til Bedrageren; som seg hør og bør ga jeg *pleiersken* tre rubler.

In modern prose, *njanja* in its prototypical meaning can also be translated as *dagmamma*:

(21) R Знаешь какой-нибудь нормальной девочке — я ищу няню для Сони. (Andrej Kurkov, “Piknik na l’du”)  
N “Hør her,” Viktor så spørrende på vennen, “du kjenner ikke en grei jente, jeg leter etter en *dagmamma* til Sonja.”

*Dagmamma* is always the choice of the Norwegian translator of the modern writer Andrej Kurkov, whereas the same word is never attested in the translation of “Anna Karenina”. This in itself is significant: it draws our attention to the fact that two words, although seemingly synonymous, may nevertheless reflect different writing styles or in fact different concepts. The case in point is the opposition between the words “barnepike” : “dagmamma”, rendered in Russian by the same word *njanja*, where the difference between the two is dependent to a large extent on the historical context each novel provides.

**Semantic profiling**

The next in line is one of the young researchers from the Tromsø group, Julia Kuznecova. Her paper is informed by a combined lexico-grammatical approach in which the notion of semantic profiling plays a key role.

Semantic profiling may be described as an advanced statistical use of corpora to detect the linguistic environments of certain constructions. In Kuznecova’s investigation, this method is used to examine the occurrence of the PP *on the phone* in Russian and English monolingual corpora. In studies of this kind, it is still very much an open question what purpose parallel corpora would be best used for, apart from generating innumerable examples of the construction(s) under investigation.

(22) R По телефону с ней говорить теперь невозможно. (Ljudmila Ulitskaya, “Medea i ee deti”)

---

OSLa volume 2(1), 2010
E  It’s impossible to speak to her on the telephone now.
N  Det er umulig å snakke med henne i telefonen.

Being at loss to say what one could do with such examples, we have no other choice but to resort to an old linguistic chestnut: such examples will have to be left for “future research”.

Another project from the same group is presented by Svetlana Sokolova and her co-author Wojciech Lewandowski who analyse the constructional profile of the verbal prefix za- in Russian and Polish.

Their data are extracted from the Russian and Polish National Corpora. Here, it seems rather obvious that good parallel corpora would give us valuable additional information. A search in a sufficiently large parallel corpus – if such were to exist - could shed light on the authors’ hypotheses concerning differences between za-verbs in the two languages. Our own example points to a similar use of the imperfective verbs zalewać and zalivat’:

(23)  P  Długi ten wąż w październiku, kiedy Nil zalewa cały Egipt, miałby błękitną barwę wody. (Bolesław Prus, “Faraon”)
R  В октябре, когда воды Нила заливают весь Египет, эта длинная змея принимает голубую окраску воды...

In their study, however, Sokolova and Lewandowski looked at the perfective prefixed verbs zalać and zalit’. What their analysis shows is that these verbs prefer the so-called Goal-Object construction in both languages (zalit’ “Egypt with water”), but in Russian, unlike Polish, can we also have the Theme-Object construction (zalit’ vodu v bak). So, an ideal parallel corpus study is thus predicted to show that in cases where zalać is used in the Goal-Object construction it can easily be translated by means of zalit’. However, if the Russian zalit’ is used in the Theme-Object construction it cannot straightforwardly be translated into Polish with the help of zalać.

Pragmatics in contrast
Elizaveta Khachaturyan from the University of Oslo contributes with a paper on discourse markers, in particular, the juxtaposition of tak skazat’ in Russian with diciamo in Italian. It is quite obvious that a parallel corpus would be extremely useful in such a study. It is unfortunate, however, that the Italian part of our RuN-Euro Corpus is not yet available for use, but the following example gives us an impression of what kind of data can be obtained:

(24)  R  В слове “БЛО” чувствовались неиссякаемые запасы жизненной си- 
        ли и одновременно что-то негуманоидное, поэтому Эдик берег

[17] The example below is excerpted from the Polish-Russian part of the RuN-Euro Corpus.
Part 2 of the proceedings is brought to a close by a pair of papers which could just as well have been placed in Part 1 (grammar), namely the contributions by Margje Post and Katja Jasinskaja. It is always exciting to have twin presentations at a conference: although they are both looking at the Russian conjunction a, the analyses are presented from different contrastive perspectives. Post compares the three Russian basic additive and contrastive coordinating conjunctions i, a and no with their two Norwegian counterparts og and men when used in utterance-initial position:

\[(25) \quad \text{야-то думал, что это сказки. (Michail Bulgakov, “Rokovye jajca”)}\]

\[(26) \quad \text{Это какие-то черти, а не люди. (Michail Bulgakov, “Rokovye jajca”)}\]

As Posts puts it: “An analysis in terms of core meanings needs to be supplied by contrastive studies on the basis of corpora, which show actual use of the words in almost all possible contexts.”

The title of Jasinskaja’s paper is music to our ears: “Corrective contrast in Russian, in contrast”. Her study could not be better suited for the approach to Slavic linguistics we are fervently advocating here. And naturally, we expect such and similar examples to abound in the RuN Corpus. In addition, it is possible to search for both ne a:

or the same construction with a random word in between, that is, ne X a:
It is always a great pleasure to end on a positive note and that means from the point of view of parallel corpora. The topics explored by Post and Jasinskaja are eminently suitable for a parallel corpus study: with a minimum amount of effort, thousands of hits are available at our disposal within seconds. Further ‘polishing’ of the initial crude results through clever use of statistical tools is bound to produce interesting sets of data. It is therefore imperative to continue the work on improving the search-interface which allows statistical processing. At present, it is only statistical processing of the source texts that is made available but with time we also aim to provide statistical processing of target texts, word-by-word-alignment, and so on.

For the discussion of and a short introduction to the rest of the papers from the conference, we refer the reader to the introduction in Part 3.

ACKNOWLEDGMENTS

This concludes the introduction to the first two parts of the proceedings, edited by Atle Grønn and Irena Marijanović. The third part in Russian is edited by Atle Grønn and Olga Klonova.

We would like to thank the participants at the Oslo conference, the authors and the reviewers. The technical part of the work with these proceedings deserves a special mention. As the readers may have noticed, the articles are type set in \LaTeX. This, we’re afraid, may prompt some of you to wonder why we chose to engage in an activity so time consuming and, possibly, not worth the effort we put in it. We defend our actions by appealing to the fact that if Grigorij Perel’man can do it, so can we! Since Mr. Perel’man is known for his preference to publish his work online in this particular document markup language and preparation system, we are hereby inviting him to feel free to submit his papers to our open access series OSLa. The conversion from Word to \LaTeX was this time done by Vladyslav Dorochin, Atle Grønn and Pavel Iosad.

We are proud to present these books to the linguistic community at large for, in our opinion, they are certainly the closest to the very best we are capable of in the field of Russian in contrast. Our conference Russian in Contrast was organized by the RuN project (2008-2010) at the University of Oslo with funding from the Norwegian Centre for International Cooperation in Higher Education (SIU).
REFERENCES


Partee, B. 2009. Perspectives on Semantics: How philosophy and syntax have shaped the development of formal semantics, and vice versa. Invited lecture, Oslo.


AUTHOR CONTACT INFORMATION

Atle Grønn
Department of Literature, Area Studies and European Languages
University of Oslo
Norway
atle.gronn@ilos.uio.no

Irena Marijanović
Department of Literature, Area Studies and European Languages
University of Oslo
Norway
irena.marijanovic@ilos.uio.no