Russian Deverbal Nouns: Lexical Denotation, Argument Structure & Translation Mismatches

Glòria de Valdivia Pujol

Russian Deverbal Nouns: Lexical Denotation, Argument Structure & Translation Mismatches

Glòria de Valdivia Pujol

Dissertation presented within the doctoral program

Cognitive Science and Language

Departament de Lingüística

Universitat de Barcelona

Under the supervision of

Dr. Joan Castellví Vives
Universitat de Barcelona

Dra. Mariona Taulé Delor
Universitat de Barcelona

Universitat de Barcelona

May 2014

Show me the place

Leonard Cohen

No direction home
like a complete unknown
like a rolling stone

Bob Dylan

And Romeo wanted Juliet
And Juliet wanted Romeo
And Romeo wanted Juliet
And Juliet wanted Romeo

Lou Reed

Per als meus pares

Abstract

This thesis presents a descriptive empirical study of deverbal nominalizations. Very often what can be expressed by means of a deverbal noun can also be expressed by means of a verbal construction and vice-versa. Deverbal nouns are hybrid categories that have a mixture of verbal and nominal features. These nouns can denote either the action named by the base verb, being, in this case, a paraphrase of a verbal construction, or the result of that action. If nouns denoting processes are closer to verbs, nouns denoting a result, that is, a concrete or an abstract entity resulting from the action, are closer to nouns. Both result and event nouns inherit the argument structure of the base verb. These analyses of the lexical denotation of deverbal nouns and their argument structure are two of the main aims of this thesis. The third goal is the descriptive and comparative study of translation mismatches between Russian and Spanish deverbal noun constructions. The thesis is structured in the following way in order to cover these three points.

In the first chapter, we present a brief introduction of the whole thesis, we highlight the main goals and motivations for carrying out this threefold study. In the second chapter, we describe the linguistic resources used in the development of this project. We first introduce the monolingual and bilingual corpora, and then, the lexicons. In the third chapter, we discuss the relationship between the morphological and lexical aspects of the base verb and the lexical denotation of the deverbal noun. We also study other factors that may influence the lexical denotation of the deverbal noun. In the fourth chapter, we present the study of the argument structure of deverbal nouns focusing on the type of constituent that can be an argument, how the arguments are realized and their possible combinations, which ends in the obtention of the more frequent syntactico-semantic patterns. In chapter five, we present the descriptive and comparative study of translation mismatches of deverbal nouns between Russian and Spanish. We provide our classification of the translation mismatches depending on the interrelation and number of linguistic changes (morphologic, syntactic, semantic and pragmatic) involved. Finally, in the last chapter we present our conclusions and ideas for further research.

Resum

Aquesta tesi presenta un estudi descriptiu empíric sobre les nominalitzacions deverbals del rus. Molt sovint, allò que es pot expressar a través d'un nom deverbal es pot expressar també a través d'una construcció verbal. Els noms deverbals són categories mixtes que barregen trets verbals amb trets nominals. Aquests noms poden expressar l'acció denotada pel verb base, en aquest cas es poden considerar paràfrasis d'una construcció verbal, o bé el resultat de l'acció. Així doncs, els noms deverbals que denoten l'acció del verb estan més a prop del verb, mentre que els noms deverbals que denoten el resultat de l'acció, és a dir, una entitat concreta o abstracta, estan més a prop del nom. Partim de la hipòtesi que ambdós tipus de nom hereten l'estructura argumental del verb base. Aquestes dues anàlisis, és a dir, l'estudi de la denotació lèxica del nom deverbal i de l'estructura argumental dels noms deverbals són dos dels tres objectius principals de la tesi. El tercer objectiu és l'estudi descriptiu i comparatiu dels desajustos de traducció entre les construccions amb noms deverbals en rus i castellà. La tesi s'estructura de la manera següent:

En el primer capítol, presentem una breu introducció a la tesi on describim els objectius i les motivacions principals d'aquest estudi. En el segon capítol, es descriuen els recursos lingüístics que hem utilitzat en el transcurs del treball. Primer, introduïm els corpora monolingües i bilingües i, finalment, els lexicons. En el tercer capítol, analitzem la relació entre l'aspecte morfològic i lèxic del verb base. També estudiem altres factors que poden influir en la denotació lèxica del nom deverbal. En el quart capítol, presentem l'estudi de l'estructura argumental dels noms deverbals centrant-nos en el tipus de constituent que pot ser argumental, com es realitzen els arguments i les seves possibles combinacions. D'aquesta anàlisi, en resulta l'obtenció dels patrons sintàctico-semàntics dels noms deverbals més freqüents. En el cinquè capítol, presentem l'estudi descriptiu i comparatiu dels desajustos de traducció entre el rus i el castellà. En aquest capítol proposem una classificació de desajustos de traducció en funció de la interrelació i el nombre de canvis lingüístics (morfològics, sintàctics, semàntics i pragmàtics). Finalment, en el darrer capítol presentem les nostres conclusions i idees per a una futura recerca.

Agraïments

I would like to do the aknowledgements in the language that I use to speak with each person that has helped me to build this thesis.

Moltes gràcies, Mariona, per ajudar-me a tirar endavant aquesta tesi. Sense tu m'hagués estat impossible acabar-la. Les hores que hem dedicat a fer i desfer camins traçats plegades no es poden calcular… en són tantes d'hores dedicades. Sé del cert que bona part de la força que m'ha permès seguir aquest camí tan dur me l'ha donat la teva tendresa, comprensió i motivació. Sempre mires endavant i aquesta qualitat m'ha ajudat a veure-m'hi entre tanta nominalització sovint encegadora.

Moltes gràcies, Joan, per ser el llum que sempre s'ha encès quan ha calgut donar un gir o un toc al camí. Gràcies per tants moments divertits al teu costat. Sempre recordaré el suport que em vas donar prenent una cervesa en aquella plaça d'Aix-en-Provence. Amb el teu realisme i humilitat m'has ensenyat que cal estar satisfet i prendre's amb humor la feina feta si s'ha dut a terme amb compromís i sinceritat.

Moltes gràcies, Toni! La teva força i entrega cap a la feina fan que sempre estiguis en moviment, sempre atenta a tot i a tothom que pugui necessitar la teva ajuda. Gràcies per la teva generositat!

Moltes gràcies, Montse. La teva presència, sentir-te teclejar o clicar el mousse m'han ajudat a no perdre el fil de la tesi. Les converses entre hores han fet de l'STeL un lloc molt agradable, on et sents a gust, malgrat el soroll de fons dels servidors. Sempre ha sapigut que si tenia algun problema (de tesi o personal) podia comptar amb tu.

Moltes gràcies, Santi. En els darrers temps quasi no ens hem vist. Tanmateix, tinc un record molt agradable de les converses que he mantingut amb tu. La teva manera pausada i equil·librada d'entendre la vida i la feina m'han ajudat molt en aquest camí.

Moltes gràcies, Martins! La teva alegria sempre m'ha carregat de forces quan les piles s'havien de carregar. Ens uneixen experiències, patiments i angoixes, però, sobretot, una caixa plena de somriures i xerrades.

Moltes gràcies, Aina! Recordo la primera vegada que em vas somriure. Crec que a partir d'aquell moment es va obrir una porta sincera que sempre ens connectarà.

Moltes gràcies, John! Contigo he compartido también muchos dolores de cabeza pero también el interés por el mundo gráfico. Recuerdo nuestras conversaciones como puntos suspensivos entre tanta tesis.

Moltes gràcies, Marta (la SúperReca)! Admiraré sempre el teu cap clar i el teu tracte sincer i amable, sempre disposat a donar un cop de mà.

Moltes gràcies, Oriol! Sempre et consideraré com un home capaç d'afrontar qualsevol problema amb una visió que descomposa el problema en petites partícules i que el recomposa donant el gir necessari a les coses perquè funcionin.

Rita, gràcies per ser tan real, tan sincera, tan amable, tan acollidora. Sempre recordaré els nostres cafès rescatadors al pati de la uni.

Moltes gràcies, Anna! Recordo arribar a l'STeL i trobar-te asseguda treballant, aixecar la mirada i amb un somriure donar-me els bons dies!

Moltes gràcies, Patri! Tu em vas ensenyar a obrir nous camins. Gràcies, per la teva complicitat i comprensió.

Moltes gràcies, Ester! Moltes gràcies, Santiago! A vosaltres dos us agraeixo l'amabilitat amb la qual compartiu el vostre inacabable interès per les vostres passions.

Moltes gràcies, Mònica! Compartim Mariona, ordinador i taula… i una pila de moments divertits! Ets molt guay, Mònica!

Moltes gràcies, Eli! Sempre plena d'energia i passió! Una papallona il·luminada!

Gràcies, Manu i Ricardo pel vostre suport informàtic!

Nora! Fue fantástico encontrarte en los primeros años de tesis, me ayudaste con humor y cariño a tomarme el camino desde puntos de vista distintos.

Valentina, thank you so much! You have been a friend whose talks were so kind that opened a door in the heart of everybody. We all know that we have a friend in Siberia.

Gràcies a la secció d'eslaves per oferir-me ajuda i un somriure sempre que ens hem creuat. A més, voldria agrair l'ajuda que m'han donat, en diferents moments de la tesi, el Ricard amb les seves traduccions i el Marc, l'Andrei i la Natàlia en els experiments.

Gràcies al departament de lingüística, especialment, a la Irene, la Joana i la Itziar.

Paqui, muchas gracias! Tu manera de querer hacer bien las cosas es un modelo a seguir para cualquiera. También lo es tu cariño que siempre me ha parecido infinito!

Marga, no només m'has donat la clau o m'has dit que la Montse ja estava a dalt, també has pintat el hall de la universitat amb els teus bons dies d'un color agradable i acollidor. Gràcies! I tu, Carlos, també moltes gràcies per les teves claus acompanyades de bromes esbojarrades!

Moltes gràcies, Montse! Sempre l'una per l'altra en color i en blanc i negre! Gràcies per estar sempre al meu costat donant-me el suport, la comprensió i el carinyo que necessitava.

Fanny, siempre comiendo con señora tesis… gracias, porque siempre encontrabas la manera de ponerle una sonrisa a tesis. Y, sobra decir, que gracias por ayudarme en mil y una cosas de la Massana.

Maga, quanta comprensió i carinyo… i 'chuchitea'! Gràcies, Mags!

Alba, moltes gràcies per les nostres converses. Sempre tracen una línia blanca a la meva vida.

Míriam, malgrat la nostra història sempre té alguna cosa de punts suspensius, gràcies per deixar-me entendrem a través de la teva ànima, moltes vegades un mirall on veure-m'hi.

Chencho, sempre recordaré les teves paraules 'quan trobis allò que necessites, ho sabràs, et cridarà'.

Cristina, ens veiem poc però, quan quedem i fem terapia vital, sempre fem una mica de net. Gràcies, Cris!

Liudmila, muchas gracias! Siempre has estado allí, aunque estuvieras en otra parte, porque sabía que podía contar contigo para cualquier cosa! Gracias por darme siempre aliento!

Carlos, moltes gràcies per donar-me forces a través de les classes i del Projecte de fi de cicle de la Massana. Les teves paraules, observacions i gestos sempre han estat el contrapunt que la meva matèria gris necessitava per obrir pas a coses i perspectives noves, refrescants i apassionants. Gràcies, per obrir-me una porta tan reveladora.

Barbacoa team, vosaltres, sempre esteu a punt per fer una barbacoa, un barça o un el que sigui! Sempre allà, omplint de *joie de vivre* tots els anys nous que vinguin.

Gràcies, Berthus, per tota la poesia que has aportat a la meva vida.

Mamà, sempre allà, sempre donant aire, sempre aturant el meu cervell quan es passava de voltes i anava com un *tsunami*. Gràcies a la vida per donar-me una mare com tu.

Papá, fuerza y valor! Gracias a ti, este es el *leitmotiv* que no dejo de repetirme en los momentos que tengo que ser valiente. Gracias por los cimientos sobre los que me ayudas a edificarme.

Alejandra, petita boleta, en tu tinc sempre l'abraçada que necessito quan estic a punt de caure.

Avi, ets tot ganes de viure! I això no és fàcil de trobar. Gràcies, per omplir de moments preciosos i dolços la meva vida.

Àvia, estimo l'àvia d'ara, la d'ahir, la de fa un any, la de demà i, sobretot, la de sempre. Sempre la meva nena del bosc!

Gracias, tía Lidia, por tu amor inmenso. Gracias por estar siempre allí.

Gràcies als tiets i als cosinets… Sempre atents i preocupats per la meva tesi! I ara també als petits Amir i Waleed per haver arribat a les nostres vides.

Lídia, gràcies perquè, malgrat tu, ets una lleona molt dolça!

Gràcies, Josep, per la teva falta de por a les coses noves per més extravagants que puguin semblar!

Tusca, Argo, Ludo, Safo, Trasto, Nemo, Nina, Miqui, gràcies per les vostres babes carinyoses! I a tu, Moshe, les teves abraçades peludes.

Adrià, ets el suport diari, el rellotge, el cronòmetre, el compàs, el regle, l'escaire, el cartabó, els totxos, el ciment, les tisores, la pega, els colors, les flors, els arbres i el cel. Gràcies.

…….

THIS THESIS HAS BEEN FUNDED BY THE *AJUTS PER A LA FORMACIÓ DE PESONAL INVESTIGADOR NOVELL* (FI-DGR) FROM THE GENERALITAT DE CATALUNYA

General contents

List of acronyms

ACC	Accusative case
Adj	Adjective
ADJ	Adjunct
AdvP	Adverbial phrase
agt	Agent
AgtC	Agentive complement
AP	Adjective phrase
Arg(0-1-2-3-4)	Argument 0, 1, 2, 3, 4
ArgM	Adjunct
AS	Argument structure
atr	Attribute
ben	Beneficiary
cau	Cause
coagt	Coagent
coexp	Coexperiencer
cot	Cotheme
DAC	Deeply Annotated Corpus
DAT	Dative case
Def	Definite determiner
Demonst	Demonstrative determiner
dest	Destination
DO	Direct object
ef	Final state
ei	Initial state
exp	Experiencer
ext	Extent
FEM	Femenine
$F_{IPF \& PF}$	Biaspectual form
GEN	Genitive case
goal	Goal
iarg(1-2-3-4)	Implicit argument
iargM	Implicit adjunct
Indef	Indefinite determiner
InfC	Infinitive clause
INS	Instrumental case
inst	Instrument
IO	Indirect object
IPF	Imperfective
loc	Location
MiniRuSp	Subsample of the Russian-Spanish corpus

mnr	Manner
MT	Machine Translation
n	Noun
$N_{[+1]}$	One or more than one nominalization derived
NC	Noun complement
Neg	Negative determiner
N_N	Neutralized nominalization
NOM	Nominative
NP	Noun phrase
N_S	Symmetrical nominalization
N_U	Uniaspectual nominalization
p	Preposition
pat	Patient
PF	Perfective
PL	Plural
PoS	Part of Speech
Poss	Possessive determiner
PP	Prepositional phrase
PREP	Prepositional case
R	Russian
RelC	Relative clause
RNC	Russian National Corpus
RuSp	Russian-Spanish corpus
S	Spanish
SG	Singular
SL	Source language
Spec	Specifier
src	Source
SubC	Subordinate clause
Subj	Subject
tem	Theme
TL	Target language
tmp	Time
V_B	Biaspectual verb
VP	Verbal phrase
V_P	Paired verb
V_U	Uniaspectual verb

Chapter 1

Introduction

The richness of natural language is reflected in the variety of different ways to express similar concepts. Very often what can be expressed by means of a verb, can be also expressed by means of its corresponding deverbal noun. Example (1) shows that the same process can be expressed by means of the verb *perevodit'* 'to translate' and its corresponding deverbal noun *perevod* 'translation'. The action of translating could be generalized in both constructions as the formal representation 'X-Agent event Y-Theme'. This representation is valid for the sentence in (1.i) and for the nominal phrase in (1.ii). Both predicates share the same number and type of arguments, that is, an Agent and a Theme. However, the way in which these arguments are syntactically realized is different. In (1.i) the Agent (agt) of the action is expressed by means of a personal pronoun *on* 'he' acting as the subject (Subj); while in (1.ii) the Agent is expressed by means of a possessive determiner *ego* 'his' acting as the specifier (Spec) of the noun phrase (NP). In (1.i), the Theme is expressed by means of a NP in accusative *Gamlet* 'Hamlet' acting as an object, whereas in (1.ii) it is expressed by means of a NP in genitive *Gamleta* 'of Hamlet', which acts as a noun complement (NC). The adjunct indicating the Manner (mnr) in which is translated the text is expressed by means of the adverb *akkuratno* 'carefully' in (1.i) and by means of the adjective *akkuratnyj* 'careful' in (1.ii). The study of deverbal nouns is necessary if we want to do a semantic analysis of a text, specially, if we are interested in establishing the semantic relations between the arguments and their predicates.

(1)
(i) *On*$_{\text{NP\_Subj\_Arg0\_agt}}$ [***perevodil*** *Gamlet*$_{\text{NP\_DO\_Arg1\_tem}}$
HE **TRANSLATED** HAMLET
akkuratno$_{\text{AdvP\_ADJ\_mnr}}$]$_{\text{VP}}$
CAREFULLY
'He **translated** Hamlet carefully.'

(ii) [*Ego* $_{\text{Poss\_Spec\_Arg0\_agt}}$ *akkuratnyj*$_{\text{AP\_NC\_ArgM\_mnr}}$ ***perevod***
HIS CAREFUL **TRANSLATION**
HAMLET
Gamleta$_{\text{NP\_NC\_Arg1\_tem}}$]$_{\text{NP}}$
'His careful **translation** of Hamlet'

Therefore, (1.i) and (1.ii) can be viewed as two alternative ways of expressing the same event, that is, by means of a verbal and a nominal predicate. This would be an example of a diathesis alternation in terms of Levin (1993), two different patterns or constructions in terms of Goldberg (1995), and an intralinguistic mismatch in terms of Mel'chuk & Wanner (2006).

This thesis presents an empirical descriptive study of Russian deverbal nouns, a subtype of nominalizations, which have been derived from verbs. Agentive deverbal nouns are not studied in this thesis, because, despite of being derived from verbs, they do not present ambiguity in their denotation and in this thesis we are focused on deverbal nouns that denote events, results and states. For instance, an agentive deverbal noun such as *tancovščik* 'dancer', which denotes the person who carries out the action named by the base verb *tancevat'* 'to dance'. This verbal origin involves that the deverbal noun is a hybrid or a mixed category, that is, it has both nominal and verbal properties. Because of the hybrid character of these nouns, their linguistic representation has been a controversial point. As a common noun, a deverbal noun is the head of a NP and can be specified and complemented in order to restrict and determine the noun reference. But, in contrast to common nouns, the reference of a deverbal noun is not always static, that is, the reference is not always a concrete entity. As its corresponding base verb, a deverbal noun can denote events, states and results, and shares the argument structure of their base verb. By event, we mean a dynamic situation that takes place in time (2); and by result, we mean the concrete or abstract entity that results from an action or a process (3) and, finally, by state deverbal nouns, we mean those nouns that denote the non-dynamic situation named by the verb (4).

(2) [*Ego*$_{\text{Poss\_NC\_Arg0\_agt}}$ ***perevod*** *sonetov* $_{\text{NP\_NC\_Arg0\_agt}}$]$_{\text{NP}}$ *zanjal* *tri* *goda*
HIS **TRANSLATION** SONNETS TOOK THREE YEARS
'His **translation** of the sonnets took three years.'

(3)[*Ego* $_{\text{Poss\_Spec\_Arg0\_agt}}$ ***perevod*** *na* *stole*$_{\text{PP\_NC\_ArgM\_loc}}$]$_{\text{NP}}$
HIS **TRANSLATION** ON TABLE
'His **translation** is on the table.'

(4) *No* *eto* *bylo* *nastol'ko* [[*moim*] Poss_Spec_Arg0_exp ***pereživaniem***],
BUT THIS WAS SO_MUCH MY **WORRY**
čto *ja* *ne* *rešilsja* *sprosit'* *u* *papy*
THAT I NOT DECIDED TO_ASK TO DADDY
'And my **worry** was so great that I did not make up my mind to ask my father.'

In (2) the deverbal noun *perevod* 'translation' denotes an event since it expresses the action, which could be expressed by means of a verb, that is, *on perevodil sonety na tri goda* 'He had been translating the sonnets for three years'. In (3) the deverbal noun *perevod* denotes a result since names the resulting object of the action of translating. In (4) the deverbal noun names a state, that is, the non-dynamic situation named by the base verb.

Consequently, one of the main goals of this research is focused on studying to what extent morphological and lexical aspects of the base verb determine the lexical denotation of its corresponding deverbal noun.

As we will see later, we can define a continuum between verbs and nouns. Result nouns are closer to prototypic nouns, while event nouns are closer to the verb. The closeness of result nouns to common nouns has provoked a linguistic discussion about the argument ability in result deverbal nouns. Authors such as Grimshaw (1990) and Zubizarreta (1987) claim that result deverbal nouns do not have argument structure, whereas authors such as Mel'chuk *et al.* (1984), Pustejovsky (1995), Picallo (1999), Alexiadou (2001), Meyers (2007) and Peris (2012) claim that result nouns also have arguments. On the other hand, there is no discussion about the existence of argument structure in event deverbal nouns, which are closer to verbs and inherit the argument structure of its corresponding base verb. Following these authors, in our study we assume the existence of argument structure for both types of nouns. A NP headed by a deverbal noun denotes a predication similar to its corresponding verb and express the same arguments than its base verb.

A second aim of this thesis is the analysis of the syntactico-semantic structure of this type of predicates, concretely, we describe which types of constituents are their semantic arguments, how they are realized and which are their possible combinations with the aim of describing the most frequent patterns of these nouns.

The study of the argument structure is the base on which we will ground the comparative study of Russian-Spanish translation mismatches regarding deverbal noun constructions in both languages, which is the third aim of this thesis. In our approach, we consider a translation as a type of paraphrase, that is, as a way to convey the same information in a different form. We can distinguish between intralinguistic and interlinguistic paraphrases. In the former, paraphrase takes place in the same language. We understand an event deverbal noun construction as an intralinguistic paraphrase, since the same meaning could have been conveyed by means of a verbal predicate. On the other hand, translation mismatches between

Russian and Spanish deverbal noun's constructions can be considered as interlinguistic paraphrases. The difference between intra and interlinguistic paraphrases is similar to the intra and interlinguistic mismatches proposed by Mel'chuk & Wanner (2006). It is relevant to notice that paraphrasing is not conveying the exactly same meaning. As Vila, Martí & Rodríguez (2013:9) point out "paraphrasing must be situated in the field of approximation, opening the path to different semantic similarity or paraphrasality degrees". Following this path, we conceive translation mismatches as different degrees of semantic similarity between the source and the target languages. At one edge of the continuum, we would have a word-for-word translation of a sentence; while at the opposite edge we would have a freer translation, that is, a continuum from more to less literal translation.

Therefore, the study presented in this thesis is threefold:

(1) the relationship between the morphological and lexical aspect of the base verb and the lexical denotation of the deverbal noun derived;

(2) the argument structure of deverbal nouns and whether it is related to the aspectual features of the lexical denotation, and

(3) Russian-Spanish translation mismatches within the NP headed by the deverbal noun.

These are the main three goals of this research and it has been carried out following a corpus-based approach. Therefore, we have grounded our hypotheses and observations on real data. We have used monolingual and bilingual (even multilingual) Spanish and Russian corpora and lexicons. All the presented classifications are based on real examples extracted from these corpora.

1.1 Contributions of the research

In the development of this study on deverbal noun constructions we have done several contributions that are highlighted in this section.

1. As we have already said, our conclusions are based on the results of studying a sample of real language by means of corpora. Therefore, our observations and descriptions always emerged and are contrasted with examples extracted from corpora, that is, from real data. In order to carry out this study we have created a parallel Russian-Spanish corpus, named RuSp, which consists of 710,622 tokens of written texts. Basically, the parallel corpus is composed of literary texts and juridical texts. Texts have been parallelized at the paragraph level and have been partially and manually annotated. We have conducted a syntactico-semantic annotation of the NPs headed by deverbal nouns. (Chapters 2, 4 and 5)

2. An empirical linguistic study of Russian deverbal nouns (Chapter 3) has been carried out which results in:

(a) a classification of deverbal nouns according to the morphological aspect of the base verb;
(b) an analysis of the lexical denotation of deverbal nouns and its relation with the base verb morphological and lexical aspect;
(c) a revision of the widely used criteria to distinguish between the lexical denotation of deverbal nouns, and the detection of denotative selectors to discriminate between event and result nouns in Russian.

3. We have determined which constituents can be argumental in Russian and we provide a detailed description of the typical patterns of the argument structure of deverbal nouns. (Chapter 4)

4. A subsample of RuSp, named MiniRuSp and consisting of 500 occurrences of deverbal nouns in Russian and its corresponding translations to Spanish has been syntactico-semantically analyzed and annotated manually. The annotation for both languages is partial since it only includes NP's constructions headed by the deverbal noun under study and their parallel constructions in Spanish. The annotation includes morpho-syntactic information (constituency and function structures) and semantic information (argument structure and its corresponding thematic roles). Moreover, we have also annotated the lexical denotation of the deverbal nouns under analysis. The analysis of the argument structure has allowed us to determine the different argument structure patterns of deverbal nouns in Russian. (Chapter 3)

5. The analysis of MiniRuSp at a syntactico-semantic level of both languages is the base on which we ground our Russian-Spanish comparative study from which the identification and classification of the frequent translation mismatches has resulted. This study has resulted in a database with 114 different deverbal nouns classified according to the translation mismatches and the linguistic changes involved. We do not provide the database in the appendices because of its size. Some of the regular mismatches are typological and then could be possibly detected automatically. (Chapter 4)

1.2 The structure of the dissertation

This thesis is structured in six chapters: (1) the present introduction; (2) a brief description of the linguistic resources used in this study; (3) the analysis of the relationship between the aspect of the base verb and the lexical denotation of the

deverbal noun; (4) the study of the argument structure; (5) the analysis of translation mismatches and (6) our final conclusions. Moreover, we include two appendices.

Before going any further, it is important to notice that we have included the state of the art related to our work in each chapter.

In the second chapter, we present the different linguistic resources used. The presentation of these resources is not an exhaustive revision of all the existing ones, we have only described those resources that have been used in this research. We have used different monolingual and bilingual corpora depending on the linguistic aspect that we were looking at. For instance, to study the lexical denotation of Russian deverbal nouns we used basically *The Russian National Corpus* (Apresjan *et al.,* 2005) and *The Essex database of Russian verbs and their corresponding deverbal nouns* (Spencer & Zaretskaya, 2010), as well as, *AnCora-Es corpus* (Peris & Taulé, 2012) and the Spanish nominal lexicon *AnCora-Nom* (Peris & Taulé, 2011). To study the argument structure of deverbal nouns and the translation mismatches between Russian and Spanish deverbal nouns we used bilingual corpora, concretely, *RuSp* (De Valdivia, Castellví & Taulé, 2013) and *UNGAR* (Rafalovich & Dale, 2009). All these linguistic resources are described in chapter 2.

In the third chapter, we analyze to what extent the morphological and the lexical aspect of the base verb determines the lexical denotation of its corresponding nominalization, in our case its denotation as an event, as a result or as a state. The initial hypothesis of our research is that it is not possible to say that morphological and lexical aspect of the base verb determine the lexical denotation of the deverbal noun, although it seems that both have significant influence on the lexical aspect of the nominal. In order to examine this hypothesis we proceed as follows. First, we analyze the different types of nominalizations on the basis of the traditional verbal classification. The result of this analysis is a deverbal nominalization classification, which enables us to determine from which class of verb the nominalization is derived and whether the nominalization has inherited morphological aspectual marks – either imperfective or perfective – from the corresponding verb. Second, we analyze each type of nominalization in examples from real data in order to establish their denotation or lexical aspect (i.e. event, result or state). The main two goals in doing this analysis are, first, to determine the influence of the morphological aspect of the base verb on the lexical denotation of the nominalization, and, second, to determine the influence of the lexical aspect of the base verb on the lexical denotation of the deverbal noun. We review the most widespread criteria used in the literature on nominalizations in order to distinguish the denotation of deverbal nouns, such as the expression of the internal argument, the ability to pluralize and the presence of denotative selectors. We analyze whether these criteria work or not for Russian deverbal nouns.

In the fourth chapter, we study the argument structure of deverbal nouns. We assume that nouns derived from verbs inherit its argument structure. The arguments

can be realized explicitly inside the NP headed by the deverbal noun, incorporatedly inside the deverbal noun lexeme or, finally, they can be realized implicitly and be expressed outside the NP headed by the deverbal noun. We have annotated manually the MiniRuSp corpus syntactically (constituent and function structures) and semantically (arguments with their corresponding thematic roles). Our study is validated by the results of the annotation of 500 occurrences. We determine which constituents can be arguments and which combinations of constituents, arguments and thematic roles can take place in the NP headed by a deverbal noun in Russian. In the study of the argument structure of deverbal nouns, we can define the internal structure of deverbal nouns by defining the possible syntactico-semantic patterns. This analysis is the base upon which we analyze translation mismatches.

In the fifth chapter, we present the different types of interlinguistic mismatches between Russian and Spanish deverbal noun's structures, obtained by means of a corpus-based analysis of the data extracted from RuSp and UNGAR parallel corpora. To do this we have taken into account the properties of translated texts and previous classifications of translation mismatches for other languages. These previous proposals were focused on verbs, whereas our proposal adapts and enlarges them to fit Russian deverbal nouns. Our classification is mainly interested in giving account of those productive and regular mismatches. We describe and classify mismatches depending on the linguistic changes produced rather than on the reasons behind these linguistic changes –typological, pragmatic, cultural or subjective. In this chapter, we propose a list of linguistic changes (Determiner change, Part of Speech change, Syntactic Function change, Explicitation Argument Structure change, Correference-elliptic change, *etc.*) and a classification of mismatches based on the number and type of linguistic changes involved.

In the sixth chapter, we present our final conclusions and further research.

Finally, we add two appendices. Appendix A corresponds to complementary information including the list of deverbal nouns used in the study accompanied by the frequency of appearance in RuSp and MiniRuSp. In Appendix B, we find translation correspondences between deverbal nouns' arguments in Russian and its counterpart in Spanish.

This research will proof the general initial hypothesis which claims that despite the morphological-syntactico-semantic relation between the verb and the noun, the noun has its own features and because of that, even if the base verb can have to some extent an influence on the noun, the effects in the noun are not the same as in the verb.

1.3 Some formal aspects

Examples are given in Russian using the transcription convertion ISO, namely ISO/R9 in italics. At the level of the Russian transcription, we include the linguistic

information required depending on the issue under study (aspectual tags, syntactic and semantic tags). Regarding glosses, that is, literal translations, they are in small caps aligned word-for-word with the Russian transcription. Glosses show the number and gender of nouns (but not the case), and, regarding verbs, glosses show their tense and person (but, not the aspect). Finally, we provide the non-literal translation in English. All the examples given in this thesis have been extracted from different corpora, see chapter 2 to read about the different linguistic resources used.

Chapter 2

Linguistic resources

To carry out the linguistic analysis of Russian deverbal nouns, we have used different linguistic resources, basically corpora but also lexicons, in order not only to describe the phenomena under consideration but also to contrast and build our proposals from the observations of real data. Therefore, we follow a corpus-based approach to describe Russian deverbal nouns. In this chapter we only present the corpora for Russian language that have been employed in this work. Different monolingual and bilingual corpora have been used depending on the linguistic aspect that we were studying. For instance, to study the lexical denotation and the argument structure of Russian deverbal nouns we basically have used two monolingual resources: a corpus and a lexicon, concretely, the *Russian National Corpus* (Apresjan *et al.,* 2005) and the *Essex database of Russian verbs and their corresponding deverbal nouns* (Spencer & Zaretskaya, 2010), whereas bilingual corpora such as *RuSp* (De Valdivia, Castellví & Taulé, 2013) and *UNGAR* (Rafalovich & Dale, 2009) have been mainly employed to study translation mismatches between Russian and Spanish deverbal nouns. The Spanish nominal lexicon *AnCora-Nom* (Peris & Taulé, 2011) was mainly consulted to compare deverbal nouns in Spanish and Russian.

There is a number of linguistic resources which directly or indirectly have contributed to the present work. For instance, we have taken into account the annotation scheme of *NomBank*[1] (Meyers *et al.,* 2004b; Meyers, 2007) and *AnCora* (Martí *et al.*, 2008; Taulé *et al.,* 2008; Recasens & Martí, 2010; Peris & Taulé, 2012) and the frames of *FrameNet*[2] (Baker *et al.,* 2003). The other resources presented below are related work. There exist other linguistic resources for other languages that deal with deverbal nominalizations. Regarding the English language, there are *Nomlex*[3] (Macleod *et al.,* 1998) and *NomBank*. Nomlex focuses on the argument structure of deverbal nouns. *NomBank* focuses on the argument structure of all nouns appearing

[1] http://nlp.cs.nyu.edu/meyers/NomBank.html
[2] https://framenet.icsi.berkeley.edu/fndrupal/
[3] http://nlp.cs.nyu.edu/nomlex/

in the *Penn Treebank* (Palmer *et al.,* 2005). There is a special *NomBank* (Xue 2006, 2008) for Chinese deverbal nouns. Another well-known resource is *FrameNet*, which is a project based on the Frame Semantics (Fillmore, 1968) and it is focused on the syntactico-semantic representation of nouns, verbs and adjectives. Therefore, deverbal nouns are represented. *FrameNet* is not only concerned with English, there are other *FrameNets* for other languages, such as German (Burchardt *et al.* 2009), Japanese (Ohara, 2009) and Spanish (Subirats, 2009). For French deverbal nouns there is a project named *NOMAGE* (Balvet *et al.* 2010, 2011). Finally, there is an ongoing project, the *Copenhaguen Dependency Treebank* (Hoeg Müller, 2011), which is aimed at the annotation of the argument structure in a parallel corpus for Danish, German, English, Italian and Spanish.

However, the present chapter is only focused on those Russian and Spanish resources that have been used in our research.

This chapter is structured in the following way: In section 2.1, we present the three monolingual corpora (the *Russian National Corpus*, the *Essex database of Russian verbs and their corresponding deverbal nouns* and the *AnCora-Es*); and in section 2.2, we present the bilingual corpora (*RuSp* and *UNGAR*).

2.1 Monolingual corpora

In this section we present the monolingual resources used in the present research: the *Russian National Corpus* and the *Essex database of Russian verbs and their corresponding deverbal nouns* (hereinafter, the *Essex database*) for Russian language, whereas *AnCora-ES* and its corresponding *AnCora-Nom* nominal lexicon for Spanish language. These linguistic resources contain different levels of annotation and they are available and consultable online.

2.1.1 The Russian National Corpus

The *Russian National Corpus* (hereinafter, RNC) is probably the largest corpus of modern Russian language, consisting of 364,881,378 words (tokens), created by the Institute of Russian language of the Russian Academy of Sciences[4] with the goal of being a reference corpus.

RNC contains written and spoken texts of different genres (memoirs, essays, journalistic works, scientific and popular scientific literature, public speeches, letters, diaries, documents, among others), different registers (literary, colloquial and vernacular) and different dialectal varieties. Written texts are more widely represented and, specially, literary texts are the commonest (representing 40% of the total), which range from the 18th century to the early 21st.

[4] http://www.ras.ru/en/index.aspx

The RNC is composed of seven subcorpora automatically PoS tagged and lemmatized: the *main corpus*, the *deeply annotated corpus* (hereinafter, *DAC*), the *parallel corpus*, the *poetry corpus,* the *dialectal corpus*, the *educational corpus* and the *corpus of spoken Russian*.

The *main corpus* and the *DAC* are the corpora that we have used in our study, because of that we describe them in more detail in the following two sections. The *parallel corpus* is composed of parallelized corpora from the following languages (both directions): English-Russian, German-Russian, French-Russian, Spanish-Russian, Italian-Russian, Polish-Russian, Ukrainian-Russian, Byelorussian-Russian and contains a total of 37,822,091 words. Actually, half of the words of these corpora correspond to the pair Russian-English (15,842,627). The Spanish-Russian pair contains 177,836 words. The *dialectal corpus* consists of recordings of dialectal speech from different regions of Russia presented in standardized orthography and it contains 194,283 words. The *poetry corpus* contains poems ranging from 1750 to 1890, but also some poets of the 20th century. It is morphologically tagged and it also has special tags such as the poetic meters. The *educational corpus* is a small corpus adapted for the Russian educational program, including works of fiction on the school reading list and contains 664,751 words. Finally, the *corpus of spoken Russian* includes the recordings of public and spontaneous spoken Russian and transcripts of Russian movies, this corpus contains 10,361,579 words.

The main corpus

The *main corpus* includes texts representing standard Russian and contains 229,968,798 words. It can be subdivided into 3 parts: Modern written texts (from 1950 to the present day), a subcorpus of real-life Russian speech (from 1950 to the present day) and early texts (from the first half of 18th to the first half of 20th century). Every text included in the main corpus is meta and morphologically tagged, moreover, most of the words are also semantically tagged. Meta information gives information about the author, title, chronology and genre of the text. Morphological information (Liashevskaya O. N. *et al.* 2005) is based on the model proposed by Zalizniak (2003 [1977]) in *Grammatical dictionary of Russian*. Morphological tagging (lemma and PoS[5]) was carried out automatically. From this corpus a subgroup of texts (6,000,000 tokens approximately) has been manually disambiguated and validated. Regarding morphological tagging, every word is tagged with the following information: lemma, PoS, gender, animacy, number, tense, and aspect, among others. Regarding semantic tagging, words are tagged with semantic features based on the classification developed by Paducheva & Rakhilina (1992) such as concrete, abstract, proper noun, human or animal, among others. Moreover, we also find information about word formation: prefixes, suffixes, root.

[5] Part of Speech

In figure 1, we present a screenshot of the *main corpus* in the RNC interface. The figure shows the search results for the word *perevod* 'translation', which has been found in 5,143 documents and it has 16,547 occurrences in the main corpus.

Figure 1: Russian National Corpus

Figure 2 shows the results obtained by clicking on the word *perevoda* 'translation'. After clicking, a window with morphosemantic information pops up. The window provides us the following information:

Regarding morphological information[6], in the first row, the lemma (*pervod* 'translation'), and in the second row, grammatical information such as the PoS (noun), the animacy (*perevod* is inanimate), the gender (masculine), the number (singular) and the case (genitive).

Regarding semantic information[7], in the third row, the deverbal noun *perevod* 'translation' is an abstract noun derived from a verb belonging to the class of movement. Finally, in the fourth row, we find additional semantic information that informs us that if the word is an object, that is, a concrete entity, then it belongs to the textual class (t:text).

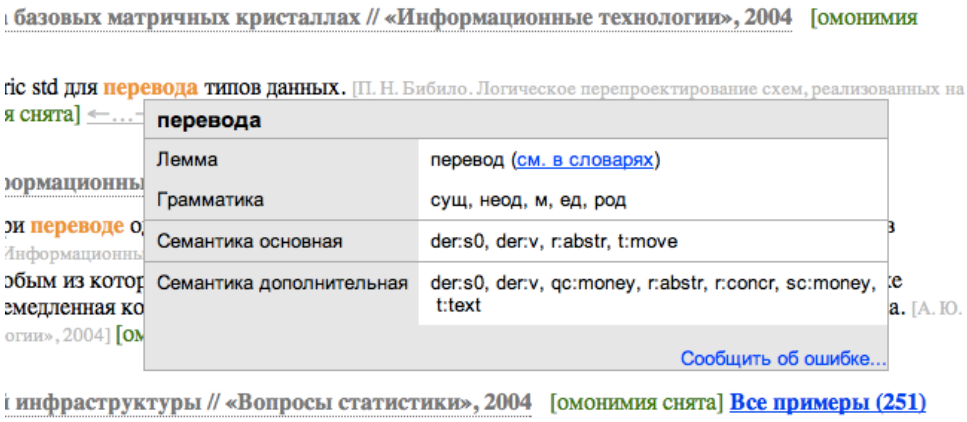

Figure 2: Morphosemantic information provided in the main corpus of RNC

The *main corpus* has been the reference corpus that we have used to study the different morphological and semantic aspects of Russian deverbal nouns such as the relationship between morphological and lexical aspect of verbs and the lexical denotation of their corresponding nominalizations.

The deeply annotated corpus[8] (hereinafter, DAC)

DAC (Apresjan *et al.,* 2005) is composed of 757,794 different tokens. This corpus includes literary texts of the 20th century, scientific literature, journalistic articles and articles about politic life. Every sentence was annotated automatically at a morpho-

[6] http://ruscorpora.ru/en/corpora-morph.html
[7] http://ruscorpora.ru/en/corpora-sem.html
[8] We thank members of the Russian National Corpus for providing access to us DAC for academic purposes.

syntactic level by means of ETAP-3[9], which includes a morphological analyzer and a syntactical analyzer. After the automatic tagging, *DAC* was fully disambiguated and validated by hand.

DAC uses dependency trees as its annotation formalism: in nodes we find words and at the edges we find the type of syntactic relationship (first complement, predicate, adverbial adjunct, among others). This syntactic annotation scheme is based on the "Meaning–Text" linguistic model by Igor A. Mel'chuk (1974). *DAC* is organized as separated files in xml with morphological and syntactic information. In figure 3, we can see an example of the syntactic structure of the sentence *britanskij učenyj i pisatel' Endrju Kollins detal'no izučil različnyje varianty perevoda dialogov Platona* 'The British researcher and writer Andrew Collins studied in detail the different variations of the dialogues of Plato'. In the analysis, the main node is the verb *izučat'* 'to study' which has a subject (*predikativnoe SintO*), which is the noun *učenyj* 'researcher' and a direct complement (*first complement*) which is *variant* 'variations'.

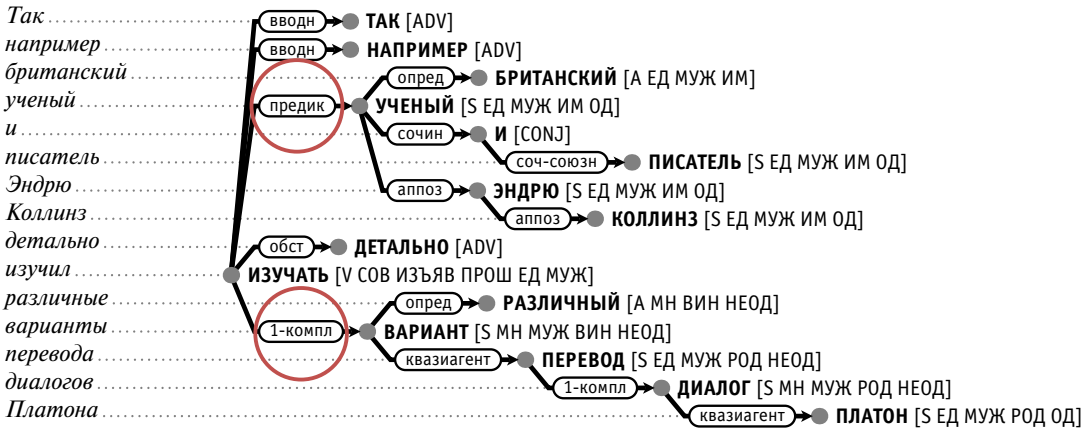

Figure 3: A sentence morpho-syntactically tagged in DAC

DAC is consultable at the address http://www.ruscorpora.ru/search-syntax.html.

2.1.2 The Essex Database of Russian verbs and their corresponding nominalizations (Essex Database)

The *Essex Database* is composed of 7,000 Russian verbs and their corresponding 5,000 nominalizations. The database consists of 3 parts: a table of verb entries, a table of nominalizations, and a table of noun-verb links. This database codifies morpho-syntactic and semantic information for both verbs and its corresponding nominalizations. The information codified is the following: (a) morphological (lemma and PoS); (b) syntactic (subcategorization frame, i.e. the different frames in

[9] http://www.ruscorpora.ru/instruction-syntax.html

which the predicates can occur); and (c) semantic (lexical aspect and argument structure). Regarding lexical aspect, verbs are classified into activities, accomplishments, achievements or states; while in the case of nouns, they are classified into complex event nominal, a simple event nominal or a result (following Grimshaw, 1990). Specifically, for verbs it is also included the information about their morphological aspect (perfective or imperfective) and their derived nominalizations, whereas in the case of nominalizations it is coded the corresponding verb from which the nominalization is derived and its nominalizing suffix.

In table 1, we present the information provided by the *Essex Database* regarding the verb *carapat'* 'to scratch'. The first column presents the identifier, which relates the verb with its corresponding deverbal noun (*carapan'e* 'scratching'). There is a special table of correspondences between verbs and their corresponding deverbal nouns. This table links the primary keys (identifier field) of the noun to a verb entry. The second column presents the infinitive form (*carapat'* 'to scratch'). The third and fourth columns are the translation of that word into English (scratch) and the definition of that word in Russian (*nanosit' carapiny kogtjami* 'to cause a scratch with the nails'). The fifth column shows the list of nominalizations derived from the verb (*carapan'e* 'scratching'). The sixth and seventh columns indicate the aspect of the verb (imperfective in the example) and its aspectual pair in case of existence (the sign '-' indicates that the verb *carapat'* 'to scratch' does not have an aspectual pair, then it is uniaspectual). The eighth and ninth columns indicate primary situation types (state, activity, achievement, accomplishment, among others) and secondary situation types (bounded process, iterative, mental, among others). The tenth column indicates the semantic type of the verb (that is, pseudo-behaviour, accompanying, body posture, etc.). The eleventh column shows a binary feature called 'control' that indicates whether the subject has the control of the event named by the verb. The twelfth column indicates whether the situation type is expressed morphologically or not (for instance, by means of the prefix *po-* is possible to indicate attenuation). The thirteenth column indicates the predicate argument structure (that is, the thematic roles: Agent and Theme in the case of *carapat'*). The fourteenth and fifteenth column indicate the verb root and the basic verb (that is, the unprefixed verb) from which the verb has been derived. The sixteenth column indicates the syntactic subcategorization of the verb. The seventeenth column indicates information about the lexical conceptual structure of the verb ('cause change of position' for *carapat'*). The eighteenth column indicates the prefixes used to form the verb (here '-' means, in the case of *carapat',* that there are not prefixes). Finally, the nineteenth column indicates the conjugation class (1st for *carapat'*).

ID	Verb	Gloss	Gloss2	Nominalization	Aspect	Aspectual pair	Sit Type 1	Sit Type 2	Semantic Type
353	carapat'	scratch	*nanosit' carapiny kogtjami*	carapan'e	impf	-	act	-	pseudo-behaviour
Control	**Aktionsart**	**PAS**	**Root**	**Basic verb**	**Subcat**	**LCS**	**Morpho logy**	**Conj class**	
+		Ag-1, Th-2	carap	carapat'	NP-1, NP-2	cause change of state	-	1aj	

Table 1: The verb carapat' 'to scratch' in the Essex Database

In table 2, we present the information provided by the *Essex Database*, regarding the deverbal noun *carapan'e* 'scratching'. As we see from table 2, different information is provided: the base verb (*carapat'sja* 'to scratch oneself') and the basic verb (*carapat'* 'to scratch'), a verb and a noun gloss, that is, the definition of the verb and the noun derived (*Skrestis' pytajas' proniknut' kuda-l* 'to scratch trying to penetrate somewhere' and *dejstvie po glag.* 'the action denoted by the verb', respectively), the argument structure and changes in the argument structure (in table 2, it is indicated that in *carapan'e* 'scratching' the Agent is deleted), the aspect (imperfective), nominal suffix (*–an'e*) added to the verbal root in order to get the deverbal noun, and lexical denotation of the deverbal noun (here CE, that is, complex event).

ID	Noun	Verb	Verb Gloss	Gloss	Notes	PAS
15391	carapan'e2	carapat'sja2	*Skrestis'pytajas' proniknut' kuda-l*	dejstvie po glag.		
PAS change	**Aspect**	**Basic verb**	**Noun suffix**	**G-change**	**S-change**	
Ag deleted	impf	carapat'	an'e	-	CE	

Table 2: The deverbal noun carapan'e 'scratching' in the Essex Database

We have used this database to select nominalizations derived from perfective and imperfective verbs and formed by means of different nominalizing suffixes. By doing this selection, we have obtained a sample of 296 deverbal nouns derived from 294 different verbs used to classify the deverbal nouns. (See chapter 3, section 3.3).

Although, nowadays, it seems to be no longer available, the *Essex Database* was downloadable when we were doing our proposal of deverbal noun classification at http://privatewww.essex.ac.uk/~spena/res_interests.htm.

2.1.3 AnCora-Nom and AnCora-Es v.2.0

AnCora (Martí *et al.*, 2008; Taulé *et al.,* 2008; Recasens & Martí, 2010; Peris & Taulé, 2012) is a multilingual corpus composed of two subcorpora: *AnCora-CA* for the Catalan language and *AnCora-ES* for the Spanish language. Each corpus contains 500,000 words mainly from journalistic texts. *AnCora* is annotated at different linguistic levels: morphology (lemma and PoS), syntax (constituents and syntactic functions), semantics (argument structure with its corresponding thematic roles, semantic class of verbs, denotative types of deverbal nouns, nominal and verbal WordNet (Miller, 1995) senses and named entities) and, finally, pragmatics (coreference relations). The annotation process was carried out from lower-to-upper-level layers of linguistic description, which means that morphology was tagged first, second syntax, then semantics and, finally, pragmatics. The annotation was carried out manually, semi-automatically and fully automatically, depending on the linguistic level of analysis. Each level of annotation implied checking and completing the previous levels in order to guarantee high quality and minimize the error rate. Each layer of annotation was considered independently from the others.

Out of *AnCora*, two verbal lexicons *AnCora-Verb* (Aparicio *et al.,* 2008) and one nominal lexicon *AnCora-Nom* (Peris & Taulé, 2011) were built. Regarding the verbal lexicon, there are two parts: one devoted to the Spanish language with 2,647 entries and the other part with 2,143 entries for the Catalan language. Regarding the deverbal noun's lexicon, it is composed of 1,658 entries for the Spanish language. The information provided in each lexicon is the following: the verbal entries for each verbal sense of *AnCora-Verb* are annotated with the semantic class of the verb, its subcategorization scheme mapped to the argument structure with the corresponding thematic roles. Nominal entries of each nominal sense of *AnCora-Nom* are annotated with the denotative type of the noun (event, result and underspecified), with the WordNet synset, the argument structure with its thematic roles, and it is linked to the verb from which the noun is derived. All this information is complemented with examples of use.

Figure 4 shows a screenshot of *AnCora-Nom*. The lexicon provides the lexial entry of *preparación* 'preparation': the sense 2 of this noun is derived from the base verb *preparar* 'prepare', its denotation type is event, and it corresponds to the WordNet synset "16:00593220". The argument structure of this noun realizes an Arg1 with the thematic role of a Patient by means of a PP introduced by the preposition *de* 'of' and an optional ArgM with the thematic role Goal by means of a PP introduced by the preposition *en* 'in'. Moreover, the NP headed by the deverbal noun *preparación* 'preparation' can be specified by an article determiner and cannot pluralize. The information included in *AnCora-Nom* has been obtained from the annotated data of the *AnCora* corpus, that means that in *AnCora* the second sense of the noun *preparación* 'preparation' never appeared in plural and always was accompanied by the article determiner.

The *AnCora* corpora and the lexicons derived are freely available for research and can be downloaded from the main website of *AnCora* (http://clic.ub.edu/ancora/).

The annotation scheme used to analyze *RuSp* follows the same of *AnCora*. Moreover, *AnCora-Nom* lexicon has been used to contrast the information obtained for Russian language with the Spanish language. (See chapter 3, section 3.5.2 and chapter 4, section 4.2).

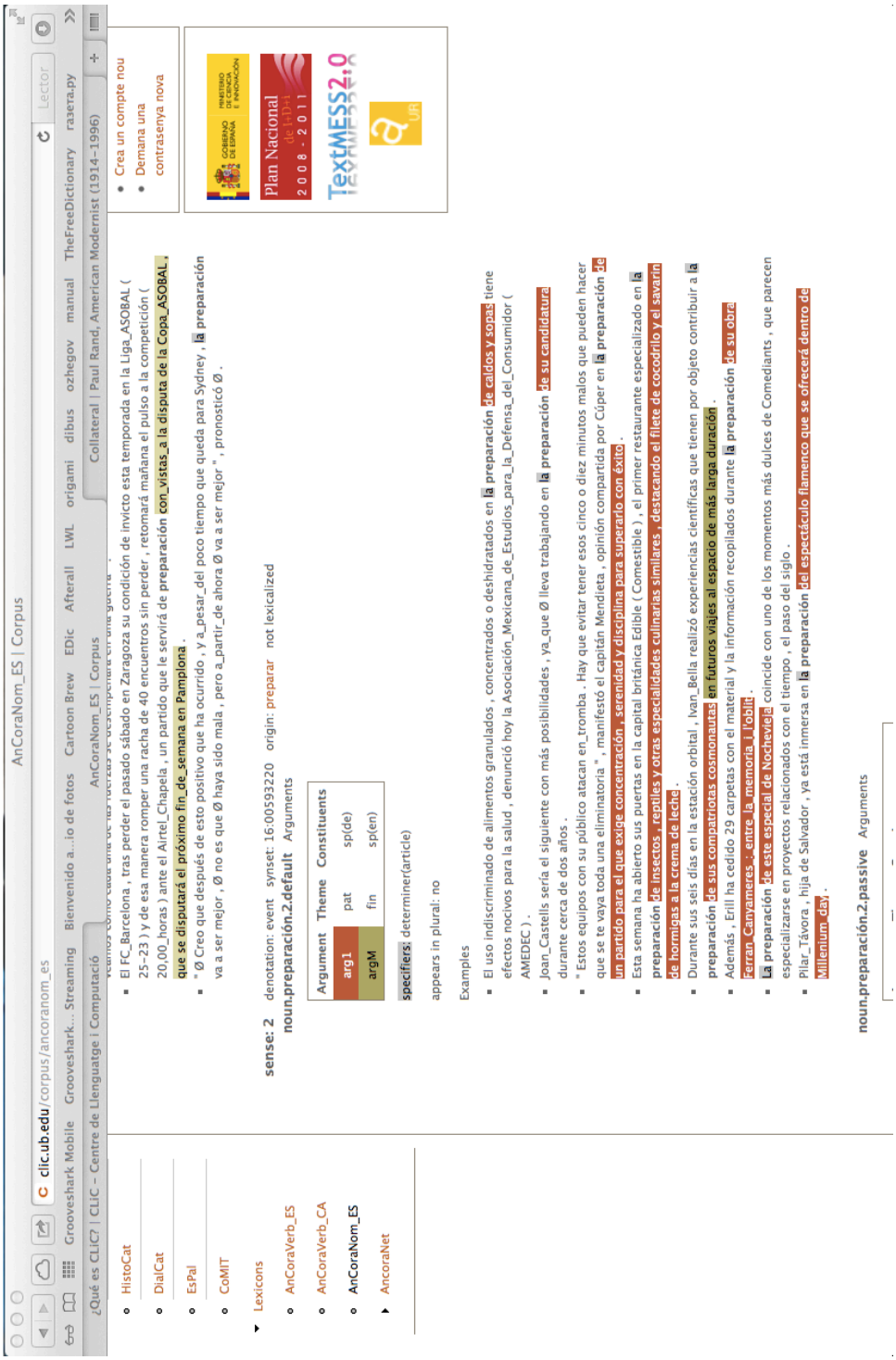

Figura 4: AnCora-Nom analysis of preparación 'preparation'

2.2 Bilingual corpora

In this section we present those corpora of two or more languages. This type of corpus parallelizes the same information in two different languages (or more languages if it is multilingual). This is useful in the comparative study that we have carried out between Russian and Spanish deverbal noun constructions. Bilingual and multilingual corpora used in this study are: *RuSp*, which is not available online, and *UNGAR*, which is available at http://www.uncorpora.org/.

2.2.1 RuSp[10]

RuSp (de Valdivia, Castellví & Taulé, 2013) is a Russian-Spanish parallel corpus mainly composed of literary texts of four different Russian authors from the 19th and 20th centuries: Aleksandr S. Puixkin, Anton P. Txèkhov, Serguei D. Dovlàtov and Varlam T. Xalàmov. *RuSp* contains a total of 710,622 tokens: 304,802 tokens, 52,962 types and 22,702 lemmas belong to the Russian part, while 405,820 tokens, 30,070 types and 15,697 lemmas belong to the Spanish translation (see table 5).

Table 3 shows the different pieces selected from each author (column 2), and the percentage of each piece in the general volume of the corpus (column 3).

Author	Russian and Spanish titles	%	
		Russian	Spanish
A. S. Puixkin	*Pikovaja dama* (La dama de picas)	17.96%	18.09%
	Kapitanskaja dočka (La hija del capitán)		
	Povesti pokojnogo Ivana Petroviča Belkina (Los relatos de Belkin)		
S. D. Dovlatov	*Naši* (Los nuestros)	14.98%	15.84%
	Inostranka (La extrangera)		
A. P. Txekhov	*Moja žizn'* (Mi vida)	25.05%	24.56%
	Čelovek v futljare (El hombre enfundado)		
	Mužiki (Muzhiks)		
	Po delam služby (Por asuntos del servicio)		
	Dama s sobačkoj (La dama del perrito)		
	Ionič (Ionich)		
	Nevesta (La novia)		
	Novaja dača (La nueva dacha)		
V. T. Šalamov	*Kolymskie rasskazy* 1 (Relatos de Kolimá 1)	43.77%	46.65%
	Artist Lopaty 1-2 (El artista de la pala 1-2)		

Table 3: RuSp literary texts

[10] Acknowledgements to the Slavic Philology department of Universitat de Barcelona, specially, to professor Ricard San Vicente.

RuSp is composed mainly of literary texts (99.5%), but also of juridical texts (0.48%), which include penal documents (16.06%), civil documents (38.84%) and administrative documents (43.77%). In table 4, we present the different documents selected from each juridical field (column 2) and the percentage of each document in the general volume of the corpus (column 3).

Type of document	**Russian and Spanish titles**	**%**	
		Russian	**Spanish**
Penal Documents	*Spravok o naličii (otsutstvii) u nikh sudimosti* (Certificado de antecedentes penales)	16.06%	11.16%
Civil documents	*Soglasie* (Consentimiento) *Spravka v tom, čto on dejsvitel'no rabotaet* (Certificado en que se da fe de la ocupación laboral)	38.84%	46.20
Adminstrative documents	*Spravka o dokhodakh fizičeskogo lica* (Certificado de ingresos de personas físicas)	39.25%	69.76%
	Svidítel'stvo o gosudarstvennoj registratsii prava (Certificado sobre el registro de propiedad)		
	Dogovor peredači (Contrato de transmisión)		

Table 4: RuSp juridical documents

These two types of texts were selected since they offered a good number of deverbal noun's structures and translation mismatches.

Russian and Spanish parts have been put in parallel at the paragraph level, that is, a paragraph in Russian corresponds to the same paragraph in Spanish. This task has been carried out manually. In the case of the Russian part, texts were automatically analyzed morphologically and lemmatized by means of *lemmatizer.ru*[11]. In the case of the Spanish part, the lemmatization and the morphological analysis have been carried out automatically by means of Hs_Morfo (Padró *et al.* 2010).

The total number of deverbal nominalizations in RuSp is 2,965 tokens, 1,243 types and 476 different lemmas in the Russian part and 7,776 tokens, 857 types and 674 different lemmas in the Spanish part. All this corpus is in excel format. In order to compute the number of deverbal nouns in *RuSp*, we use, in the case of Russian language, a list of deverbal nouns extracted from Ozhegov & Shvedova (1992).

In table 5, we present the number of tokens, types and lemmas (rows) of *RuSp* for both Russian and Spanish taking into account, on the one hand, the general content, and on the other hand, the deverbal nouns (columns). As a general feature of Russian-Spanish translations, Spanish language uses more deverbal nouns than the

[11] The resource is available at: http://lemmatizer.org/

original Russian. Moreover, in general, the translation into Spanish uses more words than the original Russian text.

	General content		Deverbal nouns	
	Russian	**Spanish**	**Russian**	**Spanish**
Tokens	304,802	405,820	2,965	7,776
Types	52,962	30,070	1,243	857
Lemmas	22,702	15,697	476	674

Table 5: Tokens, types and lemmas in RuSp

The parallel corpus Russian-Spanish of the *RNC* has not initially been used in our study since the parallelized languages Russian-Spanish was not consultable at the time of our study. Nowadays, it is already consultable and we use it to contrast our findings.

2.2.1.1 MiniRuSp

*Mini*R*uSp* is a subsample of R*uSp* consisting of 84,375 tokens, which contains 230 different deverbal nouns from which we deeply analyze a subsample consisting in 114 different deverbal nouns in Russian. This means that almost a 20% of the deverbal nouns in *RuSp* are represented in *Mini*R*uSp*. These 114 lemmas include the 40 most frequent deverbal nouns of *RuSp* for which we have extracted from 10 to 15 occurrences depending on the variety of contexts, that means that similar contexts have not been taken into account. The number of analyzed occurrences for the other 74 nominalizations has depended on the number of occurrences in the *RuSp* corpus, since being less frequent, the amount of occurrences is smaller. Moreover, during the analysis we have discarded manually those nouns that are apparently deverbal nouns, but they are not (1).

(1) *tuberkuleznye* ***otdelenija***
TUBERCULOSIS **DEPARTMENTS**
'tuberculosis **departments**'

In example (1), the deverbal noun *otdelenie* 'department', although being derived from a verb, has lexicalized its meaning and it does not denote neither an action nor its result.

In table 6, we present the content in terms of tokens, types and lemmas (rows) for the analyzed subsample for both Russian and Spanish, computing, on the one hand, all the content and, on the other hand, only deverbal nouns (columns).

	General		Deverbal nouns	
	Russian	**Spanish**	**Russian**	**Spanish**
Tokens	40,416	43,959	774	1,057
Types	11,548	8,381	418	366
Lemmas	6,775	5,359	230	332

Table 6: Tokens, types and lemmas in MiniRuSp

MiniRuSp: syntactico-sematic analysis

MiniRuSp has been manually analyzed at a syntactico-semantic level. It has been only tagged the NP which is headed by the deverbal noun. However, to do the analysis the whole context (sentence and fragment) has been taken into account. The NP headed by the deverbal noun has been tagged with the following information: constituents, syntactic function and the argument structure with its corresponding thematic roles. As we have already said the annotation scheme used for *RuSp* has been taken from the *AnCora-Nom* (See chapter 4, section 4.2).

In figure 5, we present an example of the analysis of the argument structure carried out in *MiniRuSp*.

- In the first column, we find the name of the source from which we have extracted the occurrence (*El artista de la pala*);
- in the second column, we find the deverbal noun under analysis (*vpravlenie* 'reposition');
- in the third column, we find the identifier number of the occurrence (7_R);
- in the fourth column, we find the occurrence in which the deverbal noun takes place, that is, the whole paragraph in which the deverbal noun occurs;
- in the fifth column, we find the mapping between the syntactic structure (constituent and function) and the argument structure of the deverbal noun. In the example, the word *vyvikha* 'dislocation' has one NP (NP1) in genitive (GEN) with the syntactic function of a noun complement (NC) which corresponds to the argument (Arg1) with a thematic role of Theme (tem);
- in the sixth column, we find the phrase structure of the deverbal noun (since we take into account the context, here, for instance, the NP headed by the deverbal noun is embedded in a PP);
- and, finally, in the last column we have the denotation type of the deverbal noun (an event, in the exemple).

Source	Noun	ID	Fragment of occurrence	Syntactico-semantic information	Constituent structure	Lexical Denotation
El artista de la pala	ВПРАВЛЕНИЕ	7_R	Но... однажды меня позвали вправить вывих плеча. Врач давал наркоз “рауш”, а я вправлял ногой -по Гиппократову способу. Под пяткой что-то мягко щелкнуло, и плечевая кость вошла на свое место. Я был счастлив. Татьяна Михайловна Ильина, присутствовавшая **при вправлении вывиха**, сказала:	[вывиха] NP1_GEN_NC_Arg1_tem	при [вправлении [вывиха] PP[p[NP[n[NP1[n]]]]]	EVENT
El artista de la pala	ВПРАВЛЕНИЕ	7_S	Pero... una vez me llamaron para colocar en su lugar un hombro dislocado. El médico inyectaba una anestesia “Rausch” y yo recolocaba el hombro con el pie, por el método de Hipócrates. Debajo de la planta de mi pie algo chasqueó suavemente y el hueso del hombro se reintegró en su lugar. Me sentí feliz. Tatiana Mijáilovna Iliná, presente **en la operación**, me dijo:	--	en [la operación] PP[p[NP[spec,n]]]	EVENT

Figure 5: An example of syntacticosemantic annotation of MiniRuSp

MiniRuSp: translation mismatches analysis

MiniRuSp has also been used to analyze translation mismatches (see chapter 5, section 5.1). We have analyzed 500 deverbal nouns contextualized corresponding to 114 different lemmas with their corresponding translations to determine and annotate the different types of linguistic changes taking place.

MiniRuSp has been used to analyze the argument structure and translation mismatches between Russian and Spanish. In table 7, we present the translation mismatches analysis of the deverbal noun *vpravlenie* ‘reposition’. We have marked as (+) or (-) the different linguistic changes according to the presence (+) or absence (-) of the change. The deverbal noun *vpravlenie* ‘reposition’ presents the following mismatches: the determiner change (that is, in the translation we have the determiner article), the explicitation change (that is, Russian instance has one argument explicitly realized, whereas this argument is implicit in Spanish), the lexical change (that is, Spanish translation has chosen a word that it does not correspond directly with the source).

Source	Deverbal noun	ID	Russian example	Spanish example	Determiner change	PoS change	Constituent change	Function change	Order change	Incorporation change
El artista de la pala	ВПРАВЛЕНИЕ	7_R	Но... однажды меня позвали вправить вывих плеча. Врач давал наркоз “рауш”, а я вправлял ногой - по Гиппократову способу. Под пяткой что-то мягко щелкнуло, и плечевая кость вошла на свое место. Я был счастлив. Татьяна Михайловна Ильина, присутствовавшая **при вправлении вывиха,** сказала:	Pero... una vez me llamaron para colocar en su lugar un hombro dislocado. El médico inyectaba una anestesia “Rausch” y yo recolocaba el hombro con el pie, por el método de Hipócrates. Debajo de la planta de mi pie algo chasqueó suavemente y el hueso del hombro se reintegró en su lugar. Me sentí feliz. Tatiana Mijáilovna Iliná, presente **en la operación**, me dijo:	+	-	-	-	-	-
					Explicitation change	Lexical change	Head swapping change	Number change	Coref-anaphoric change	Coref-elision change
					+	+	-	-	-	-

*Table 7: An example of the translation mismatches analysis in Mini*RuSp

In Appendix A, there are the frequencies of the deverbal nouns analyzed both in *RuSp* and in *Mini*RuSp.

2.2.2 UNGAR

UNGAR is a multilingual corpus that contains 2,100 resolutions of the United Nations General Assembly for the following languages: Arabic, Chinese, English, French, Russian and Spanish. Texts has been originally written in English and then translated into the other languages. Each text has been aligned at the level of paragraphs, with just over 74,000 paragraphs in each language. The corpus contains an average of around 3 million tokens for each language, specifically 2,748,898 tokens in Russian and 3,581,566 tokens in Spanish. Rafalovich & Dale (2009) warn that a complete tokenization of the corpus has not been carried out. The corpus is encoded in XML using the Translation Memory eXchange format (TMX), with some of the significant sections and texts segments marked to assist future research. TMX format has been selected as a storage format as it is standard used in Computer–Assisted Translation tools and has a structure that is simple and sufficient for their needs. The corpus is available in three different formats: (1) a machine-friendly version, which contains no newlines or insignificant whitespaces; (2) a human-friendly version, a slightly larger version, which has been pretty-printed to make it easier to review its content or to process it with non-XML tools; and, finally, (3) a

plain TM version: In this version, voting segments, footnotes, symbols and lead markers are removed. (See Rafalovich & Dale, 2009 for more information.)

In this research, this corpus has been used as a linguistic resource to consult and contrast translation mismatches. However, it is important to bear in mind that it is not possible to use this corpus as a resource to compare Russian and Spanish texts as translation parallel since both languages are translations of original English texts.

In figure 6, we present a screenshot of UNGAR where it is possible to see the sentence 'Adopted at the 31st plenary meeting, on December 2000, on the recommendation of the Committee' parallelized to all the languages mention previously. As we see, it is merely parallelized and marked with XML tags (which do not correspond to any linguistic information).

The corpus is available at http://www.uncorpora.org/

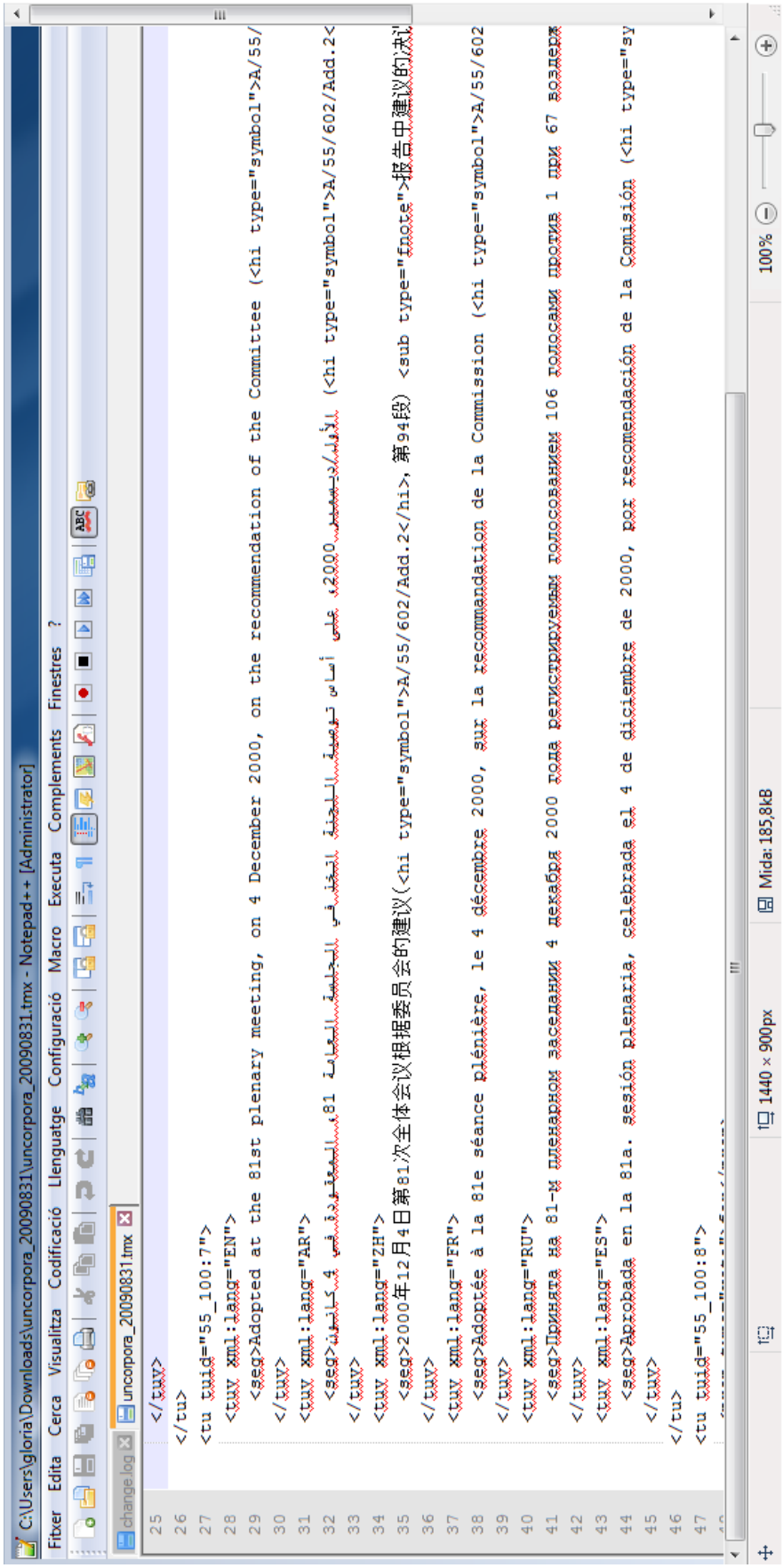

Figure 6: UNGAR Corpus

2.3 Summary

We present in a schematic way the basic resources used, the analyzed sample and the aim of the research.

Linguistic resource	Sample	Aim of the study
The Essex database **(Spencer & Zaretskaya, 2010)**	We took a sample of 296 different deverbal nouns derived from 294 different verbs.	Classification of Russian deverbal nouns in symmetric, neutralized, biaspectual and uniaspectual deverbal nouns. (Section 3.3)
The Russian National Corpus **(Apresjan *et al.*, 2005)**	Out of the sample of 296 deverbal nouns, we took 109 different deverbal nouns and we extracted 323 occurrences from the RNC (between 1 and 3 instances for each noun).	Relationship between aspect of base verb and the the lexical aspect of the deverbal noun (Section 3.4.1):
	Out of the 323 occurrences we took 152 occurrences that have obtained total and partial agreement. For this sample, biaspectual nouns were excluded.	- morphological aspect of the base verb and the lexical denotation of the deverbal noun (Section 3.4.3);
	Out of the 323 occurrences, we took 177 occurrences that have also obtained total and partial agreement, however, here, biaspectual nouns were included.	- lexical aspect of the base verb and the lexical denotation of the deverbal noun (Section 3.5);
	We took 9 different deverbal nouns derived from accomplishment base verbs and we look for 135 occurrences.	- denotative preferences of deverbal nouns derived from accomplishment base verb (Section 3.5.1).

	The previous samples of 177 occurrences and 135 occurrences were put together resulting in a sample of 312 occurrences.	The wide-spread criteria to distinguish between deverbal noun denotation (Section 3.6).
MiniRuSp **(de Valdivia, Castellví & Taulé, 2013)**	We gathered 114 different lemmas of deverbal nouns and we extracted 500 occurrences along with their corresponding translations into Spanish.	Analysis of the argument structure of Russian deverbal noun constructions. (Section 4.3.2)
		Analysis of the translation mismatches between Russian and Spanish deverbal noun constructions. (Section 5.4)

Chapter 3

Morphological and lexical aspect in Russian deverbal nominalizations

Deverbal nominalizations in Russian inherit the presence of the aspectual morphological marks of the base verb from which the nominalization derives. The main focus of this chapter is to analyze to what extent the morphological and the lexical aspect of the base verb determines the lexical denotation of its corresponding nominalization. In our case its denotation as an event, as a result or as a state.

The initial hypothesis of our research is that it is not possible to say that morphological and lexical aspect of the base verb determine the lexical denotation of the deverbal noun, although it seems that both have significant influence on the lexical aspect of the nominal.

In order to examine this hypothesis we proceed as follows. (a) First, we analyze the different types of nominalizations on the basis of the traditional verbal classification into aspectual paired, biaspectual and uniaspectual verbs. The result of this analysis is a deverbal nominalization classification, which enables us to determine from which class of verb the nominalization is derived and whether the nominalization has inherited morphological aspectual marks –either imperfective or perfective – from the corresponding verb. (b) Second, we analyze each type of nominalization in examples from real data in order to establish their denotation or lexical aspect (i.e. event, result or state). The main two goals in doing this analysis are, first, to determine the influence of the morphological aspect of the base verb on the lexical denotation of the nominalization, and, second, to determine the influence of the lexical aspect of the base verb on the lexical denotation of the deverbal noun. The intent of this research is to understand more deeply the relationship established between the nominalization and its base verb.

Finally, we review the most widespread criteria used in the literature to distinguish the denotation of deverbal nouns, such as the expression of the internal argument, the ability to pluralize and the presence of denotative selectors.

This chapter is structured in the following way. In section 3.1, we briefly introduce verbal aspect. In section 3.2, we describe our approach to the morphological and lexical aspect of deverbal nouns. In section 3.3, we present our proposal of classification of deverbal nouns in Russian. In section 3.4, we analyze the relationship between morphological verbal aspect and deverbal lexical denotation. In section 3.5, we analyze the relationship between verbal lexical class and the denotation of the deverbal noun. In section 3.6, we focus on the widely accepted criteria to distinguish between deverbal denotations. In section 3.7, we study the nominalizing suffixes and the lexical denotation of the nominalization. Finally, in section 3.8, we present our conclusions.

3.1 An introduction to verbal aspect

Morphological and lexical aspects are different ways of viewing the internal temporal constituency of a situation (Holt, 1943 (*op cit.* Comrie, 1976)). The first makes reference to a property of a specific verb form, whereas the second makes reference to an inherent property of an eventuality. Therefore, lexical aspect is invariant, while grammatical aspect can be changed according to the whims of the speaker.

3.1.1 Morphological aspect in verbs

Regarding morphological aspect, Russian verbs express by means of two different forms the morphological aspectual opposition between imperfective and perfective[12]. Comrie (1976) claims that, on the one hand, the imperfective looks at the situation from the inside, and it is crucially concerned with the internal structure of the situation, since it can both look backwards to the start of the situation and it can look forwards to the end of the situation. Besides this backward-forward movement, the imperfective meaning is equally appropriate if the situation lasts through all time, without any beginning and without any end. On the other hand, the perfective looks at the situation from outside and as a whole complete unit, without necessarily distinguishing the internal structure of the situation. This means that the imperfective forms denote an ongoing process, while the perfective forms refer to an action either with a culmination point or finished. In Russian, this aspectual opposition can be obtained through changing the accent position (*razrezat'*[PF][13]-*razrezát'*[IPF][14] 'to cut'), with a suppletive form (*skazat'*[PF]-*govorit'*[IPF]'to say') and by affixation (*sdelat'*[PF]-*delat'*[IPF] 'to do'). A more fine-grained distinction can be done between primary and

[12] Comrie (1976) claims that very often the term 'perfective' is confused with the term 'perfect'. By perfective it is meant "a situation in its entirety, without regard to its internal temporal constituency". On the other hand, by perfect it is meant "a past situation which has present relevance, for instance, the present result of a past event (his arm has been broken)."

[13] [PF] stands for perfective verbs.

[14] [IPF] stands for imperfective verbs (both primary and secondary imperfective).

secondary imperfective; and primary and secondary perfective. This distinction is done taking into account the process of derivation in which the verb has been generated. Therefore, a primary imperfective refers to the unprefixed imperfective form (1a), which in turn is the base for generating a perfective form by means of a prefix attachment (1b). A secondary imperfective refers to those verbal forms derived from a prefixed perfective form by means of the attachment of the imperfectivizing suffix (1c). Furthermore, prefix *'po-'* can form a perfective verb with the aspectual meaning 'a little'. This type of perfective is also called 'secondary perfective verb' or 'aktionsart verb'. This prefix can be added either to a primary imperfective (1d) or to a secondary imperfective (1f). We leave aside secondary perfective verbs, since they do not form deverbal nominalizations as has been claimed by Pazelskaya & Tatevosov (2003) and Schoorlemmer (1995).

(1)

a. *pisat'* [IPF]	'to write'
b. *zapisat'* [PF]	'to write down'
c. *zapisyvat'* [IPF]	'to write down'
d. *popisat'* [PF]	'to write a little'
f. *pozapisyat'* [PF]	'to read a little'

Taking into account the morphological aspect, Russian verbs are traditionally classified into three different types: aspectual paired, biaspectual and uniaspectual verbs.

A. Aspectual paired verbs

The majority of Russian verbs are aspectual paired verbs, this means that they are organized in pairs *(delat'* [IPF]*-sdelat'* [PF] 'to do' in 2*),* one member of the pair is imperfective (2a), whereas the other is perfective (2b). Both forms share the same lexical meaning.

(2)

a. Soldaty	*-eto*	*te,*	*o*	*kom*	*my*	*ne*	*znaem,*	*što*
SOLDIERS-	THIS	THOSE	ABOUT	WHOM	WE	NOT	KNOW	WHAT

oni ***delajut***
THEY **DO**

'Soldiers are those about whom we do not know what they **do**.'

b. *Vy* *khotite,* *čtoby* *ja* ***sdelal*** *doklad* *na* *temu*
YOU WANT THAT I **DID** REPORT ABOUT ISSUE
razbitija *rossijsko-meksikanskikh* *ekonomičeskikh* *svjazej?*
EVOLUTION RUSSO-MEXICAN ECONOMIC CONNECTIONS
'Do you want me **to do** a report about the evolution of the connections between Russia and Mexico?'

However, sometimes a perfective verb can have two aspectual pairs (*čitat'*$_{[IPF]}$ – *pročitat'*$_{[PF]}$ – *pročityvat'*$_{[IPF]}$ 'to read' in 3), i.e. a primary imperfective (*čitat'*$_{[IPF]}$-*pro*čitat'$_{[PF]}$ in 3a) and a secondary imperfective (*pročitat'*$_{[IPF]}$ –*pročityvat'*$_{[PF]}$ in 3c).

(3)
a. *Ona* *vsegda* *byla* *gotova* *i* *vsegda* *vsë* ***čitala***$_{[IPF]}$.
SHE ALWAYS WAS READY AND ALWAYS EVERYTHING **READ**
'She always was ready and **read** everything.'

b. *Ja* ***pročital***$_{[PF]}$ *i* *ponjal,* *čto* *ničego* *podobnogo*
I **READ** AND UNDERSTOOD, THAT NOTHING SIMILAR
mne *ran'še* *ne* *popadalos'.*
TO_ME BEFORE NOT HAPPENED
'I **read** and understood that nothing similar had happened to me before.'

c. *Obyčno,* *prežde* *čem* *otpravit'* *delovoe* *pis'mo,* *on*
GENERALLY, BEFORE THAT TO_SEND BUSINESS LETTER, HE
pročityval$_{[IPF]}$ *ego* *Ljudmile* *vslukh.*
READ IT TO_LIUDMILA ALOUD.
'Before sending a business letter, he would often **read** it aloud to Liudmila.'

The verbal forms *čitat'* 'to read', *pročitat'* 'to read' and *pročitat'* 'to read' have the same lexical meaning 'to read'. The difference between the first imperfective (3a) and the perfective (3b), or between the perfective form (3b) and the secondary imperfective form (3c) is purely aspectual. This means that both imperfective forms denote an ongoing process, while the perfective form refers to an action either with a culmination point or finished. However, if the difference between verbal forms is not only aspectual, but lexical too, then the forms do not constitute an aspectual pair (4). In example (4), *čitat'*$_{[IPF]}$ (3a) is not considered the aspectual pair of *perečitat'*$_{[PF]}$ (4a). On the contrary, *perečityvat'*$_{[IPF]}$ (4b) is considered the aspectual pair of *perečitat'*$_{[PF]}$ (4b). This is because between (3a) and (4a) there is a change in the meaning, since there is a new nuance implying repetition 'to read again'; while between (4a) and (4b) there is no change in the meaning.

(4)

a. *So*	*strakhom*	*i*	*otvraščeniem*	*on*	***perečital***$_{[PF]}$	*svoi*
WITH	HORROR	AND	DISGUST	HE	**REREAD**	HIS
dve	*stranicy.*					
TWO	PAGES					

'He **reread** his two pages with horror and disgust.'

b. *Po-moemu,*	*eë*	*nado*	***perečityvat'***$_{[IPF]}$	*periodičeski.*
IN_MY_OPINION,	HER	ONE_OUGHT	**REREAD**	PERIODICALLY

'In my opinion, it has to be **reread** periodically.'

B. Biaspectual verbs

Biaspectual verbs are those that have a unique form with both imperfective and perfective meaning, thus they do not participate in aspectual pairs and the aspectual distinction of these verbs emerges exclusively from the context. We label them '$F_{IPF\ \&\ PF}$' (*ženit'sja* 'to marry' in 5) in order to indicate that the form is biaspectual. In example (5a) the verb is perfective, while in the example (5b) is imperfective.

(5)

a. *Včera*	*on*	*nakonec*	***ženilsja***$_{[PF]}$.
YESTERDAY	HE	FINALLY	**MARRIED**

'Finally, he got **married** yesterday.'

b. *On*	***ženilsja***$_{[IPF]}$	*neskol'ko*	*raz.*
HE	**MARRIED**	SEVERAL	TIMES

'He got **married** several times.'

C. Uniaspectual verbs

Finally, uniaspectual verbs are either perfective or imperfective and, consequently, cannot participate in aspectual pairs. To indicate that the form is uniaspectual we use the label 'F_{IPF}' (*davit'*$_{[IPF]}$ 'to press' in 6) or 'F_{PF}' (*sglazit'*$_{[PF]}$ 'to put the evil eye' in 7).

(6) *Vašington*	*i*	*Brjussel'*	***davit'***$_{[IPF]}$	*na*	*Putina*	*ne*	
WASHINGTON	AND	BRUSSELS	**TO_PRESS**	TO	PUTIN	NOT	
budut.							
WILL							

'Washington and Brussels won't **press** Putin.'

(7) *Ljusja* *stučit,* *čtob* *ne* ***sglazit'*** [PF], *na*
LIUSJA KNOCKS, IN_ORDER NOT **TO_GET_EVIL_EYE**, ON
skamejke.
BENCH
'Liusja knocked on the bench in order not **to prevent the evil eye**.'

3.1.2 Lexical aspect in verbs

As we have already pointed out, lexical aspect is an inherent and an invariant property of an eventuality. Since Vendler (1967), situations described by predicates can be classified between states (such as *know* or *want*), activities (such as *run* and *laugh*), accomplishments (such as *build* and *open*) and achievements (such as *discover* and *arrive*). For authors such as Mourelatos (1978) and Verkuyl (1989, 1993), this classification in four types can be reduced into three types since accomplishments and achievements can be grouped into one group named events. Vendler classifies situations into these four types by means of two features: processuality and telicity. Activities and accomplishments would be qualified as processual since they are constituted by different phases; whereas states and achievements have only one phase then they are not processual. However, the processuality parameter may be problematic. States and achievements are not processual in the same way, that is, states do not have temporal phases, whereas achievements make reference to punctual situations. Comrie (1976) prefers to classify them maintaining the telicity parameter and using the dynamicity parameter instead of the processuality one. Comrie (1976) describes states as static (they continue as before unless changed), and events and processes as dynamic (require a continual input of energy if they are not to come to an end). Events are dynamic situations viewed as a complete whole (perfectively), whereas processes are dynamic situations viewed in progress, from within (imperfectively). Lexical aspect is characterized by semantic properties such as durativity *vs.* punctuality, telicity *vs.* atelicity, and stativity *vs.* dynamicity. Durativity, according to Comrie, refers to the fact that the given situation lasts for a certain period of time; whereas punctuality makes reference to the quality of a situation that does not last in time, that is, one that takes place momentarily. A telic situation has a terminal point; whereas an atelic situation has not such a terminal point, and can be protracted indefinitely or broken off at any point. A dynamic situation involves necessarily a change and requires an effort (that is, an input of energy) to hold it; whereas a static situation does not necessarily involve change and, besides this, to remain in a certain state there is no need of any effort, that is, the state will continue unless something happens to change it.

Morphological and lexical aspects can be interrelated. Marín (2002), for instance, relates the inner aspect (lexical aspect) with the outer aspect (morphological aspect). He postulates that the progressive meaning is compatible only with events and processes, but not with states. However the compatibility with either the event or the process is different, since when it is combined with a process then the meaning is not

changed, whereas when it is combined with an event then the meaning is altered since the event loses its telic character becoming a process. Comrie (1976) postulates that perfectivity and imperfectivity are both compatible with the notion of duration. However, there is a special type of perfective meaning which make reference to situations that have no duration, that is, that are punctual (for instance, 'to reach the summit' or 'to give a slap'). Moreover, he claims that the concept of telicity does not connect a perfective meaning with a telic interpretation and an imperfective meaning with an atelic interpretation.

But according to Paducheva (1996), states and processes are connected with the imperfective meaning; whereas achievements are connected with the perfective meaning. However, in the case of accomplishments the relation is not so straightforward since the aspect depends on whether the target of the situation has been reached or not.

3.2 Morphological and lexical aspect in deverbal nouns

Russian deverbal nouns can preserve morphological aspectual marks from the corresponding base verb. In the case of verbs, those aspectual marks have a grammatical function. In fact, depending on the expression of the morphological aspect, verbs are traditionally classified into aspectual paired (*pisat'/napisat'* 'to write'), biaspectual (*ženit'sja* 'to marry') and uniaspectual (*sglazit'* 'to put the evil eye'). A logical question is then to ask whether or not nominalizations have morphological aspect because of the presence or the absence of such morphological marks. In fact, in other Slavic languages such as Polish, verbal nouns and deverbal nouns are distinguished by the presence of morphological aspect (Komur, 2005). Most authors (Vinogradov, 1972; Schoorlemmer, 1995; Spencer and Zaretskaya, 2010; Zimmermann, 2002; Pazelskaya and Tatevosov, 2003) consider that Russian nominalizations do not have such an aspect, thus aspectual marks inherited by the nominals do not have a grammatical function. Some authors claim that, even if there is not morphological aspect in the nominalizations, some inherited verbal affixes have an influence on the lexical denotation of the nominalization. Vinogradov (1972) postulates that the secondary imperfectivizing suffix *-(y/i)va-* gives a repetitive nuance to the meaning of the nominalization. Schoorlemmer (1995) and Zimmermann (2002) suggest that this *-(y/i)va-* suffix adds some complexity to the nominal giving some unambiguous event character to it. Spencer and Zaretskaya (2010) argue that morphological aspect is not directly reflected in the nominalization, but what is reflected is the semantic interpretation commonly associated with particular grammatical verbal aspects. For instance, in the case of nouns derived from a secondary imperfective what is reflected is the process meaning of their imperfective base verb. Pazelskaya and Tatevosov (2003) do not see such an influence and claim that the lexical aspect of the nominalization depends on the event structure of the corresponding verb and its internal argument. Therefore the inheritance of the

aspectual marks by the nominal poses the question about the influence of such marks on the lexical meaning of the nominalization.

Zalizniak (1977) claims that deverbal nouns preserved a connection with their base verb not only at a morphological level but also at a semantic level. This author claims that deverbal nouns can preserve part of the base verb aspectual meaning. Deverbal nouns, which were derived from the imperfective base form, had a tendency to denote an on-going or iterative process, while nouns, which were derived from the perfective base form, had a tendency to denote a nonrepetitive completed action as well as its result or product. However, this aspectual opposition has disappeared. Deverbal nouns do not inherit the reciprocal or passive mark *–sja*, the opposition of forms such as *proščenie* 'forgiveness' and *prošč̌anie* 'parting' although at the beginning the opposition between them was aspectual, later it was used to mark the voice. The deverbal noun *proščenie* 'forgiveness' is derived from the base verb *prostit'* 'to forgive', while *prošč̌anije* 'departing' is derived from the verbal form with *–sja*, that is, *prostit'sja* 'to say goodbye'. Another example of the disappearance of the aspectual opposition in deverbal nouns is a noun such as *sočinenie* 'composition'. At the beginning, it only denoted a completed non-repeated action and its result; however, nowadays *sočinenie* 'composition' can denote a process.

Roy & Soare (2011) claim that "aspect (either the *Aktionsart* inherited from the verbal base or the built-in aspect inside nominals) determines the count properties of derived nominals, and this again shows that typically verbal information is accessible inside derived nominal". Alexiadou *et al.* (2010) postulate that deverbal nouns' ability to pluralize is the consequence of a competition between Aspect and Number in these nominals. She relates it to Romanian or French, whose deverbal nouns encode the morphological aspect. In Romanian, the infinitive and the supine nominalization encode different aspectual values, that is, telicity and atelicity, respectively. In French, event nominals in *–age* encode imperfectivity, whereas event nouns in *–ée* encode perfectivity.

Following Vendler (1957), lexical aspect traditionally classifies verbs into states, activities, achievements and accomplishments. Nominalizations derived from verbs also have lexical aspect, which can be studied through the property of telicity (Pazelskaya and Tatevosov, 2003) or through the denotative differences such as event, result and state. By an event we mean the action expressed by the corresponding base verb, a result names the concrete or abstract entity related to the action and a state refers the non-dynamic situation named by the corresponding verb. Our proposal is parallel to what is proposed in literature. What we call event nouns corresponds to what authors such as Apresjan (1971), Grimshaw (1990), Pustejovsky (1995), Alexiadou (2001), Eberle *et al.* (2009), Peris & Taulé (2009), Balvet *et al.* (2010) have called complex event nouns, process nouns, event nouns and action nouns. Our result nouns correspond to result or result-object nouns and, finally, state nouns correspond to what in the literature has been called as state-object.

3.3 Classification of deverbal nominalizations

This section presents a deverbal nominalization classification based both on the different verbal types and on the morphological aspect of the corresponding base verb. The main point in proposing a classification of deverbal nominalizations is to see whether or not there is any systematic relationship between morphological aspect of the base verb and the lexical aspect of the deverbal noun. The analysis is based on a sample of 296 different types of nominalizations derived from 294 different verbs extracted from the *Essex Database* (Spencer & Zaretskaya, 2010)[15] (See chapter 2, section 2.1.2). In order to create our sample, we have taken into account from the *Essex Database,* the following information: the type of base verb (paired, biaspectual and uniaspectual) from which the nominalization derives, its corresponding aspectual pair (if there is one), and the nominalizations derived. The selected nominalizations are extracted according to a list that includes the most productive Russian nominalizing suffixes, i.e. the group of suffixes ending in *-ij(e)*[16]*, -k(a), -stv(o)* and masculine and femine null suffix*,* that is, - Ø and - Ø(a), respectively as proposed by Pazelskaya (2009) based on Shvedova (1982: 157-166). In order to provide a balanced sample of the nominalizing suffixes we have extracted 300 nominalizations, from which 296 instances were finally selected. These correspond to 8% of instances in the database containing each of these nominalizing suffixes. The selection of 296 nominalizations includes every productive nominalizing suffix. In the case of the least productive suffixes (such as, *-iš, -išč(e),-otn(ja), -ot(a)*) they are represented in almost its totality since there are few deverbal nouns with these suffixes. In table 1, it is shown the number of suffixes represented in the selection of nominalizations used for the present analysis.

[15] Database's documentation and downloading at: http://privatewww.essex.ac.uk/~spena/res_interests.htm

[16] By nominal suffixes ending in -ij(e) we mean the following nominalizing suffixes: *ni(e), -eni(e), -ani(e), -ti(e), -an'(e), -en'(e), -n'e, -iti(e), -t'e, -vi(e)*. Being the most productive -*eni(e)* and -*ani(e)* with a total of 934 and 796 nominalizations containing this suffix in the database respectively.

Suffix	Number	Suffix	Number
-ij(e)	149	*-išče*	3
-k(a)	30	*-otnja*	3
-stv(o)	29	*-ot(a)*	2
- Ø	22	*-ovnja*	1
- Ø (a)	11	*-izn'*	1
-ok	14	*-ot*	1
-cij(a)	12	*-až*	1
-b(a)	6	*-on*	1
-ež	5	*-in(y)*	1
-yš	3	*-č(a)*	1

Table 1: Nominalizing suffixes

The classification of Russian deverbal nominalizations takes into account the type of verb –aspectual paired, biaspectual and uniaspectual- and the morphological aspect -perfective or imperfective- of the base verb. This classification groups deverbal nominalizations into symmetrical, neutralized, biaspectual and uniaspectual.

A. Symmetric Nominalizations are derived from a paired verb, and are those in which different deverbal nouns are derived from each member of the aspectual paired verb. Consequently, the aspectual morphological opposition of the corresponding base verb is preserved (8 and 9).

(8) *izživat'*$_{[IPF]}$ 'to eliminate' > *izživanie*$_{<--IPF}$[17] 'elimination'
izžit'$_{[PF]}$ 'to eliminate' > *izžitie*$_{<--PF}$ 'elimination'

(9) *darit'*$_{[IPF]}$ 'give a present' > *dar*$_{<--IPF}$ 'gift', *darenie*$_{<--IPF}$ 'donation'[18]
podarit'$_{[PF]}$ 'give a present' > *podarok*$_{<--PF}$ 'gift'

[17] Symbols '$_{<--IPF}$' and '$_{<--PF}$' indicate that the nominalization is derived from either an imperfective or a perfective base verb.
[18] It is worth to note that a verb can generate one or more than one nominalization from each member of its aspectual pair (for instance, from *darit'* (9) two different nominalizations are derived *dar* and *darenie*).

As seen in the examples (8) and (9), in this kind of nominalization each member of the aspectual pair, that is, *izživat'-izžit'* and *darit'-podarit',* generates its own nominalizations. The nominalizations *izživanie, izžitie, podarok* preserve the presence of the aspectual marks, that is, the suffix *-va-* from the (secondary) imperfective (8) and the prefixes *iz-* (8) or *po-* (9) from the perfective respectively. Those three nouns have been derived by means of different nominalizing suffixes, that is, *-nij(e)* (8) and its allomorph *-tij(e)* (8), and *-ok* (9). On the contrary, the nouns *dar* and *darenie* in (9) do not have any morphological aspectual mark. They have been derived from the corresponding imperfective base verb by attaching the nominalizing suffix, a zero morpheme and *-nij(e)*, respectively.

B. Neutralized Nominalizations are also derived from a paired verb, which are characterized by a simplification of the aspectual opposition, because derived nominalizations[19] come from one member of the aspectual pair (10, 11).

(10)	*obživat'*$_{[IPF]}$	'make habitable by living'	> *obživanie*$_{<--IPF}$	'action of the V'
	obžit'$_{[PF]}$	'make habitable by living'	> Ø	
(11)	*kroit'*$_{[IPF]}$	'cut out (a garment)'	>*krojka*$_{<--IPF}$	'action of the V',
			>*kroenie*$_{<--IPF}$	'action of the V'
			>*kroj*$_{<--IPF}$	'action of the V'/'style',
	skroit'$_{[PF]}$	'cut out (a garment)'	> Ø	

The verbs *obživat'-obžit'* and *kroit'-skroit'* generate their corresponding nominalizations from only one member of the aspectual pair. In the case of (10) the presence of the morphological aspectual mark (*-va-*) is inherited by the nominalization from the imperfective verb *obživat'*; whereas in the case of (11) there is not any aspectual mark inherited from the base verb *kroit'*. These nominalizations *kroj, krojka, kroenie* and *obživanie* have been derived by means of different nominalizing suffixes, that is, zero morpheme, *-ka* and *-enij(e)* in (11), and *-nij(e)* in (10), respectively.

C. Biaspectual Nominalizations are generated from a unique biaspectual verbal form. Consequently, there is not aspectual opposition of morphological marks neither in the original verb nor in the nominalization.

(12)	*demoralizovat'*$_{[IPF \& PF]}$	'demoralize'	> *demoralizacija*$_{<--IPF \& PF}$	'demoralitzation'

The verb *demoralizovat'* (12) generates a nominalization from their unique verbal form. The nominalization has been derived by means of the nominalizing suffix *-cij(a).*

[19] One or more than one nominalization.

D. Uniaspectual Nominalizations are generated from a uniaspectual verbal form, thus not showing any aspectual opposition.

(13) *sglazit'*[PF] 'to put the evil eye' > *sglaz*<--PF 'evil eye'

In the example (13), the uniaspectual verb *sglazit'* gives rise to the nominalization *sglaz*, by means of the zero nominalizing suffix.

Table 2 summarizes each type of nominalization and how they are derived. 'V' stands for the type of verb: V_P 'paired verb', V_B 'biaspectual verb', and V_U 'uniaspectual verb'. 'N' refers to the type of nominalization derived, where N_S stands for 'symmetrical nominalization', N_N for 'neutralized nominalization', N_B for 'biaspectual nominalization', and N_U for 'uniaspectual nominalization'. '$N_{[+1]}$' indicates that one or more than one nominalization can be derived.

Type of nominalization	Schema
Symmetric	$V_P: F_{PF} > N_{S[1+]}$ $: F_{IPF} > N_{S[1+]}$
Neutralized	$V_P: F_{PF} > N_{N[1+]}$ $: F_{IPF} > Ø$ $V_P: F_{PF} > Ø$ $: F_{IPF} > N_{N[1+]}$
Biaspectual	$V_B: F_{IPF \& PF} > N_{B[1+]}$
Uniaspectual	$V_U: F_{IPF \text{ or } PF} > N_{U[1+]}$

Table 2: Nominalization classification

The results obtained in the analysis of the 296 nominalizations show the number of nouns in relation to the number of verbs and the morphological aspectual marks inherited by the noun (table 3). Paired verbs in Russian are more common than the biaspectual and uniaspectual verbs, so it is not rare to have more nominalizations coming from paired verbs than coming from biaspectual and uniaspectual.

Type of verb	Type of nominalization	Aspectual mark
Paired (194)	Symmetric (83)	IPF: 46 PF: 37
	Neutralized (102)	IPF: 62 PF: 40
Biaspectual (29)	Biaspectual (32)	---
Uniaspectual (71)[20]	Uniaspectual (79)	IPF: 74 PF: 5
TOTAL: 294	TOTAL: 296	

Table 3: Classification of deverbal nominalizations

As shown in table 3, there is a tendency to derive nominalizations from the imperfective base verb rather than from the perfective base verb. In fact, 182 nominalizations out of the 296 are derived from an imperfective verb, while only 82 of them are derived from a perfective verb. In the case of uniaspectual nominalizations, the difference between those derived from the imperfective verb (74 in total) and those others derived from the perfective verb (5 in total) might be misleading. If we consider the imperfective uniaspectual verbs (67 in total) and the number of perfective uniaspectual verbs (4 in total), the number of nominalizations from both imperfective and perfective is comprehensible. Moreover, in Russian, uniaspectual verbs tend to be imperfective rather than perfective.

3.4 Morphological verbal aspect and deverbal lexical denotation

The proposed classification of Russian deverbal nominalizations provides us with information related both to the verbal type and to the morphological aspect of the base verb to analyze in more detail whether there is a systematic correlation between the aspectual distinction in the base verb and the lexical aspect in the nominalization. In order to see whether there is such a relationship, we have carried out an experiment in which participants had to determine the lexical aspect of a set of nominalizations. By lexical aspect we mean the established distinction between deverbal nouns denoting events (the action named by the corresponding verb base as in 14), results (the concrete or abstract entity resulting from the corresponding verb as in 15) and states (the non-dynamic situation by the corresponding verb as in 16).

[20] Regarding uniaspectual verbs 67 are imperfective verbs and 4 are perfective verbs.

(14) *Ego* ***perevod*** *sonetov* *zanjal* *tri* *goda.*
HIS **TRANSLATION** SONNETS TOOK THREE YEARS
'His **translation** of the sonnets took three years.'

(15) ***Perevod*** *na* *stole* *redaktora.*
TRANSLATION ON TABLE EDITOR
'The **translation** is on the editor's table.'

(16) *No* *eto* *bylo* *nastol'ko* *moim* ***pereživaniem,*** *čto*
BUT THIS WAS SO_MUCH MY **WORRY** THAT
ja *ne* *rešilsja* *sprosit'* *u* *papy.*
I NOT DECIDED TO_ASK TO DADDY
'And my **worry** was so great that I did not make up my mind to ask my father.'

The problem in establishing the lexical denotation of a nominalization emerges in cases like (17), since it is not clear whether the noun in this context denotes the event of the action expressed by the base verb 'to devalue' or the result of this action. Cases like (17) are examples of nominalizations that we annotate as 'unspecified' (since their aspectual reading is not clear enough).

(17) *V* *slučae* *destabilizacii* *miravoj* *ekonomiki* *i*
IN CASE DESTABILIZATION WORLDWIDE ECONOMY AND
obescenenija *dollara* *eti* *bereženija* *mogut* *sil'no*
DEVALUATION DOLLAR THESE SAVINGS CAN HEAVILY
postradat' *libo* *budut* *vovše* *utračeny.*
TO_SUFFER OR WILL IN-GENERAL LOST
'In the case of a worldwide economic destabilization and a **devaluation** of the dollar, these savings may be reduced or lost.'

Briefly, the experiment consisted of the assignment of one of these four possible denotation types (event, result, state and unspecified) to each deverbal nominalization of the selected sample.

3.4.1 The experiment: dataset, subjects and procedure

We selected a sample of 109 nominalizations out of those 296 previously analyzed in section 3. This selection takes into account 54 nominalizations derived from paired verbs, 28 from biaspectual verbs, and 20 from uniaspectual verbs and it preserves a balanced representation of the most productive Russian nominalizing suffixes. These 109 nominalizations are analyzed in examples from real data. We randomly extracted 3 instances for each deverbal noun from the *Russian National Corpus* (RNC, Apresjan

et al. 2005). It resulted in a sample of 323 examples[21] (See chapter 2, section 2.1.1), where each instance includes the sentence in which the nominalization appears as well as the previous and subsequent sentences to provide a wider context to interpret the meaning of the nominalizations. Four Russian native speakers with a background in Linguistics[22] have participated in the test in which they were asked to assign the corresponding denotation type to each nominal instance. To carry out the experiment, participants did not attend to a training process. They were only given brief instructions to complete the task along with unambiguous examples of each denotation type in order to avoid influences on their linguistic intuitions. Moreover, the test was carried out in parallel and participants were required to annotate denotations individually.

3.4.2 Inter-annotator agreement results

Once the test is finished, the inter-annotator agreement is assessed taking into account the observed agreement (Scott, 1955) and the Fleiss' kappa (Fleiss, 1971). As it is expressed in the following formula, Fleiss' kappa measures the agreement between raters by removing from the observed agreement the agreement obtained by chance.

$$\text{Fleiss' kappa} = \frac{\text{Observed agreement} - \text{Expected agreement}}{1 - \text{Expected agreement}}$$

We use the R environment[23] for statistically measuring both the total (i.e. the agreement among the four annotators) and the pairwise agreement (i.e. the agreement between each one of the possible pairs of annotators). In table 4, the agreement percentages obtained are presented. Columns show the result for each pair of annotators (pairwise agreement) and between all the annotators (total agreement). The rows show the observed agreement and kappa coefficient.

Pairwise agreement								**Total agreement**
Annotator pairs	**A-B**	**A-C**	**B-C**	**A-D**	**B-D**	**C-D**	**Average**	**A-B-C-D**
Observed agreement	45.37%	50%	51.85%	52.63%	62.54%	57.28%	53.27%	26.62%
Fleiss' kappa	22.20%	27.00%	24.90%	31.00%	41.70%	30.50%	29.56%	29.80%

Table 4: Results of agreement percentages

[21] In a small number of instances, we could not find 3 examples for each nominalization. For this reason we have obtained 323 examples.

[22] Acknowledgments to the Slavic Philology Department of the University of Barcelona.

[23] It is avalaible at http://www.r-project.org/

Following the interpretations of Fleiss' kappa proposed by Landis & Koch (1977) the inter-annotator agreement (both total and pairwise agreement) reached in the current experiment is a fair agreement. The total agreement obtains 29.80% kappa and the average pairwise agreement reaches 29.56% kappa. Table 5 presents the kappa of pairwise and total agreement (columns) obtained for different denotation types (rows)[24].

	Pairwise agreement							Total agreement
Annotator pair	**A-B**	**A-C**	**B-C**	**A-D**	**B-D**	**C-D**	**Average**	**A-B-C-D**
Event	40.90%	37.80%	36.10%	38.80%	57.90%	37.40%	41.48%	41.70%
Result	7.60%	25.60%	14.50%	28.70%	40.40%	27.10%	23.98%	24.40%
State	27.60%	23.60%	43.90%	28.30%	44.50%	30.60%	33.08%	32.10%
Unspecified	6.20%	5.90%	-1.70%	21.20%	2.30%	9.10%	7.16%	8.90%

Table 5: Kappa coefficient of pairwise and total agreement for denotation types

As in the figures of table 5, the agreement rate between the different denotation types ranges from moderate (41.70% for the event reading) to slight (an 8.90% for the unspecified reading). Regarding the total agreement and the average pairwise agreement, the highest rate corresponds to the event type which is 41.70% and 41.48% respectively, followed by the state type which is 32.10% and 33.08%, and the result type which is 24.40% and 23.98%. The lowest rate belongs to the unspecified type, which is 8.90% and 7.16%. Those readings identified with more agreement are the event and the state denotations, whereas those with a result reading have a lower agreement. Connecting these results with what we have found about disagreement, the highest rate of disagreement is found between the event and the result readings, whereas the lowest rate of disagreement is found between the event and the state readings. This may be due to the fact that boundaries between those readings more closely related to predicative interpretations are better established than those between the action and the product of that action. This result confirms Comrie (1976) who claims that "in practice one finds a large measure of agreement between individuals who are asked to classify situations as static (that is, states) or dynamic (that is, events)". Moreover, the slight agreement in the case of the unspecified reading may be due to the fact that the context usually gives enough information to assign a

[24] Rates of kappa agreement can be higher in the case of total agreement than in the case of pairwise agreement, since the effect of the measure of agreement by chance is more penalizing in the pairwise agreement than in the total agreement. Moreover, in cases of low agreement it is possible to obtain negative kappa values. A kappa value can be negative since it 'goes' from -1 to 1.

specified denotation to the nominalization and also due to the fact that annotators may prefer to assign a reading such as event, result and state rather than an unspecified reading.

Our inter-annotator agreement results are similar to those obtained for Spanish under the same conditions (without specific training). We meant to study as clear as possible instances of deverbal nominalizations, which respond to the speaker's intuition of what a state, an event and a result are. Therefore, it is likely that, after training, Russian annotators would reach the same rate as those for Spanish (65% Fleiss' kappa, Peris *et al.,* 2012), but this higher rate would respond to external judgments and specific training, rather than the coincidence in the intuitive meaning that speakers share.

3.4.3 Morphological aspect of base verb and nominal lexical denotation

Given our fair inter-annotator agreement, we have extracted a subsample to carry out our next analysis to study how the lexical reading of a deverbal noun is determined. This subsample consists of 152 nouns that have obtained total or partial agreement (that is, when three of the four annotators agree), excluding biaspectual nominalizations since they lack aspectual opposition (see section 3). Table 6 presents the number of nominalizations relating the morphological aspect of the base verb and the lexical aspect of the deverbal noun.

	Deverbal noun denotation				
Aspect of the base verb	**Event**	**Result**	**State**	**Unspecified**	**Total**
IF	76	10	18	4	108
PF	10	33	0	1	44
Total	86	43	18	5	152

Table 6: Morphological aspect and lexical denotation

The results in table 6 show that there is a tendency to have more nouns with an event or state reading coming from an imperfective base verb (90.38%[25]), and more nouns with a result reading coming from a perfective base verb (76.74%[26]). These results agree with what was pointed out by Schoorlemmer (1995). It is worthy to note that all nominalizations with a state reading in the experiment come from an imperfective base verb. However, we cannot say that the morphological aspect of the base verb determines the lexical aspect of its corresponding nominalization. In fact, we find nouns with an event reading from both imperfective and perfective base verbs (88.37% and 11.62%, respectively), and nouns with a result reading from both perfective and imperfective base verbs (76.74% and 23.25%, respectively). Example

[25] Out of the total number of nominalizations with an event and a state reading.
[26] Out of the total number of nominalizations with a result reading.

(18) shows a nominalization derived from a perfective base verb (*opisanie* 'description'), which in one context has an event reading (18a) and in the other context has a result reading (18b). Example (19) shows a nominalization derived from an imperfective base verb (*obescenivanie* 'devaluation'), which in one context has an event reading (19a) and in the other context it has a result reading (19b).

(18)

(a) *Avtomatičeskij*	*sintez*	*strukturnogo*	***opisanija***	*konstrukcii.*
AUTOMATIC	SYNTHESIS	STRUCTURAL	**DESCRIPTION**	CONSTRUCTION

'The automatic synthesis of the structural **description** of the construction'

(b) *Iskhodnymi*	*dannymi*	*javljajutsja:*	*sistema*	*semantičeskik*	*(smyslovykh)*
SOURCE	DATA	ARE	SYSTEM	SEMANTIC	(SEMANTIC)

opisanij	*svojstv*	*i*	*osobennostej*	*klassa*	*izdelij*
DESCRIPTIONS	PROPERTIES	AND	SPECIAL	CLASS	PRODUCT

i	*formalizovannoe*	*tekhničeskoe*	*zadanie.*
AND	FORMALIZED	TECHNICAL	PRODUCTION

'The source data are a system of semantic **descriptions** of the properties and particularities of the product's class and the formalized technical production.'

(19)

(a) *Každyj*	*posledujuščij*	*etap*	*stroitsja*	*na*	*obolganii,*
EACH	POSTERIOR	PERIOD	BUILD-ITSELF	ON	SLANDERING,

ogluplenii	*i*	***obescenivanii***	*perioda*	*predyduščego.*
DISTORTION	AND	**DEVALUATION**	PERIOD	PREVIOUS

'Each posterior period is built on the slandering, distortion and **devaluation** of the previous period.'

(b) *Polnaja*	*dostupnost'*	*informacii,*	*v*	*osobennosti*	*rasprostranennaja*
FULL	AVAILABILITY	INFORMATION,	IN	SPECIAL	EXTENDED

na	*proizvedenija*	*iskustva,*	*nesët*	*v*	*sebe*	*opasnost'*
ON	CREATION	ART,	BRING	IN	ITSELF	DANGER

ikh	***obescenivanija.***
THEIR	**DEVALUATION**

'Complete availability of information, which is especially common in works of art, entails the risk of their **devaluation**.'

Therefore, it is not possible to draw a systematic correlation between the morphological aspect of the base verb and the lexical aspect of a nominalization.

Moreover, if we look at the difference between nominalizations derived from a primary imperfective and nominalizations derived from a secondary imperfective, we observe that in both kinds of imperfective there are nominalizations that denote events, results and states (or an unspecified reading). This fact suggests that the different imperfective types are not specialized in any special denotation. So, in contrast to what Schoorlemmer (1995:313) and Zimmermann (2002) claim, nominalizations derived from secondary imperfective verbs do not always denote an event (we find 8 nominalizations out of 46 nominalizations derived from secondary imperfectives which denote a result or a state). Nevertheless, it can be seen that the morphological aspect of the base verb has an influence on the lexical denotation of the deverbal noun.

3.5 Verbal lexical class and denotation of the nominalization

Having seen the influence of the morphological aspect of the base verb on the derived nominalization, we have also analyzed wether there is any relationship between the lexical class of the verb (state, activity, achievement and accomplishment) and the lexical denotation of the deverbal noun. The initial hypothesis, following Picallo (1999), Alexiadou (2001), Peris & Taulé (2009), Jezek & Melloni (2009) and Fábregas & Marín (2011) is that the aspectual class of the base verb can determine the lexical denotation of the nominalization. According to these authors, nouns derived from event verbs (accomplishments and achievements) tend to have an event or a result reading, whereas nouns derived from activity and state verbs tend to denote results. Fábregas & Marín (2011) claim that "the question whether a nominalization denotes an event, a state or is ambiguous between the two readings depends on the *Aktionsart* of the predicate, not on the properties of the nominalizer". This suggests, for instance, that the availability of a state nominalization depends on the existence of a state subevent in the verbal base.

In Italian, those correspondences between the *Aktionsart* and the deverbal nouns' denotation are not exactly the same. Aligned with what has been posed before, Jezek & Melloni (2009) claim that causatives and other accomplishments are optimal candidates for yielding ambiguous nouns between the event and the result reading. These authors claim that a deverbal noun derived from an activity base verb can have an event interpretation. This is explained by the fact that activity verbs do not have, in their Event Structure, a state subevent. The state subevent is the responsible for yielding a result reading. However, nouns derived from activity base forms can refer to concrete or abstract objects such as agents in a construction such as 'the administration of the company' or locations 'the US administration' (Jezek & Melloni, 2009).

To analyze whether there is a relationship we added to the previous sample of 152 nouns, which had obtained total or partial agreement, those biaspectual nouns that were excluded in the first sample. Biaspectual nouns were not relevant in the previous analysis since they do not show morphological aspectual opposition.

However, in the current analysis biaspectual nouns are included. This makes a sample of 177 occurrences whose analysis is presented in table 7.

	Lexical denotation		
Lexical aspect of base verb	**Event**	**Result**	**State**
State	3	0	9
Activity	42	10	5
Accomplishment	44	1	0
Achievement	21	41	1

Table 7: Lexical aspect of a deverbal noun and its base verb

The results obtained show clear tendencies that connect denotative readings with the lexical aspectual class of the base verb. In Russian, state verbs tend to generate nouns with a state reading. Achievement verbs can generate nouns with both denotative readings (event and result), as in English or Spanish. Regarding nouns derived from activities, they can also have the two readings (with a special preference to denote events), contrary to what have been postulated for English (Alexiadou, 2001) or Spanish (Peris & Taulé, 2009), but similar to the Italian findings (Jezek & Melloni, 2011). Russian nouns derived from accomplishment base forms can potentially denote results and events, as in Spanish and English, but in our sample this type of nominalization denotes almost exclusively events. This is because the double denotation does not necessarily imply both readings. In order to study this fact in more detail we have enlarged the sample with deverbal nouns derived from accomplishment base verbs, which have the possibility to denote both events and results. To carry out the analysis of these deverbal nouns we have proceeded as follows (section 3.5.1).

3.5.1 Extension1: Enlarging the sample

In order to study the denotative preferences of nouns derived from accomplishments, we have gathered 135 occurrences of 9 different deverbal nouns extracted from the RNC. The analysis consisted of the assignment of the denotative reading of each deverbal noun in its context by an annotator supervised by external experts. The results obtained show that deverbal nouns derived from accomplishment base verbs can denote both events and results. In table 8, we present the denotative tendencies of each deverbal noun under analysis.

Deverbal noun	Event	Result	Unspecified
Razryv 'rupture'	8	7	0
Formirovanie 'formation'	12	2	1
Obslužívanie 'service'	9	6	0
Upravlenie 'direction'	8	7	0
Zakaz 'order'	0	15	0
Sozdanie 'creation'	15	0	0
Izdanie 'publication'	1	14	0
Rešenie 'decision'	3	12	0
Ismenenie 'change'	1	13	1
TOTAL	59	75	2

Table 8: Nominalizations from accomplishment verbs

As shown in table 8, 3 out of 9 deverbal nouns have a balanced tendency to denote either a result or an event (*razryv* 'rupture', *obslužívanie* 'service', *upravlenie* 'direction' and *ismenenie* 'change'), while the rest show preferences to denote one of the two possible denotative readings, an event (*formirovanie* 'formation' and *sozdanie* 'creation') or a result (*zakaz* 'order', *izdanie* 'publication' and *rešenie* 'decision').

Therefore, deverbal nouns derived from accomplishment base verbs can denote events and results, but, despite this possibility, most of them seem to be specialized in one particular reading. To analyze this observation in more detail, we conducted a new experiment to see how a noun with a 'non-prefered' reading is translated. For instance, which will be the option of the translator in a context in which the deverbal noun *formación* 'formation' appears as a result: will he use the deverbal noun *formirovanie* 'formation' with a result reading (which is a 'non-preferred' reading) or will he use a different noun, verb or even another construction? And in this second case, what motivates this election?

3.5.2 Extension2: Translation test

In order to explain the preference of deverbal nouns derived from accomplishment base verbs for one particular denotation (event or result), we have prepared a second experiment, with the aim to translate into Russian 54 examples of 9 Spanish deverbal nouns derived from accomplishments with a clear either event or result denotation. For each of these 9 deverbal nouns, we extracted 6 occurrences from the Spanish lexicon *AnCora-Nom* (Peris & Taulé, 2011): three had an eventive reading and the other three had a result reading. The Spanish examples were translated to Russian to analyze what happens with those nouns in contexts which clearly have their 'non-preferred' reading.

The results obtained were similar to the prior experiment: it seems that some nouns are specialized in one particular reading. We also observed the tendency to select a word, which is not the deverbal noun under analysis. For instance, in the case of the

Spanish *formación* 'formation' with a resultative reading is not translated by the corresponding Russian *obrazovanie* 'formation' with a result reading. The words picked up can be deverbal and non-deverbal nouns (which in any case were those under study), and verbs (both in finite and infinite forms).

The election of a different word can be explained by the following facts:

a. There is another more frequent and usual word to express one of these two possible denotations. This most of the times is determined by the context in which the nominalization takes place:

(i) Most of the times the internal argument of the deverbal noun denotes the outcome of the action, this internal argument alone can be used as an alternative way to denote the result reading of the deverbal noun. In these cases, the deverbal noun is specialized in the event reading and the internal argument is specialized in the result reading. In examples (20b. ii) and (21b.ii), *zdanie* 'building' and *karies* 'caries' are used to denote the result reading, whereas *obrazovanie kariesa* 'caries formation' (21b.i) and *strojka Velikoj Steny* 'construction of the Great Wall' (20b.i) are used to denote the event reading. In these cases, the deverbal noun (*obrazovanie* 'formation' and *strojka* 'construction') focuses on the event, whereas the non-deverbal noun (*karies* 'caries' and *Velikaja Stena 'Great Wall'*) expresses the outcome of the action, which is equivalent to the internal argument of the verbal or deverbal construction. These nouns express more clearly the result denotation and they are a mechanism that the language uses to reduce denotative ambiguity.

(20)

(i) *Soldat,*	*mobilizovannykh*	*na*	***strojku***	*Velikoj*	*Steny*
SOLDIERS	MOBILIZED	IN	**CONSTRUCTION**	GREAT	WALL

'Soldiers mobilized for the **construction** of the Great Wall.'

(ii) *Pered*	***zdaniem***	*muzeja*	*stojal*	*bronzovyi*	*Puškin*
BEFORE	**BUILDING**	MUSEUM	STOOD	BRONZE	PUSHKIN

'In front of the museum's **building** the bronze Pushkin stood.'

(21)

(i) *V*	*kakom*	*vozraste*	*naibolee*	*velik*	*risk*	***obrazovanija***	*kariesa?*
IN	WHICH	AGE	MORE	BIG	RISK	**FORMATION**	CARIES

'At which age there are bigger risks of caries **formation**?'

(ii) *lečenie*	***kariesa***
TREATMENT	**CARIES**

'**caries** treatment'

The use of *zdanie* 'building' is better for the result reading, whereas *stroenie/strojka zdania* 'construction of a building' is more frequently use for the event reading. The same would happen in the case of Spanish, where a sentence such as 'la construcción del edificio es sólida' would be ambiguous between the event and the result reading, while constructions such as *la construcción es sólida* 'the construction is solid' and *el edificio es sólido* 'the building is solid' has an unambiguous reading as a result. As in the Spanish example, it is also possible in Russian to express the result reading using the deverbal noun only, since it can have the internal argument incorporated (22).

(22) *Sejčas* ***strojka*** *zakonservirovana*
NOW **CONSTRUCTION** PRESERVED
'Now the **construction** is preserved.'

(ii) In some occasions, the Spanish noun has been translated by a verb in order to denote the event reading (23) instead of its corresponding Russian deverbal noun. This also would respond to the natural tendency of a language to avoid ambiguity: the use of a verb instead of a noun to express an event is preferred.

(23) *Mne* *poručili* ***opublikovat'*** *biografiju* *Stiva* *Džobsa*
TO_ME COMMISSION **TO_PUBLISH** BIOGRAPHY STEVE JOBS
'I was commissioned **to publish** the biography of Steve Jobs'

If we change the infinitive *opublikovat'* 'to publish' for its corresponding deverbal noun *opublikovanie* 'publication', the construction would be ambiguous between a result and an event reading.

(iii) The existence of two deverbal nouns with the same meaning where one is specialized in denoting events and the other is specialized in denoting results. In examples (24) and (25), we observe that the deverbal noun *formirovanie* 'formation' is specialized in denoting events, whereas *obrazovanie* 'formation' is specialized in denoting results. In example (25), the fact of being in plural reinforces the result reading.

(24) *Povlijajut* *na* ***formirovanie*** *pravitel'stva*
INFLUENCE ON **FORMATION** GOVERNMENT
'They influence on the **formation** of the government.'

(25) *Vozmožno* *zafiksirovat'* *gornye* ***obrazovania***
POSSIBLY TO_DETECT MOUNTAINOUS **FORMATIONS**
'It is possible to detect mountainous **formations**.'

This specialization can be determined by the context of occurrence and the frequency of use. For instance, *formirovanie pravitel'stva* 'formation of the government' is more frequent than the combination *obrazovanie pravitel'stva* 'formation of the government'. As we will study in more detail in section 3.6.4, words that are in combination with the deverbal noun, usually determine the election of the deverbal noun. Moreover, a deverbal noun wiith a high frequency in the language has more chances to be selected than another with a low frequency, independently of their reading.

Words with a high total frequency[27] usually have general meanings, while words with a low frequency are more specialized. *Obrazovanie* 'formation' has 30,078 coincidences in the RNC and 379,000,000 of coincidences in Google, while *formirovanie* 'formation' has 9,992 coincidences in the RNC and 97,600,000 in Google coincidences. These figures can be explained by the fact that deverbal nouns generated from Latin roots tend to have low frequencies and a higher register; whereas deverbal nouns generated from Slavic roots tend to have higher frequencies and standard register.

(26) *Organizacija* *nezakonnogo* *vooružennogo* ***formirovanija***
ORGANIZATION ILLEGAL ARMED **FORMATION**
'the organization of an illegal armed **formation**'

(27) *Sistema* *ne* *sposobna* *raspoznat'* *eto* ***novoobrazovanie***
SYSTEM NOT ABLE TO_RECOGNIZE THIS **NEW_FORMATION**
'The system is not able to recognize this **new formation**.'

From what we have presented above, we can make the following observations:

The lexical aspect of the base verb (as we have seen in the case of morphological aspect) has influence on the lexical denotation of Russian deverbal nouns. Therefore, nouns derived from states denote states; nouns derived from activities tend to denote mainly events; nouns derived from achievements tend to denote results and, finally, nouns derived from accomplishments tend to denote events. However, as in the case of morphological aspect, these correspondences are not systematically held. A noun such as *strojka* 'construction' can denote the event and the noun *stroenie* 'construction'

[27] By total frequency we mean the frequency that a particular word has in the language without taking into account the context where it occurs.

can denote the result. Both deverbal nouns have been derived from the accomplishment base verb *stroit'* 'to build'.

Regarding nouns derived from accomplishments, we have observed that they can express both events and results, but they can be specialized in one of the particular denotations. This can be explained by the tendency of any language to avoid ambiguity. Therefore, the tendency is to use words that express clearly one denotation. Then the internal argument of the deverbal noun is used to denote the result reading and a verb is used to denote the event reading. Moreover, sometimes a same meaning can be expressed by means of two different deverbal nouns. However, one deverbal noun is specialized in an event reading, whereas the other is specialized in the result reading and this may respond to the context surrounding the deverbal noun and the frequency of the deverbal noun in the language.

Therefore, all these factors can determine the tendencies of deverbal nouns for denoting event, states or results. In the following chapter we study the influence of other factors, which according to many authors can be considered as criteria to distinguish between denotations.

3.6 Analysis of the criteria to distinguish among deverbal denotations

As we have already seen, morphological and lexical aspects influence the lexical denotation of the deverbal noun but not systematically. However, morphological and lexical aspect can be considered as useful criteria to distinguish between denotative readings. In this section, we analyze other criteria, very widespread in literature on nominalizations, to distinguish between deverbal denotations. We analyze whether these criteria also work for Russian deverbal nouns. Among them, we are going to focus on those concerned with the elements inside the NP headed by the nominalization, concretely, with the ability to pluralize (see section 3.6.1), with the expression of the internal argument (Grimshaw, 1990 for English; Schoorlemmer, 1995 for Russian; Peris & Taulé, 2009 for Spanish) (see section 3.6.2); and with the specifiers (Grimshaw, 1990 for English and Peris & Taulé, 2009 for Spanish) (see section 3.6.3). And, finally, in section 3.6.4, we study selectors, that is, words that favour one or another reading.

To analyze these criteria we have taken into account a sample of 312 nominalizations belonging to the previous subsamples. This sample consists of the subsample extracted to analyze the relation between morphological aspect of the base verb and the denotation of the deverbal noun with the addition of biaspectual nominalizations (177) and of the subsample of 135 examples gathered to study the influence of the lexical class of the base verb with the denotation of the deverbal noun.

3.6.1 The ability to pluralize

According to Grimshaw (1990) and Peris & Taulé (2009), event deverbal nouns tend to appear in singular (28), whereas result nouns can appear both in singular and plural (29 and 30). For instance:

(28) The **assignment** of difficult problems always causes problems.
(29) The **assignment** was long.
(30) The **assignments** of the problems took a long time.

[Grimshaw, 1990]

However, Brandtner & Heusinger (2009) claim that plurality does not only select a result reading, since in examples (31) and (32), despite being in plural, deverbal nouns denote an event.

(31) *Die wiederholten* ***Messungen*** *belegen, dass keine Besserung eingetreten ist.*
'The repeated **measurements** show that there hasn't been an improvement.'

[Brandtner & Heusinger, 2009]

(32) The many destructions of Constantinople across history.

[Fábregas and Marín, in press]

Fábregas & Marín (2011) punctualize that the option to pluralize "is clearly available to nominalizations denoting objects and to nouns denoting events, provided that they are telic and interpreted as ordered in a temporal succession as in example (32)". They add that "in contrast, state nominalizations and state nouns reject the plural form. When the nominalization (or the noun) allows for a plural form, the state reading disappears and emerges the event reading or the reading of the deverbal noun as a participant in the event". By participants in the event, Fábregas and Marín (2011) mean what Pesetsky (1995) names 'target of emotion' and 'causer of emotion'. The 'target of the emotion' is the object towards which a particular psychological state is directed, as in *mis amores* 'my beloved ones'. The 'causer of the emotion' is the entity that triggers the state as in *las distracciones de los niños también distraían a los padres* 'children's hobbies also distracted their parent'.

In Russian a similar pattern seems to take place. In table 9 we present the results obtained in the analysis of this criteria.

Number	Event	Result	State	Unspecified
Singular	173	79	15	1
Plural	7	55	1	1

Table 9: Number and denotations

In the analyzed sample, we have found the same tendency in Russian as in Grimshaw (1990) for English and Peris & Taulé (2009) for Spanish. Regarding event nouns, 4% of nouns appears in plural (33) while 96% appears in singular (34). Regarding result nouns, 41% appears in plural (35) and the other 59% appears in singular (36). And regarding state nouns, 6% appears in plural while 94% appears in singular. Therefore, in Russian nouns denoting an event or a state tend to be expressed in singular, while nouns denoting results are not restricted in this respect.

(33) *V* *geroine* *Štain* *pokazal* *fisičeskij* *itog*
IN HEROIN SHTAIN SHOW PHYSICAL RESULT
nravstvennykh ***razryvov***$_{PL}$
MORAL **RUPTURE**
'In the heroin, Shtain showed the physical result of the moral **ruptures**.'

(34) *Igra* *dejatel'nost',* *v* *kotoroj* *proiskhodit* ***izživanije***$_{SG}$
PLAYING ACTIVITY IN WHICH TAKE_PLACE **ELIMINATION**
detskogo *egocentrizma.*
CHILDISH EGOCENTRICITY
'Playing is an activity where the **elimination** of childish egocentricity takes place.'

(35) *Vypuskniki* *pretendujut* *na* ***zarabotki***$_{PL}$ *vyše*
GRADUATE PRETEND TO **EARNINGS** HIGHER
10 *tysjač*
10 THOUSAND
'Graduates want **earnings** higher than 10 thousand'

(36) *Pokaz* *fil'ma* *nastojaščij* *paskhal'nyj* **podarok**$_{SG}$
SHOWING FILM REAL EASTER **GIFT**
'The sreening of the film is an real Easter **gift**.'

3.6.2 The realization of the internal argument

According to Grimshaw (1990) for English, Schoorlemmer (1995) for Russian and Peris & Taulé (2009) for Spanish, nouns denoting an event tend to express the internal argument of their base verb (37), whereas nouns denoting a result tend not to express it (38).

(37) [28] [***Razrušenie*** [*Saraeva*] NP-NC-Arg1-pat]NP *nas* *porazilo*
DESTRUCTION SARAJEVO US STRUCK
'The **destruction** of Sarajevo struck us.'

(38) [***Razrušenie***]NP *nas* *porazilo*
DESTRUCTION US STRUCK
'The **destruction** struck us.'

Regarding these tendencies, the results of the analysis are presented in table 10.

Internal argument	Event	Result	State	Unspecified
Yes	86	39	3	2
No	30	77	4	0

Table 10: The need to express the internal argument

These figures show the existence of the tendencies mentioned above, that is, deverbal nouns denoting events tend to appear with their internal argument (75%), whereas nouns with a result reading tend to appear without their internal argument (67%). However, the opposite situation is also possible. In (39) and (40), the deverbal nouns *likvidacija* 'liquidation' and *razryv* 'rupture' denote events. The deverbal noun *likvidacija* 'liquidation' in (39) expresses its internal argument, whereas *razryv* 'rupture' in (40) does not.

(39) *Ustranenie,* *izžitie,* [***likvidacija*** [*tekh* *nepravil'nykh*
REMOVAL ELIMINATION **LIQUIDATION** THOSE WRONG
otnošenij] NP-NC-Arg1-tem]NP
RELATIONS
'the removal, elimination and **liquidation** of those bad relations'

[28] Examples (37) and (38) are extracted by Schoorlemmer (1995).

(40) [***Razryv***]$_{NP}$ *dlilsja* *pjatnadsat'* *let*
RUPTURE LASTED FIFTEEN YEAR
'The **rupture** lasted fifteen years.'

Regarding results, in (41) the deverbal noun expresses its internal argument, while in example (42) it does not.

(41) *Osložnenija* *i* [***razryvy*** [*ekonomičeskikh* *otnošenij*]$_{NP\text{-}NC\text{-}Arg1\text{-}tem}$]$_{NP}$
COMPLICATIONS AND **RUPTURES** ECONOMIC RELATIONS
privodjat *k* *protivostojanijam*
BRING TO OPPOSITION
'The complications and **ruptures** of economic relations cause opposition.'

(42) *Tekhničeskoje* [***obsluživanie***]$_{NP}$ *vkhodit* *v* *tarify* *na* *gaz!*
TECHNICAL **SERVICE** ENTER INTO BILL ON GAS
'The technical **service** is included in the gas bill.'

In spite of the low number of deverbal nouns with a state reading in our sample, it seems that the same situation is held: we found examples of nominalizations with or without internal argument (43, 44, respectively). For instance:

(43) *alkanie* *i* [***zhažda*** [*bespredel'noj* *božestvennoj* *istiny*]$_{NP\text{-}NC\text{-}Arg1\text{-}tem}$]$_{NP}$
HUNGER AND **THIRST** INFINITE DIVINE TRUTH
'**thirst** and hunger of the infinite divine truth'

(44) *otnosit'sja* *k* *R. A.* *s* [***nenavist'ju*** *i* ***prezrenije***]$_{NP}$
TO_TREAT TO R.A. WITH **HATE** AND **CONTEMPT**
'to treat the R. A. with **hate** and **contempt**'

Regarding those nouns with an event reading which express their internal argument, we have observed that they have a great preference to express their internal argument by means of a noun phrase in genitive (45).

(45) *v* *slučaje* [***razryva*** [*dogovora*]$_{NP\text{-}GEN\ –Arg1\text{-}tem}$]$_{NP}$ *o* *družbe*
IN CASE **BREAK** AGREEMENT ON FRIENDSHIP
'in the case of **breaking** the agreement on friendship'

However, they can also express their internal argument by other means such as a

nominal phrase in instrumental (46), a possessive pronoun (47), an adjective (48) or a prepositional phrase (49). For instance:

(46) [***upravlenie*** [*složnymi* *sistemami*] NP-INS-Arg1-tem]NP
ADMINISTRATION COMPLEX SYSTEM
'the **administration** of complex systems'

(47) [[*ikh*]Poss-Spec -Arg1-tem ***razryva***]NP
THEIR **RUPTURE**
'their **rupture**'

(48) [[*konstruktorskikh* *i* *dizajnerskikh*] AP-NC- Arg1-tem ***rešenij***]NP
CONSTRUCTION AND DESIGN **DECISIONS**
'construction and design **decisions**'

(49) *v* *slučaje* [[*polnogo*] ***razryva*** [*s* *sovetskim* *prošlym*]PP-NC-Arg1-tem]NP
IN CASE COMPLETE **RUPTURE** WITH SOVIET PAST
'in case of complete **rupture** with the soviet past'

Regarding nouns with an event reading, which appear without their internal argument, they can express implicitly the internal argument, that is, the argument can be expressed outside the NP headed by the deverbal noun (50) and (51).

(50) *Recepty* *potrebujut* *nekotorogo* *vremeni* *dlja* *podgotovki*
RECIPES WILL_DEMAND SOME TIME FOR PREPARATION
ingredientov, NP-GEN-NC-Arg1-tem *v* *častnosti,* [***narezki*** *i* ***šinkovki***]NP
INGREDIENTS, IN PARTICULAR, **CUTTING** AND **CHOPPING**.
'Recipes take some time to prepare the ingredients, specially, to cut and chop.'

(51) *Vystupaet* *slonovja* *truppa*NP-NC-Arg1-pat *pod* [***upravleniem*** [*Jurija*
PERFORMS ELEPHANT TROUPE UNDER **DIRECTION** JURI
Durova]NP-NC-Arg0-agt]NP
DUROV
'The elephant troupe performs under the **direction** of Juri Durov.'

In example (50), the internal argument *ingredientov* 'ingredients' is expressed explicitly only for the first deverbal noun, which is *podgotovki* 'preparation', however, in the

case of the other deverbal nouns *narezki* 'cutting' and *šinkovki* 'chopping', this internal argument is implicitly realized. In example (51), the internal argument is realized implicitly outside the NP headed by the deverbal noun, concretely as the subject of the main sentence.

Regarding nouns with a result reading, they can express their internal argument by means of a noun phrase in genitive (52), instrumental (53), prenominal adjective (54), prenominal possessive determiner (55) and prepositional phrase (56).

(52) *V literature vstrečaetsja* [***opisanie*** [*sniženija*
IN LITERATURE IS_FOUND **DESCRIPTION** DECREASE
aktivnosti]$_{NP\_GEN\text{-}NC\text{-}Arg1\text{-}tem}$]$_{NP}$
ACTIVITY
'In the literature, there is a **description** about the decrease of the activity.'

(53) *Narušenie* [***upravlenija*** [*impul'sami*]$_{NP\_INS\text{-}NC\text{-}Arg1\text{-}tem}$]$_{NP}$
DISFUNCTION **CONTROL** IMPULSES
'The disfunction of the **control** of the impulses'

(54) *Pravitel'stvo ubeličilo* [[*bjudžetnyje*] $_{AP\text{-}NC\text{-}Arg1\text{-}tem}$ ***assignovanija***
GOVERNMENT INCREASED BUDGET **ASSIGNATIONS**
[*na khimčiskoje razoruženije*]$_{PP}$]$_{NP}$ *s 500*
FOR CHEMICAL DISARMAMENT FROM 500
'Government has increased the budget's **assignation** for chemical disarmament above 500.'

(55) *opasnost'* [[*ikh*]$_{Poss\text{-}Spec\text{-}Arg1\text{-}tem}$ ***obescenivanija***]$_{NP}$
DANGER THEIR **DEVALUATION**
'the danger of their **devaluation**'

(56) [***razryv*** [*s tradicionnoj filosofskoj dikhotomijej ob'jekta*
RUPTURE WITH TRADITIONAL PHILOSOPHICAL DICHOTOMY OBJECT
i sub'jekta.]$_{PP\text{-}NC\text{-}Arg1\text{-}tem}$]$_{NP}$
AND SUBJECT.
'the **rupture** with the traditional philosophical dichotomy between the object and the subject'

Regarding those nouns with a result reading, which do not express their internal argument, they can have an incorporated internal argument, that is, the argument is realized inside the deverbal noun root as in (57) and (58).

(57) *Naši* ***issledovanija*** *pokazyvajut* (…)
OUR **INVESTIGATIONS** SHOW
'Our **investigations** show (…)'

(58) *finansiruja* *eto* ***izdanie***
FINANCING THIS **PUBLICATION**
'financing this **publication**'

Regarding deverbal nouns denoting states, they can realize the internal argument by means of a noun in genitive (59) or by means of an internal argument incorporated (60).

(59) *Nakonec-* *vseobščaja* [***demoralizacija*** [*nravov*]$_{\text{NP-GEN-NC-Arg1-tem}}$]$_{\text{NP}}$
FINALLY GENERAL **DEMORALIZATION** CUSTOMS
'Finally, there is a general **corruption** of the way of living.'

(60) *Ej* *neinteresny* *eë* *sobstvennye* [***perežívanija***]$_{\text{NP}}$
TO_HER NON_INTERESTING HER OWN **SUFFERINGS**
'She is not interested in her own sufferings.'

3.6.3 Specifiers

The specifier is an important criterion to distinguish the denotation of deverbal nouns in other languages such as English or Spanish. Peris (2009) claims that definite articles and possessive determiners appear with Spanish deverbal nouns denoting results and events. Moreover, bare nouns (that is, nouns with no specifier) can have both readings too. On the other hand, demonstrative determiners, indefinite articles and numerals seem to only appear with nouns denoting result readings. However, in Slavic languages this is not an informative criteria, since most of the times NPs are bare. In Russian, the possible specifiers are: possessive determiners (*moj* 'my', *tvoj* 'your', etc.), demonstrative determiners (*etot* 'this', *tot* 'that', etc.), interrogative determiners (*kakoj* 'which', *chej* 'whose', etc.), relative determiners (*kotoryj* 'who, which', *chej* 'whose', etc.), negative determiners (*nikakoj* '*no*', *nichej* 'no-one's', etc.), definite determiners (*ves'* 'all', *kazhdyj* 'each', etc.) and indefinite determiners (*nekotoroj* 'some', *nekij* 'some', etc.), and possessors ('s car). In Russian, there are neither definite nor indefinite articles.

In the table 11, we find the number of deverbal nouns with or without specifier (and which type of specifier) according to their denotation.

Specifier	Event	Result	State	Unspecified
Definite	---	2	---	---
Indefinite	---	4	---	---
Demostrative	3	8	---	---
Possessive	6	4	2	---
Negative	---	2	1	---
Bare-NP	171	114	13	2

Table 11: Specifier and deverbal noun denotation

As said before, Russian specifiers are not enough informative, since 90.36% of deverbal nouns found in our sample appear as bare nouns.

Regarding nouns denoting events, they can be accompanied by possessive determiners and demonstrative determiners (61, 62, respectively). Contrary to what has been found in Spanish (Peris *et al.* 2009), in Russian a demonstrative determiner can accompany a deverbal noun denoting an event.

(61) *Srok* [[*eë*]$_{\text{Poss}}$ ***formirovanie***]$_{\text{NP}}$ *kratok*
PERIOD HER **FORMATION** SHORT
'The period of her **formation** is short.'

(62) [[*Eto*]$_{\text{Demonst}}$ *sliškom dolgoe* ***ugovarivanie***]$_{\text{NP}}$ *privelo nas* (…)
THIS TOO LONG **PERSUASION** BROUGHT US
'This too long **persuasion** brought us (…)'

Result nouns are far more flexible and can be accompanied by definite (63), indefinite (64), demonstrative (65), possessive (66) and negative determiners (67). For instance:

(63) [[*Vse*]$_{\text{Def}}$ ***izdanija***]$_{\text{NP}}$ *sozdajut svoi spiski*
ALL **PUBLICATIONS** CREATE THEIR-OWN RECORDS
'All **publications** create their own records.'

(64) [[*Nekotorye*]$_{\text{Indef}}$ *naučnye* ***izdania***]$_{\text{NP}}$ *pičkajut čitatelej*
SOME SCIENTIFIC **PUBLICATIONS** FILL_UP READERS
'Some scientific **publications** fill up the readers.'

(65) *Količestvo* [[*takikh*]$_{\text{Demonst}}$ ***zakazov***]$_{\text{NP}}$ *postojanno* *ubeličivaetsja*
AMOUNT LIKE_THESE **ORDERS** CONSTANTLY INCREASE
'The amount of this kind of **orders** increases constantly.'

(66) [[*Svoi*]$_{\text{Poss}}$ ***vyigryši*** *i* ***proigryši***]$_{\text{NP}}$
HIS **BENEFITS** AND **LOSSES**
'his **benefits** and **losses'**

(67) [[*Nikakikh*]$_{\text{Neg}}$ *revolucionnykh* ***izmenenij***]$_{\text{NP}}$ *v* *etom* *godu*
NO REVOLUTIONARY **CHANGES** IN THIS YEAR
ne *budet*
WON'T BE
'This year there won't be any revolutionary **change.**'

Regarding nouns with state reading, they can be accompanied by possessive determiners (68) and negative determiners (69).

(68) [[*Eë*]$_{\text{Poss}}$ *sobstvennye* ***perežívanija***]$_{\text{NP}}$
HER OWN **SUFFERINGS**
'her own **sufferings'**

(69) *Ne* *pozvoljavšim* *po* *otnošeniju* *k* *sebe* *nikakogo*$_{\text{Neg}}$ ***amikošonstva***
NOT PERMITING IN RELATION WITH HIM ANY **FAMILIARITY**
'He does not permit himself any **familiarity**.'

3.6.4 Selectors

Picallo (1999), Peris & Taulé (2009) and Brandtner *et al.* (2010) claim that adjectives, verbs and nouns surrounding the deverbal noun can select the reading of the deverbal noun. Brandtner *et al.* (2010) claim that "one of the predicates extends its meaning and thereby imposes different selectional restrictions" on the deverbal noun. Peris & Taulé (2009) proposes a group of prepositions, nouns, adjectives, adverbs and verbs in Spanish that can select a special reading of the deverbal noun.

Melloni & Jezek (2011) investigate the distributional behaviour of deverbal nouns in text, that is, the selectional properties of their verbal and adjectival collocates. According to Melloni & Jezek (2011), a verb such as *finance* selects an event reading, while a verb such as *examine* selects a result reading. In the same way, an adjective such as *possible* selects an event reading; while an adjective such as *wooden* selects a result reading.

Therefore, we are going to analyze which are these selectors in Russian: firstly, we study the influence of adjectives; secondly, we study the influence of some nouns as heads of a deverbal noun construction; thirdly, we analyze prepositions; and, finally, we study the role of the verb for which the deverbal noun can be the subject or the object.

A. Adjectives

Grimshaw (1990) for English and Schoorlemmer (1995) for Russian note that the adjectives 'constant' and 'permanent', when accompanying a deverbal noun, select an event reading (70, 71). However, despite having one of these adjectives, if the deverbal noun is in plural, then plurality rules out the interpretation of the nominal as an event (72).

(70) The constant examination of the student
(71) *The constant examination annoyed the students
(72) The constant examinations annoyed the students

[Grimshaw, 1990]

These authors claim that if a deverbal noun is modified by an agent-oriented adjective such as 'intentional' or by the adjective 'gradual' (*posteppenyj* in Russian) then the deverbal noun denotes an event.

Peris & Taulé (2009) postulate that relational adjectives, interpreted as the internal argument, select result readings, never event readings. In a construction such as *la producción quesera de los holandeses* 'The Dutch production of cheese' the relational adjective *quesera* 'cheese' selects a result reading. However, in our analysis, we have found deverbal nouns denoting events modified by relational adjectives, interpreted as internal arguments (73).

(73) *V*	*geroine*	*Štain*	*pokazal*	*fisičeskij*	*itog*
IN	HEROIN	SHTAIN	SHOW	PHYSICAL	RESULT

[[*nravstvennykh*]$_{\text{AP-NC-Arg1-tem}}$ ***razryvov***]$_{\text{NP}}$
MORAL **RUPTURES**

'In the heroin, Shtain showed the physical result of the moral **ruptures**.'

Furthermore, Peris & Taulé (2009) also claim that a deverbal noun acting as the complement of the adjective *resultante* 'resulting' is eventive. For instance, *la explosión resultante de la ebullición de estos líquidos* 'the explosion resulting from boiling these

liquids'. In (73), the deverbal noun *razryvov* 'of ruptures' acts as a complement of the noun *itog* 'result', as in Peris & Taulé (2009), the deverbal noun *razryv* 'rupture' denotes the process that brought about a consequence. In Russian, a construction such as *v rezultate* 'as a result of' accompanying a deverbal noun would select for it an event reading. The translation into Russian of the example *la explosión resultante de la ebullición de estos líquidos* 'the explosion resulting from the boiling of this liquids' would be *vzryv v rezul'tate kipenija etikh židkostej*. In a sentence like this, the explosion, that is, *vzryv* is the result of a process in which some liquids have been boiled.

Brandtner & Heusinger (2009) focus on German adjectives modifying deverbal nouns ending in *-ung*. They consider that those adjectives, which either modify the process of the deverbal noun (for instance, 'cautious') or indicate the iteration of a process of the deverbal noun (for instance, 'permanent'), select an event reading, whereas adjectives, which refer to the colour ('red'), to the shape ('round') or to the material ('wood'), select a result reading.

Fábregas & Marín (2011) claim that "adjectives such as *rápido* 'fast' and *lento* 'slow' qualify the way in which a dynamic predicate is performed and as such select events. They are not compatible with state nouns, and result nouns".

(74)

(a) [*La* *construcción* [*rápida*]$_{AP}$ [*del* *Puente*]$_{NP}$]
THE CONSTRUCTION FAST OF BRIDGE
'The fast construction of the bridge'

(b) *[*El* aburrimiento [*rápido*]$_{AP}$ [*de* *Juan*]$_{PP}$]$_{NP}$
THE BOREDOM FAST OF JOHN
'The fast boredom of John'

Regarding this, in our experiment we have found three kinds of adjectives which can appear with event nouns: process modifying adjectives (*podrobnyj* 'detailed'), that is, adjectives denoting the manner in which the action is carried out (75 and 76), adjectives concerned with the durativity (*neskončaemyj* 'never-ending'), that is, the temporal structure of the action (77 and 78) or adjectives, such as, *efektivnyj* 'efficient' referring to the consequences of the action named by the deverbal noun (79).

(75) [[*podrobnoe*]$_{AP}$ ***opisyvanie*** [*ženskogo* *tela*]$_{NP}$]$_{NP}$
DETAILED **DESCRIPTION** WOMAN BODY
'a detailed **description** of the woman's body'

(76) *na* [[*tščatel'nom*]$_{AP}$ *vyjavlenii* *i* ***obsledovanii***]$_{NP}$
ON CAREFUL DETECTION AND **EXAMINATION**
'based on a careful detection and **examination'**

(77) [[*neskončaemyje*]$_{AP}$ ***ugovory***]$_{NP}$
NEVER-ENDING **PERSUASIONS**
'never-ending **discussions'**

(78) *Eto* [[*sliškom* *dolgoe*]$_{AP}$ ***ugovarivanie***]$_{NP}$
THIS TOO LONG **PERSUASION**
'This is a too long **persuasion'**

(79) [[*efektivnogo*]$_{AP}$ ***stimulirovania***]$_{NP}$
EFFICIENT **STIMULATION**
'of an efficient **stimulation'**

Adjectives related to colour, shape, material and ordering appear generally with result nouns. They refer to a property of a concrete or an abstract entity. Adjectives such as *maksimal'nyj* 'maximal', *bol'shoj* 'big' and *krupnyj* 'enormous' select result readings. These adjectives denote either a degree of an entity (*maksimal'nyj* 'maximal') or the size of the entity (*bol'shoj* 'big' or *krupnyj* 'enormous'). The adjective in the example (81) *mešočnyj* 'of sack', denoting the material with which some entity is done, appears with result nouns.

(80) *Mogut* *proizojti* [[*krupnyje*]$_{AP}$ ***izmenenia***]$_{NP}$
MAY TO_HAPPEN BIG **CHANGES**
'Big **changes** may occur.'

(81) *Ja* *nabirala* *kovër* *iz* *starykh* *trjapok* *i*
I MADE CARPET FROM OLD RUGS AND
[[*mešočnykh*]$_{AP}$ ***ušivok***]$_{NP}$
SACK **PIECES**
'I made a carpet from old rugs and **pieces** of sack.'

However, adjectives denoting colour or material can accompany event nouns, when they are used figuratively. In example (96) the adjective *zolotoe* 'golden' is used with the meaning of 'great' and, then, it modifies a noun denoting an event, since in this context the adjective refers to the manner in which the action named by the deverbal noun is carried out.

(82) [[*zolotoe*]$_{AP}$ ***koronovanie*** [*geroev*]$_{NP}$]$_{NP}$
GOLDEN CORONATION HEROES
'The golden **coronation** of heroes.'

B. Nouns

As has been already outlined by Peris & Taulé (2009), NP with a deverbal noun embedded can influence the denotation of the embeded deverbal noun. In a sentence such as *La capacidad de adaptación* 'the ability of adaptation', the noun *capacidad* 'ability' favours the event reading of the deverbal noun *adaptación* 'adaptation'. Regarding this, we have found nouns that can have an influence on the reading of the deverbal noun. Nouns selecting event readings are close to the notion of process. Moreover, they also can make reference to the time or the temporal structure of the process named by the deverbal noun.

Nouns, which denote a process, select the event reading in the deverbal noun (*process* 'process', *opyt* 'experience', *rabota* 'work', *obuchenije* 'teaching', *sintez* 'synthesis', *akt* 'action', *dejatel'nost'* 'occupation', *funkcija* 'function').

(83) *Tam* *idut* [*processy* [*po* [***formirovaniju*** [*budyščej*
THERE GO PROCESSES ON **FORMATION** FUTURE
pandemii *grippa*]$_{NP}$]$_{NP}$]$_{PP}$]$_{NP}$
EPIDEMIC INFLUENZA
'Processes of **formation** of the future pandemic influenza take place there.'

(84) *U* *menja* *sfera* [*dejatel'nosti* –[***formirovanie*** [*ekonomičeskoj*
CLOSE_TO ME SPHERE ACTIVITY -**FORMATION** ECONOMIC
politiki]$_{NP}$]$_{NP}$]$_{NP}$
POLITICS
'My occupation is the **formation** of the economic policy.'

Nouns that name the consequence of a process denote an event (*effekt* 'effect').

(85) *ekonomičeskij* [*effekt* [*ot* [***sozdanija*** [*takikh system*]$_{NP}$]$_{NP}$]$_{NP}$]$_{NP}$
ECONOMIC EFFECT FROM **CREATION** SUCH SYSTEM
'the economic effect of the **creation** of such a system'

Nouns that name the way in which the process takes place (*sposob* 'method', *procedura* 'procedure', *mekhanizm* 'mechanism', *sreda* 'circumstances', *strategija* 'strategy', *sredstvo*

'means', *narušenije* 'disfunction') also denote an event.

(86) [*mekhanizm* [***stimulirovanija***]$_{NP}$]$_{NP}$
MECHANISM STIMULATION
'mechanism of **stimulation'**

Nouns that name the possibility, the reason, the obligation or the invitation to carry out a process denote an event (*pravo* 'right', *(ne)vozmozhnost'* '(im)possibility', *faktor* 'factor', *prizyv* 'call', *zakon* 'law').

(87) [*prizyvy* [*k* [***razryvu*** [*federativnykh* *svjazej*]$_{NP}$]$_{NP}$]$_{PP}$]$_{NP}$
CALL TO **BREAK** FEDERAL RELATIONS
'calls for the **breaking** of federal relations'

(88) *Soglasie* *bylo* [*odnim* *faktorom* [*ikh* ***razryva***]$_{NP}$]$_{NP}$
AGREEMENT WAS ONE FACTOR THEIR **BREAK**
'Agreement was one cause for their **rupture**.'

(89) *Eto* [*vozmožnost'* [***formirovanija*** [*podobnogo* *veščestva*]$_{NP}$]$_{NP}$]$_{NP}$
THIS POSSIBILITY **FORMATION** SIMILAR SUBSTANCE
'This is a possibility of **formation** of such a substance.'

Nouns that make reference to the participation in a process denote an event (*rol' v* 'a role in', *ispolzovanie v* 'the use in', *vklad* 'contribution').

(90) [*vklad* [*v* [***rešenie*** [*etoj* *problem*]$_{NP}$]$_{NP}$]$_{PP}$]$_{NP}$
CONTRIBUTION IN **SOLUTION** THIS PROBLEM
'a contribution to the **solution** of this problem'

(91) *ikh* [*ispolzovanije* [*v* [*promyšlennom* *sektore*] *i*
ITS USE IN INDUSTRIAL SECTOR AND
[*administrativnom* ***upravlenii***]$_{NP}$]$_{PP}$]$_{NP}$
ADMINISTRATIVE **DIRECTION**
'its use in the industrial field and in the administrative **direction'**

Nouns that make reference to the temporal structure of a process denote an event (*faza* 'phase', *etap* 'stage', *srok* 'period', *moment* 'moment', *den'* 'day', *let* 'year').

(92) [[*Važnym*]$_{AP}$ *etapom*]$_{NP}$ *javljajetsja* [***formirovanie*** [*navyka*]$_{NP}$]$_{NP}$
IMPORTANT PHASE IS **FORMATION** SKILL
'The **development** of the skill is an important phase.'

(93) [*Srok* [[*eë*] ***formirovanie*** [*kratok*]$_{NP}$]$_{NP}$]$_{NP}$
PERIOD HER **FORMATION** SHORT
'The period of its **formation** is short.'

(94) [*posle* [*neskol'ko* *let* [***urezania*** [*raskhodov*]$_{NP}$]$_{NP}$]$_{NP}$]$_{PP}$
AFTER SOME YEARS **CUTS** EXPENSES
'after some years of spending **cuts**'

(95) [*dni* [***koronovanija*** [*imperatora*]$_{NP}$]$_{NP}$]$_{NP}$
DAYS **CORONATION** EMPEROR
'the days of **coronation** of the emperor'

All these nouns tend to favour an event reading, since they are focused on the action rather than on the result.

On the contrary, nouns focused on the result of a process favour a result reading. These nouns make reference to an entity, to the existence of an entity, to the destruction of an entity or to the amount of an entity.

Nouns that name a concrete or an abstract entity: *predmet* 'subject'.

(96) [*Predmety* *–menedžment,* [***upravlenie,***]$_{NP}$]$_{NP}$ *ekonomika* *i*
SUBJECTS –MANAGEMENT, **DIRECTION**, ECONOMICS AND
pravo]$_{NP}$ *prepodajutsja* *lušče (…)*
LAW TEACH BETTER
'Subjects such as management, **direction**, economics and law are better taught (…)'

Nouns that name the presence, existence or destruction of a concrete or an abstract entity, such as *naličije* 'presence, avalaibility', *likvidacija* 'elimination', *pod kryšej* 'below the roof', *vyrabotka* 'elaboration' head an NP where the deverbal noun denotes a result.

(97) *zadačami*	*v*	*Turkestane*	*javljajutsja*	*ustranenie,*	*izžitie,*
TASKS	IN	TURKESTAN	ARE	REMOVAL	ELIMINATION
[***likvidacija***		[*tekh*	*nepravil'nykh*	*otnošenij*] $_{\text{NP-NC-Arg1-tem}}$]$_{\text{NP}}$	
LIQUIDATION		THOSE	WRONG	RELATIONS	

'The tasks in Turkestan are the removal, elimination and **liquidation** of those bad relations.'

Nouns that name the quantity of concrete or abstract entities such as *potok* 'flow', summa 'sum'.

(98) *potok*	[***zakazov***	[*iz*	*Evropy*]$_{\text{PP}}$]$_{\text{NP}}$]$_{\text{NP}}$
FLOW	**ORDER**	FROM	EUROPE

'the order flow from Europe'

C. Prepositions and adverbial locutions

Peris & Taulé (2009) postulate that some prepositions can influence the denotation of the deverbal noun. According to these authors, Spanish prepositions such as *tras* 'after', *durante* 'during' and *en* 'in' indicate temporality and because of that they select the event reading of the deverbal noun.

(99) [tras	[la	**presentación**	[de	la	documentación]$_{\text{NP}}$]$_{\text{NP}}$]$_{\text{PP}}$
AFTER	THE	**SUBMITION**	OF	THE	DOCUMENTATION

'after the **submission** of all papers'

In the data analyzed, we have also found that prepositions referring to temporality appear with event nouns. Prepositions such as *pri* 'during' and *pered* 'before' select an event reading in the deverbal noun (100 and 101). The eventive reading is predictable since they make reference to temporal structures which are only possible with actions but not with results.

(100) [*pri*	[***obsledovanii***	[*800*	*syvorotok*]$_{\text{NP}}$]$_{\text{NP}}$]$_{\text{PP}}$
DURING	**EXAMINATION**	800	SERUMS

'during the **examination** of 800 serums'

(101) *Vy*	*prosypajetes'*	*v*	*kamere,*	*v*	*tjurme,*	[*pered*
YOU	SLEEP	IN	CELL	IN	PRISON	BEFORE

[***kazn'ju***]$_{\text{NP}}$]$_{\text{PP}}$
EXECUTION

'You sleep in the cell, in prison, before the **execution**.'

As in Spanish (Peris & Taulé, 2009), we have found that the temporal preposition *posle* 'after', denoting temporality selects the event reading. In examples (102 and 103), the preposition *posle* 'after' focuses on the event.

(102) *Pojavit'sja* *takie* *simptomy* *čerez* *dve* *nedeli* [*posle*
TO_APPEAR THESE SYMPTOMS IN TWO WEEKS AFTER
[***začatija***]$_{NP}$] $_{PP}$
conception
'These symptoms appear two weeks after the **conception**.'

(103) [*Posle* [***začatija*** [*pervogo* *rebënok*] $_{NP}$] $_{NP}$]$_{PP}$ *umudrilas'*
AFTER **CONCEPTION** FIRST CHILD MANAGED
zaberemenet' *snova*
TO_BECOME_PREGNANT AGAIN
'After the **conception** of the first child, she managed to become pregnant again.'

Moreover, the preposition *posredstvom* 'by means' is a selector of an event reading. This preposition focuses on the action that must be carried out to achieve a particular target (104).

(104) *Dostigaetsja* [*posredstvom* [***usvojenia***]$_{NP}$]$_{PP}$
IS_ACHIEVED BY_MEANS **ASSIMILATION**
'It is achieved by means of the **assimilation**.'

The prepositional locution *v slučaje* 'in case of' selects an event reading, because it means that something will happen if something else happens first (105).

(105) [*V* *slučaje* [***razryva*** [*dogovora* *o* *družbe*]$_{NP}$]$_{NP}$]$_{PP}$
IN CASE **RUPTURE** CONVERSATIONS ON FRIENDSHIP
Rossija *postavit* *vopros*
RUSSIA POSE QUESTION
'In case of the **rupture** of the conversations about friendship relations, Russia will pose a question (…).'

D. Verbs

A verbal predicate, whose subject or complement is a deverbal noun, can act as a selector of the deverbal noun's reading. In *Die Messung ist gestern* 'The measuring was completed' (Bradtner & Heusinger, 2009), the predicate *ist gestern* 'was completed'

selects an event reading, since it refers to the completion of an action. Brandtner & Heusinger (2009) claim that time frame predicates (such as 'begin', 'stop', 'continue') or duration predicates (such as 'take place', 'last') select the event reading of the deverbal. On the other hand, verbal predicates that denote a physical change (such as 'present', 'appear') and those that indicate location ('lie on the table', 'be available') select the result reading.

Picallo (1999) proposes that predicates such as *ser inconsistent* 'to be inconsistent', *valer a algú un premi* 'to award somebody' o *enviar*-se *a algú* 'to be send' select a result reading.

Peris & Taulé (2009) propose that a predicate such as *perder* 'to lose' usually has a result reading, whereas the predicate *provenir* 'to come from' and the object of a gerund form denote an event reading.

Maienborn (2003 *op. cit.* in Fábregas & Marín, 2011) notices that, in the verbal domain, only events can be substituted by an expression such as *this happened*, which can be rendered in Spanish as *sucedió esto*. In example (106), the deverbal noun denotes an event as well as the pronoun *esto*. On the other hand, in (107), the deverbal noun *aburrimento* 'boredom' denotes a result since it cannot be substituted by the construction *sucedió esto* 'this happened'.

(106) *La* ***construcción*** *del* *puente* *fue* *largo.* ***Esto***
THE **CONSTRUCTION** OF-THE BRIDGE WAS LONG. **THIS**
sucedió porque
HAPPENED BECAUSE …
'The **construction** of the bridge was long. **This** happened because…'

(107) *El* ***aburrimiento*** *de* *Juan* *fue* *grande* *aquella* *tarde*,
THE **BOREDOM** OF JOHN WAS GREAT THAT AFTERNOON.
**Esto* *sucedió* *porque…*
***THIS** HAPPENED BECAUSE …
'The **boredom** of John was great that afternoon. *This happened because…'

In the sample analyzed, we have found verbal predicates denoting time frame, such as *končit'* 'to end', *pristupit'* 'to start', *zaveršit'sja* 'to complete' (108), which select an event reading. The same holds for predicates denoting duration or occurrence of an action such as *provodit'* 'to conduct', *privodit'* 'to result in', *zanimat'sja* 'to be occupied with' (109), *proizvodit'sja* 'carry out', *dlit'sja* 'to last' (110), *preryvat'sja* 'to be interrupted', *prekratit'sja* 'to be ceased' (111), *sostojat'sja* 'to take place', *prodolžat'sja* 'to continue' (112), *služit' dlja* 'to be used for', *igrat' rol' v* 'to play a role in', *rekomendovat'* 'to recommend' or *razbazarivat' čto-to na čto-to* 'to fool away', *ukhodit' na* 'to waste on', *prinimat' rešenie o* 'to decide about' (113), *rabotat' nad* 'to work on', *nastala pora* 'it is time to'.

(108) *Roditel'skij*	***razryv***	*nakonec*	*zaveršilsja*
PARENTAL	**RUPTURE**	FINALLY	COMPLETED

'Finally, the matrimonial **rupture** was completed'

(109) *Nel'zja*	*zanimat'sja*	*tol'ko*	***ugovarivaniem.***
PROHIBITED	TO_BE_OCCUPIED	ONLY	**PERSUASION**

'It is not permitted to be only occupied with **persuasions'**

(110) ***Razryv***	*dlilsja*	*pjatnadcat'*	*let*
RUPTURE	LASTED	FIFTEEN	YEAR

'The **rupture** lasted fifteen years'

(111) ***Obsluživanie***	*klientov*	*ne*	*prekrašalos'*
SERVICE	COSTUMER	NOT	CEASED

'Costumer's **service** was not ceased'

(112) ***rasledovanie***	*pričin*	*(…)*	*prodolžaetsja*
RESEARCH	CAUSES	(…)	CONTINUES

'the **research** of the causes (…) continues'

(113) *KUI*	*prinimajet*	*rešenie*	*o*	***prodaže***	*ob'jekta*
KUI	TAKES	DECISION	ABOUT	**SALE**	OBJECT

'KUI decides about the **sales** of the object'

However, predicates indicating location *byt' gde-nibud'* 'to be somewhere' (114)[29] or the transference or possession of some concrete or abstract entity such as *obladat'* 'to possess', *vzjat'* 'to take', *dat'* 'give' (115), *predostavit'* 'concede', *vyčitat'* 'remove', *razmestit'* 'collocate' or a copulative construction with an adjective denoting a quality of the entity, such as *byt' khoroša* 'to be beautiful' (116) denote a result. Moreover, there are some predicates, which only can have a concrete or an abstract entity in the subject or the object position to act as its agent or its theme. Verbal predicates such as *summirovat'* 'to sum up', *obnaružit'* 'to show', *somnevat'* 'to doubt', *fiksirovat'* 'to fix' act as selectors of a result reading.

(114) ***Perevod***	*na*	*stole*	*redaktora*
TRANSLATION	ON	TABLE	EDITOR

'The **translation** is on the editor's table.'

[29] Notice that in the present tense the verb *byt'* 'to be' is eliptic.

(115) *On* *podal* ***žalobu***
HE GAVE **COMPLAINT**
'He gave a **complaint.'**

(116) *Eto* ***rešenie*** *khorošoe*
THIS **DECISION** GOOD
'This is a good **decision.'**

3.7 Nominalizing suffixes and the lexical denotation of the nominalization

We have also carried out a brief analysis of 155 occurrences of deverbal nouns to study the influence of their nominalizing suffix on the lexical reading of the deverbal noun. Some nominalizing suffixes are specialized in giving nouns with a particular denotation, for instance *-ok* tends to denote results. However, most of them are quite flexible, for instance, the suffix *-ij(e)* which has a strong tendency to form nouns which denote events, but which also can form nouns denoting results or states.

Therefore, as in the case of the relationship between morphological aspect and lexical aspect, nominalizing suffixes are not determining in the identification of the denotation of a deverbal noun, but they show tendencies that correlate their presence in a noun with a special denotation. In table 12, we present the different nominalizing suffixes with their possible denotations.

Suffix	Event	Result	State	TOTAL
-ij(e)	72	14	8	**94**
-k(a)	11	12	0	**26**
-Ø	3	4	4	**11**
-stv(o)	1	1	6	**8**
-ok	0	8	0	**8**
-yš	0	3	0	**3**
-zn'	0	0	0	**2**
TOTAL	**87**	**45**	**18**	**152**

Table 12: lexical denotation of the deverbal noun and lexical reading

3.8 Conclusions

We can draw the following observations concerning the relationship between morphological and lexical aspect of the base verb and the lexical denotation of the nominalization derived. On the one hand, regarding morphological aspect, the analysis has shown that there is a tendency to express an event by means of a nominalization derived from an imperfective base verb, and a tendency to express a result by means of a nominalization derived from a perfective base verb. However, this fact is not systematic since there are deverbal nouns derived from imperfective verbs with a result reading and nouns derived from perfective verbs with an event reading. It is important to notice that, in spite of being influenced by their awareness of the morphological aspect of the base verb, the annotators did not take it as a crucial factor to discern the lexical denotation of deverbal nouns. On the other hand, there is a tendency to relate the lexical denotation of a deverbal noun to the lexical class of its base verb: in general, state verbs derive state deverbal nouns, activity verbs derive event nouns, accomplishment and achievement verbs derive both event and result nouns, but the latter usually derives result nouns. Therefore, the analyses of the relation between the morphological and lexical aspect of the base verb and the lexical denotation of the deverbal noun confirmed our initial hypothesis that neither the morphological nor the lexical aspect of the base verb determines the lexical denotation of the deverbal noun, although it has a significant influence on the lexical denotation of the nominal and can be considered an important criteria for denotative interpretation in Russian. As for the aspectual marks, our results confirmed the claim that the aspectual marks do not have grammatical function (Vinogradov 1972; Schoorlemmer 1995; Pazelskaya & Tatevosov 2003). Rather, the lexical denotation of the deverbal noun is more connected to the event structure of the base verb than the morphological aspect.

Deverbal nouns can be specialized in having one particular denotation instead of the other. Factors that can affect this specialization are: the existence of a different word to express the other denotation, such as the internal argument of the deverbal noun, or a different deverbal noun with the same meaning but different denotation. Moreover, the corresponding base verb can be the one, which is specialized in the event reading, whereas the deverbal noun is specialized in the result. These are means of the language to avoid ambiguity. On the other hand, the context of appearance (that is, selectors) and the frequency of the deverbal noun in the language are important factors that determine the denotative reading and the preference for a particular reading.

Since Grimshaw (1990), some other criteria to distinguish the denotation of deverbal nouns have been analyzed: the pluralization, the expression of the internal argument and the presence of specifiers. Regarding to the ability to pluralize, we have observed that deverbal nouns with an event reading avoid the plural, while result nouns can appear in singular as well as in plural. Regarding the expression of the internal argument, nouns denoting an event tend to express it, whereas nouns denoting a

result tend no to. Regarding specifiers, they are not informative in Russian since most of deverbal nouns appear as bare nouns (without a specifier).

We have listed some nouns, verbs, adjectives and prepositions that can act as selectors of a particular reading. For instance, we have observed that a verb establishing the location or existence of an element selects a concrete or an abstract entity, that is, a result deverbal noun; whereas a verb that names the beginning, the development or the ending of the deverbal noun selects an event reading.

Finally, we have observed that some nominalizing suffixes show preferences for a particular reading. However, they are only tendencies, since a suffix such as *–nij(e)* can have both readings in most of the cases.

Therefore, a deverbal noun has a particular denotation, depending on the intrinsic characteristics of morphological or lexical aspect of its base verb, but, mainly, depending on the context where it occurs.

Chapter 4

The argument structure of deverbal nouns

The aim of this chapter is to propose a descriptive study of the argument structure of deverbal nominalizations in Russian, the base for our analysis of translation mismatches. We study the argument structure of state, event and result deverbal nouns. The mapping between syntactic constituents and arguments of the deverbal nouns depends on the syntactico-semantic structure of the base verb from which the deverbal noun is derived. Therefore, we follow the commonly accepted hypothesis that claims that deverbal nouns inherit the argument structure of their base verb. With this aim, we have carried out a corpus-based study by means of the parallel Russian-Spanish MiniRuSp corpus, to ground our observations and descriptions on real language data (see chapter 2 to read about the resources used).

Firstly, we present a brief introduction to the Russian case since they have to be taken into account in order to understand the particularities of the noun phrase (hereinafter, NP) in Russian (section 4.1). We will see how, depending on the case and the presence or not of prepositions the syntactic function and the semantic interpretation of the NPs is different. Secondly, we present how the syntactico-semantic analysis was carried out (section 4.2), as well as the annotation scheme used (section 4.2.1). Thirdly, we present the observations obtained from the analysis (section 4.3). We describe in detail the syntactico-semantic structure of deverbal nouns. Concretely, the mapping between syntactic constituents and semantic arguments (section 4.3.1) and the patterns or combinations of arguments obtained (section 4.3.2). Finally, we give the conclusions (section 4.4).

4.1 Basic facts on Russian Morphology[30]

Before analyzing the argument structure of Russian deverbal nouns, in this section we will present a non-exhaustive introduction to the basic functions and meanings of Russian case. In Russian, nouns, adjectives, determiners, pronouns and numerals decline in gender, number and case. There are three genders (masculine, feminine and neuter), two numbers (plural and singular) and 6 cases (nominative, accusative, genitive, dative, instrumental and prepositional). Morphological case can be expressed by means of inflectional suffixes (1) and also by means of prepositions plus inflectional suffixes (2). The inflectional components of the Russian NP agree in gender, number and case. In example (1) the feminine nominative suffixes (that is, *-ja* for the possessive determiner, *-aja* for the adjective and *–a* for the noun) agree with each other. In example (2) the PP is introduced by the preposition *na* 'in' and the noun has the prepositional suffix *–e* in prepositional case.

(1) [*Moja* $_{\text{Poss-FEM-SG-NOM}}$ *xorošaja* $_{\text{AP-FEM-SG-NOM}}$ *podruga*$_{\text{noun-FEM-SG-NOM}}$]$_{\text{NP}}$ *ždala*
MY GOOD FRIEND WAITED
'My good friend waited.'

(2) *Ona* *ždala* [*v* *komnate* $_{\text{NP-FEM-SG-PREP}}$]$_{\text{PP}}$
SHE WAITED IN ROOM
'She waited in the room.'

Nominative (NOM) case always appears without preposition. Nouns in this case have the syntactic function of subject[31] (3) or the syntactic function of nominal predicate[32] (4). In (3), the noun *rebënok* 'child' is in nominative and acts as the subject. In (4) the nouns *dočka* 'daughter' and *studentka* 'student' are both in nominative and act as the subject and the nominal predicate, respectively.

(3)[*Rebënok* $_{\text{noun-NOM}}$]$_{\text{NP-Subj}}$ *spit*
CHILD SLEEPS
'The child sleeps.'

(4) [*Moja* $_{\text{Poss-NOM}}$ *dočka*$_{\text{noun-NOM}}$]$_{\text{NP-Subj}}$ -[*studentka* $_{\text{noun-NOM}}$]$_{\text{NP-Nominal predicate}}$
MY DAUGHTER STUDENT
'My daughter is a student.'

[30] For a more detailed description in English about the Russian case see Wade T. (1992) *A Comprehensive Russian Grammar* and Offord D. (1993) *Modern Russian. An advanced grammar course,* among others.
[31] Hereinafter 'Subj'.
[32] Hereinafter 'Nominal predicate'.

Accusative, genitive, dative and instrumental cases can appear with or without preposition. Regarding the accusative (ACC), the basic syntactic function is the direct object[33], which is expressed by means of a prepositionless form (5). On the other hand, the accusative preceded by a preposition (for instance *na* 'onto, to, at'*, v* 'to, in, on, at', or *čerez* 'over, across') has the syntactic function of an adjunct[34], whose meaning depends on the prepositions (destination (6), time, *etc.*).

(5) *stroit'* [*dom* noun-ACC]NP-DO
TO_BUILD HOUSE
'to build a house'

(6) *položit'* [*knigu* noun-ACC]NP-DO [*na* stol noun-ACC]PP-ADJ
TO_PUT BOOK ON TABLE
'to put the book on the table'

The main syntactic functions of Genitive (GEN) without preposition are the noun complement[35] denoting the possessor (7) and the object with a partitive meaning (8). A genitive can be headed by several prepositions, and depending on the preposition (*dlja* 'for', *okolo* 'near' or *iz* 'from'), the NP can have a big range of meanings. For instance, a genitive headed by *dlja* 'for' denotes the beneficiary of the action, that is, the indirect object[36] (9).

(7)[*kniga* noun-NOM [*sestry* noun-GEN]NP-NC]NP
BOOK SISTER
'my sister's book'

(8)*kupit'* [*khleba* noun-GEN]NP-DO
TO_BUY BREAD
'to buy some bread'

(9)*kupit'* [*putevku* noun-ACC]NP-DO [*dlja* *syna* noun-GEN]PP-IO
TO_BUY VOUCHER FOR SON
'to buy a voucher for the son'

[33] Hereinafter 'DO'.
[34] Hereinafter 'ADJ'.
[35] Hereinafter 'NC'.
[36] Hereinafter 'IO'.

Dative (DAT) without preposition has the basic syntactic function of the indirect object denoting the beneficiary of the action (10). A dative headed by different prepositions has adjunct functions with a variety of meanings depending on the preposition (*k* 'to', *po* 'at, on', among others). For instance, with the preposition *k* 'to', the PP names the destination of the action (11).

(10) *dat'* [*emu* pron-DAT]NP-IO [*knigu* noun-ACC]NP-DO
TO_GIVE HIM BOOK
'to give him the book'

(11) *plyt'* [*k* *vostoku* noun-DAT]PP-ADJ
TO_SAIL TO EAST
'to sail to the east'

The instrumental (INS) case has the basic prepositionless syntactic function of an adjunct[37] complement naming the instrument or the means by which an action is carried out (12) as well as the agent of an action in a passive sentence or in a deverbal nominalization (13). As with the other cases with prepositions, the different prepositions give different meanings to the NP. With the preposition *s* 'with', the PP refers to the company with whom the action was carried out (14).

(12) *pisat'* [*karandašom* noun-INS]NP-ADJ
TO_WRITE PENCIL
'to write with a pencil'

(13)[*dom* noun-NOM]NP-Subj *stroit'sja* [*brigadoj*noun-INS]NP-AgtC
HOUSE TO_BE_BUILT BRIGADE
'The house is built by the brigade.'

(14) [*brat*noun-NOM [*s* *sestroj* noun-INS]PP-NC]NP-Subj *ušli*
BROTHER WITH SISTER WENT_AWAY
'Brother and sister went away.'

As the name indicates, prepositional (PREP) case is always headed by a preposition. It has the syntactic function of an adjunct. Depending on the preposition, the PP can acquire a great variety of meanings. For instace, with the preposition *v* 'in', the PP refers to a location (15) or to a point in time (16).

[37] Hereinafter 'ADJ'.

(15) *sidet'* [*v* *komnate* $_{\text{noun-PREP}}$]$_{\text{PP-ADJ}}$
TO_SIT IN ROOM
'to sit in the room'

(16) *On* *umer* [*v* *1981* *godu* $_{\text{noun-PREP}}$]$_{\text{PP-ADJ}}$
HE DIED IN 1981 YEAR
'He died in 1981.'

To sum up, nominative always goes without preposition and has a core function (subject and nominal predicate), whereas prepositional case always goes with a preposition and has non-core functions (adjunct). In the case of genitive, accusative, dative and instrumental case they can appear with or without preposition. As basic syntactic function without preposition, genitive acts as noun complement, the accusative as a direct object, the dative as indirect complement, and, finally, an instrumental acts as an agentive complement or adjunct. All these cases with preposition are adjuncts. However, in cases such as (9), the preposition *dlja* 'for' with a noun in genitive has the syntactic function of an IO.

4.2 MiniRuSp: Syntactico-semantic annotation

In order to analyze the argument structure of deverbal nouns in Russian, the bilingual MiniRuSp corpus was partially tagged with syntactic and semantic annotations carried out by one annotator manually. MiniRuSp, a subsample of RuSp, is composed of 500 occurrences of Russian deverbal nouns, which correspond to 114 different lemmas and their corresponding translations into Spanish. The aim of this study is to determine which type of arguments the nominalizations have, in which order they are realized, and by means of which constituent. The annotation consisted of:

(a) the identification of the constituents of the noun phrase headed by the deverbal noun, namely, a NP, an adjective phrase (AP), a prepositional phrase (PP), a subordinate clause (SubC), which includes substantive, infinitive and relative clauses, and, finally, a possessive determiner (Poss);

(b) the assignation of a syntactic function, that is, noun complement (NC), specifier (Spec), subject (Subj), direct object (DO), indirect object (IO) and adjunct (ADJ); and

(c) the determination of whether a constituent is argumental and, if so, the assignment of an argument position Arg0, Arg1, Arg2, Arg3, Arg4 and ArgM; and its corresponding thematic role, for instance, agent (agt), experiencer (exp), theme (tem), among others.

The syntactico-semantic analysis carried out was limited to the NP headed by the deverbal noun in the case of Russian, while in the case of Spanish the analysis copes with other structures such as subordinate or infinitive clauses since the translation of

a Russian deverbal noun does not always correspond to a Spanish deverbal noun. In some occasions, the deverbal noun (17.i) is translated into a non-deverbal noun in Spanish (17.ii) and, consequently, it does not have argument structure, or the Russian deverbal noun (18.i) can be translated into a verb (18.ii).

(17)
(i)[***gorenie***$_{\text{noun}}$ [*rastajavšej* *sveči*$_{\text{GEN}}$]$_{\text{NP-NC}}$]$_{\text{NP}}$
BURNING MELTED CANDLE
'the **burning** of a candle melted'

(ii) [*la* ***llama***$_{\text{noun}}$ [*de* *una* *vela* *derretida*]$_{\text{PP-NC}}$]$_{\text{NP}}$
THE **FLAME** OF A CANDLE MELTED
'the **flame** of a melted candle'

(18)
(i) *sposobnost'* [***vydači***$_{\text{noun}}$ [*znanij*]$_{\text{NP-NC}}$]$_{\text{NP}}$
SKILL **DELIVERY** KNOWLEDGE
'the skill of knowledge **transmission**'

(ii) *la* *facultad* *de* [***exponer***$_{\text{inf}}$ [*algún* *conocimiento*]$_{\text{NP-DO}}$]$_{\text{InfC}}$
THE SKILL OF **TO_PRESENT** SOME KNOWLEDGE
'the skill to **present** some knowledge'

In example (17), the Russian deverbal noun *gorenie* 'burning' is translated into the Spanish non-deverbal noun *llama* 'flame'. In (18), the Russian deverbal noun *vydači* 'delivery' is translated into the Spanish infinitive *exponer* 'to present'. These two examples will be explained in much more detail in chapter 5, which is devoted to translation mismatches.

In section 4.2.1, we present the annotation scheme adopted and, in section 4.2.2, we present the results obtained in the linguistic analysis.

4.2.1 Argument structure annotation scheme

Our annotation scheme for the argument structure follows the proposal for the annotation of deverbal nouns in Spanish AnCora corpus (Peris & Taulé, 2011b), which is based on the PropBank proposal of argument annotation (Palmer *et al.,* 2005) and on the proposal of thematic annotation in VerbNet (Kipper *et al.,* 2006), and posterior reformulations in SemLink[38] (Loper *et al.* 2007, Yi *et al.* 2007, Palmer *et*

[38] http://verbs.colorado.edu/semlink/

al., 2009). This proposal distinguishes between core and adjunct arguments. The core arguments are numbered according to their degree of proximity in relation to the predicate, that is, Arg0, Arg1, Arg2, Arg3, Arg4. Therefore, Arg0 is the nearest and Arg4 is the farthest. And the adjunct arguments are labeled as ArgM. Each argument has assigned a thematic role. The list of thematic roles includes 20 tags widely accepted in Linguistics: agent (agt), causer (cau), experiencer (exp), patient (pat), theme (tem), cotheme (cot), beneficiary (ben), attribute (atr), extension (ext), location (loc), coagent (coagt), coexperiencer (coexp), instrument (inst), source (src), initial state (ei), goal (goal), final state (ef), destination (dest), time (tmp) and manner (mnr). All the tags used in our annotation correspond to those used in AnCora. We have just included two additional thematic roles coagent and coexperiencer, which have been specially designed and added to the existing list, and the 'adv' thematic role, which corresponds to non-specific adjuncts, has been not used in our annotation.

Table 1 presents the different thematic roles that can correspond to a particular argument. The combination of arguments and thematic roles results in a list of 24 different semantic tags.

Argument	Thematic Role	Definition[39]
Arg0	**Agent (agt)**	The agent (agt) corresponds to an animate subject, which volitionally controls the predicate (we may include internally controlled subjects such as forces and machines).
	Cause (cau)	The cause (cau) is the causing event that yields to some action or to a change of state.
	Experiencer (exp)	The experiencer (exp) is the participant that is aware or experiencing something.
Arg1	**Patient (pat)**	A patient (pat) is the participant that is undergoing a process or that has been affected in some way by an action or a state. Moreover, a patient must participate in the possibility of the conversion of an active construction into a passive.
	Theme (tem)	A theme (tem) is used for participants that undergoing a process or a change but do not participate in the possibility of a conversion of an active into a passive construction.
	Cotheme (cot)	A cotheme (cot) is used for participant that is participating in the action with the theme.

[39] http://verbs.colorado.edu/~mpalmer/projects/verbnet.html

Arg2	**Beneficiary (ben)**	A beneficiary (ben) is the participant that is benefited by some action.
	Attribute (atr)	An attribute (atr) refers to a quality of someone or something that is being changed.
	Extent (ext)	An extension (ext) is used to specify the range or degree of change.
	Location (loc)	Location (loc) is used to specify the place where the action takes place.
	Coagent (coagt)	Coagent (coagt) is used for the participant that is participating in the action with the Agent.
	Coexperiencer (coexp)	Coexperiencer (coexp) is used for the participant that is experiencing the process or the change of state with the Experiencer.
Arg3	**Instrument (inst)**	Instrument (inst) is used for objects or forces that come in contact with an object and cause some change in them.
	Source (src)	Source (src) is used for the start point of an action.
	Initial state (ei)	The initial state (ei) is used for the state, which is prior to the change.
Arg4	**Goal (goal)**	Goal (goal) is used for the aim/motivation of an action.
	Final state (ef)	The Final State (ef) is used for the state, which has undergone the change.
	Destination (dest)	Destination (dest) is used for the end point of a motion or the direction towards which the motion is directed.
ArgM	**Time (tmp)**	Time (tmp) is used to express time.
	Location (loc)	Location (loc) is used to specify the place where the action takes place.
	Manner (mnr)	Manner (mnr) is used for the way in which an action takes place.
	Cause (cau)	The cause (cau) is the causing event that yields to some action or to a change of state.
	Goal (goal)	Goal (goal) is used for the aim/motivation of an action.
	Source (src)	Source (src) is used for the start point of an action.

Table 1: Thematic roles based on VerbNet

As seen in table 1, the possible combinations of arguments and thematic roles are the following: Arg0 can correspond to agent (agt), causer (cau) and experiencer (exp) roles; Arg1 can correspond to patient (pat) and theme (tem) roles; Arg2 to beneficiary (ben), attribute (atr), extension (ext) and locative (loc) roles; Arg3 to instrument (inst), source (src) and initial states (ei); Arg4 to goal (goal) and final state (ef) roles. Finally, ArgM can correspond to adjuncts such as time (tmp), location (loc), manner (mnr), cause (cau), source (src) and goal (goal).

4.3 Analysis of the argument structure of Russian deverbal nouns

In this section, we present the observations made in the analysis of MiniRuSp, concretely, it describes which type of argument the deverbal noun has, in which position the arguments appear, which constituents realize the arguments and how (that is, inside or outside the NP headed by the deverbal noun) and which combinations are more frequent. The arguments of a verbal or nominal predicate are the participants involved in the action named by either a verb or a noun. As we assume the commonly accepted claim that the argument structure of a deverbal noun is the same as the argument structure of their corresponding base verb, the argument of a deverbal noun "is that one, which can be semantically interpreted as one of the arguments associated to the corresponding verb" (Peris, 2012:66).

The analysis of the data gives rise to the observations presented below. The arguments of a deverbal noun can be syntactically realized in three different ways: explicitly, that is, inside the NP whose head is the deverbal noun (19); implicitly, that is, outside the NP headed by the deverbal noun, usually in the linguistic or extralinguistic context (20); and incorporatedly in the deverbal noun (21).

(19) [*Akkuratnyj* $_{\text{AP-NC-ArgM-mnr}}$	***perevod***	*Marii* $_{\text{NP-NC-Arg0-agt}}$]$_{\text{NP}}$	*na*	*stole*
ACCURATE	**TRANSLATION**	MARY	ON	TABLE

'Mary's accurate **translation** is on the table.'

(20) *Ty* $_{\text{NP-Subj-iarg0-agt}}$	*mne* $_{\text{NP-IO-iarg1-pat}}$	*ne*	*poklonilsja*	[*pri*	[***vstreče***]$_{\text{NP}}$]$_{\text{PP}}$
YOU	TO-ME	NOT	SAID-HELLO	DURING	**ENCOUNTER**

'You didn't say hello to me during the **encounter.'**

(21)[*Vpolne*	*real'nyj*	***vyigriš*** $_{\text{noun-Subj-Arg1-tem}}$]$_{\text{NP}}$
COMPLETE	REAL	**GAIN**

'a complete real **gain**'

In example (19), there are two explicit arguments overtly expressed inside the NP headed by the deverbal noun, namely, *akkuratnyj*$_{\text{ArgM-mnr}}$ 'accurate' and *Marii*$_{\text{Arg0-agt}}$ 'of

Mary'. In example (20), the arguments are not overtly expressed in the local context of the NP, but they can be inferred from the linguistic context. The implicit arguments of the deverbal noun *vstreča* 'encounter' in (20) are $mne_{\text{iarg1-pat}}$ 'to me' acting as the patient and $ty_{\text{iarg0-agt}}$ 'you' acting as the agent appearing in the same sentence but outside the NP headed by the deverbal noun. Following Gerber & Chai (2010, 2012) and Peris, Taulé, Rodríguez & Bertran (2013), we use the 'iarg' tag to label the implicit arguments and differenciate them from the explicit arguments. In example (21), *vyigriš* 'gain' has the theme argument (Arg1-tem) incorporated in the deverbal noun.

In our analysis, we are concerned with explicit arguments; the incorporated and implicit arguments of deverbal nouns are not the goal of this work and they have not been annotated in MiniRuSp.

The number of explicit arguments realized inside the NP can be up to three[40] but as we can see in figure 1, deverbal nouns with three arguments represent a negligible amount.

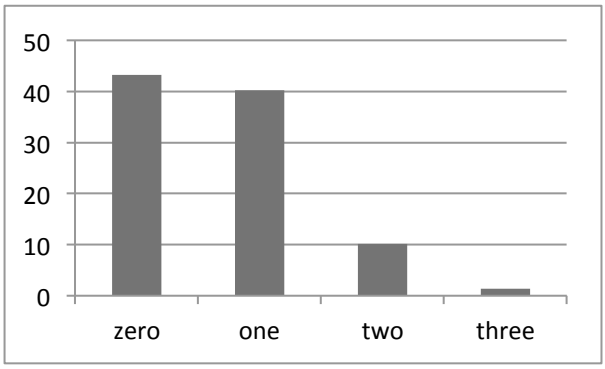

Figure 1: Percentages of deverbal nouns with zero, one, two, three arguments explicitly realized

As we can observe in figure 1, the most frequent situation is either the null realization inside the NP of explicit arguments (43.24%), or the explicit realization inside the NP of one argument (40.22%). Only in 10.15% the NP has two arguments explicitly realized and in 1.31% the NP has three arguments syntactically realized. These figures give more empirical evidences to support the claim that deverbal noun focus on the action/result denoted, that is, on the predicate meaning, rather than on the arguments taking part in the action. Moreover, they are usually recoverable from the linguistic and extralinguistic context (by means of the implicit arguments).

In the following subsections, we present the type of arguments and the constituents that realizes them. In Russian, constituents that can function as arguments of

[40] In some rare cases a deverbal noun can express more tan three arguments.

deverbal nouns are NP, AP, PP, SubC (which includes both substantive -declarative and infinitive- and relative clauses). These constituents act syntactically as noun complements. Possessive determiners can also express an argument of a deverbal noun and act syntactically as specifiers. The rest of specifiers, that is, demonstratives, numerals, indefinites, etc. cannot be arguments. Moreover, NPs in Russian tend to be bare, that is, they do not have any specifier (see chapter 3, section 3.6.3). Different constituents can express the same argument and thematic role, for instance, an Arg0, acting as experiencer, can be expressed by a NP, AP, PP and by a possessive determiner as we will see in more detail in section 4.3.1.

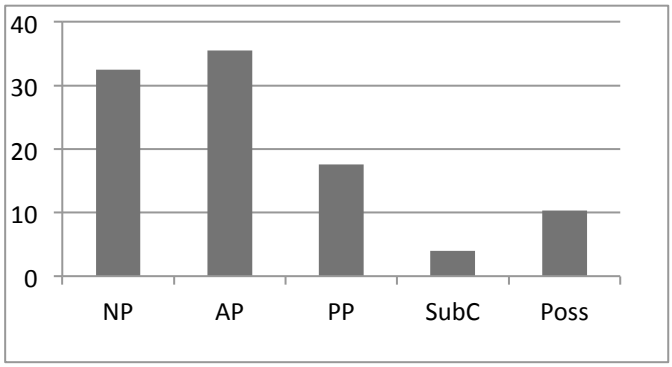

Figure 2: Frequency of constituents

As we can observe in figure 2, the commonest constituents acting as arguments are the AP (35.56%) and the NP (32.52%), followed by PP (17.62%), Possessive determiners (10.33%) and the least frequent is SubC (4%).

In the following subsections we present each constituent with its possible arguments and thematic roles (4.3.1), and which types of argumental combination can take place (4.3.2).

If we take into account the type of argument realized inside the NP, we observe that the argument explicitly realized is the core argument Arg1 (42.37%) followed by the non-core argument ArgM (33.53%) and Arg0 (16.15%). These arguments are by far the commonest since altogether they correspond to 92.05% of the explicit arguments. The least frequent arguments are Arg2 (5.18), Arg4 (2.13%) and Arg3 (0.60%).

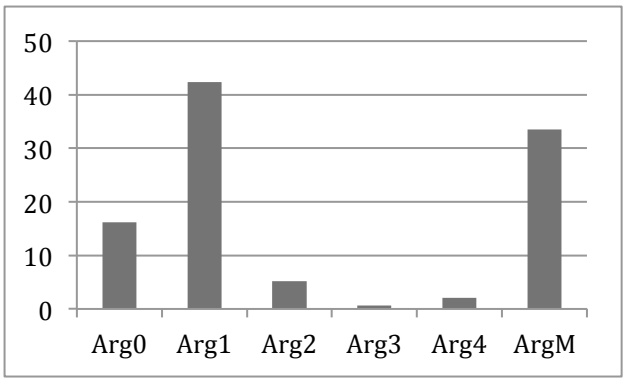

Figure 3: The frequency of the different types of arguments

This tendency responds to the fact that Arg0 and Arg1 can be implicit and inferred from the linguistic or the extralinguistic context. ArgM, on the other hand, may not be recovered from the context if it is not realized and consequently is lost.

4.3.1 Mapping between constituents, arguments and thematic roles

In this section, we present in detail which type of syntactic constituents (NP, AP, PP, SubC and Poss) realize the explicit semantic arguments inside the NP headed by the deverbal noun. In each case, we will give the argument position and the thematic role that can realize each constituent followed by examples extracted from the corpus.

Noun Phrase

In table 2, we list the different correspondences between the type of argument and the thematic role in which we find the NP.

	Arg	Thematic-role	Example
NP	Arg0	Agent (agt)	(22) *Eto* *byla* [***dobyča*** [*Savčenko*]NP-GEN-NC-Arg0-agt]NP THIS WAS PREY SAVCHENKO 'This was Savchenko's **prey.**'
		Experiencer (exp)	(23)*vypolnjaja* [***želanie*** [*inženiera* ACCOMPLISHING **DESIRE** ENGINEER *Kiseleva*] NP-GEN-NC-Arg0-exp]NP KISELEV 'accomplishing the **desire** of the engineer Kiselev'
	Arg1	Patient (pat)	(24)[***naselenie*** [*territorii*]NP-GEN-NC-Arg1-pat]NP **COLONIZATION** TERRITORY '**colonization** of the territory'
		Theme (tem)	(25)*Est'* [***iskaženie*** [*istiny*]NP-GEN-NC-Arg1-tem]NP THERE_IS **DISTORTION** TRUTH 'There is a **distortion** of the truth.'
	Arg2	Attribute (atr)	(26) *Nakhodilsja* *v* [***sostojanii*** WAS IN **STATE** [*oščelomlenija*]NP-GEN-NC-Arg2-atr]NP CONFUSION 'He was in a **state** of confusion.'
	ArgM	Temporal (tmp)	(27) *Ja* *znal* *nemnožko* *po* [***peresylke*** I KNEW A_LITTLE FROM **TRANSIT_CAMP** [*prošlykh* *let*]NP-GEN-NC-ArgM-tmp]NP PAST YEARS 'I knew (him) a little from a transit camp of years ago.'

Table 2: NP and its arguments and thematic roles

In the sample analyzed, a NP (32.52%) can be: Arg0 with the thematic role of agent (73.68%) or experiencer (26.31%); Arg1 with the thematic role of patient (60.97%) or theme (39.02%). Therefore, very frequently NPs realize core arguments, particularly, Arg1 (76.63%) and Arg0 (17.75%). In very few cases, the NP realizes less core arguments such as Arg2 (0.93%) and ArgM (3.73%). Moreover, the NP is a genitive, which acts as a noun complement (98%), whereas in a very small proportion the NP is in dative.

Adjective Phrase

In table 3, we list the different correspondences between the type of argument and the thematic role in which we have found the AP.

AP	**Arg0**	**Agent (agt)**	(28) *Vygovarivaja* [[*blatnye*]AP-NC-Arg0-agt ***vyraženija***]NP TALKING CRIMINAL **EXPRESSION** 'using a criminal **expression'**
		Experiencer (exp)	(29) [[*Sobstvennomu*]AP-NC-Arg0-exp ***želaniju***]NP OWN **DESIRE** 'his/her own **desire**'
	Arg1	**Theme (tem)**	(30) [[*Lorino*]AP-NC-Arg1-tem ***sušestvovanie***]NP LORINA'S **EXISTENCE** 'Lorina's **existence'**
		Patient (pat)	(31) *polučit'* [[*medicinskie*] AP-NC-Arg1-pat ***znanija***]NP TO_RECEIVE MEDICAL **KNOWLEDGES** 'to receive medical **knowledge'**
	Arg2	**Attribute (atr)**	(32) [[*nervnoe*] AP-NC-Arg2-atr ***sostojanie***]NP NERVOUS **STATE** 'nervous **state'**
	Arg4	**Destination (dest)**	(33) *V* [*ženskoj* [*taežnoj*] AP-NC-Arg4-dest ***komandirovke***]NP TO WOMEN'S TAIGA **EXPEDITION** *ne popala* NOT SENT 'She was not sent to a women's **expedition** to the taiga.'
	Arg M	**Location (loc)**	(34) *pol'zujas'* [*krasočnym* [lagernym]AP-NC-ArgM-loc USING COLOURFUL CAMP ***vyraženiem***]**NP** **EXPRESSION** 'using a colourful **expression** of the camp (of prisoners)'
		Manner (mnr)	(35) *trebuja* [[*skrupuleznogo*]AP-NC-ArgM-mnr ***vypolnenija*** DEMANDING ACCURATE **ACCOMPLISHMENT** *trebovanij*]NP DEMAND 'demanding an accurate **accomplishment** of the demands'
		Temporal (tmp)	(36) *sdača* [[*nočnogo*] AP-NC-ArgM-tmp ***dežurstva***]NP SWITCH NIGHT **SURVEILLANCE** 'the switch of the night **surveillance'**
		Goal (goal)	(37) [*metodičeskie* *ežednevnye* [*smertnyj*]AP-NC-ArgM-goal METHODIC DIARY MORTAL ***izbijenija***]NP **BEATINGS** 'mortal, methodic and diary **beatings'**

Table 3: AP and its arguments and thematic roles

In the case of APs (35.56%), the more frequent is ArgM (72.64%); on the other hand, they are less frequent core arguments: Arg0 (11.11%), Arg1 (11.11%), Arg2 (4.27%) and Arg4 (0.85%). ArgMs denote: location (15.29%), manner (60%),

temporal (15.29%) or goal (9.41%). Arg0 have the thematic role of agent (92.30%) or experiencer (7.69%) and Arg1 have the thematic role of theme (23.07%) or patient (76.92%). As Arg2, they act as attributes, and as Arg4 (0.85%), they act as destinations.

In Russian, adjectives agree in gender, number and case with the deverbal noun. They are not argumental when they are qualificative (*otdel'nyj vkhod* 'separate entrance'), intensional-adverbial (*real'nyj vyigriš* 'real gain') or modal *obyčnaja vstreča* 'frequent meeting'). On the other hand, relational adjectives (*blatnoe vyraženie* 'criminal expression') or adverbial adjectives (*organizovannoe vystuplenie* 'organized demonstration') can be argumental.

Prepositional Phrases

In table 4, we list the correspondences between the type of argument and the thematic role in which we have found PPs complementing the deverbal noun.

PP	**Arg0**	**Experiencer (exp)**	(38) *vyzyvala* ***vozmuščenie*** *daže* PROVOKE **INDIGNATION** EVEN [*u tekh*]$_{\text{PP-NC-Arg0-exp}}$]$_{\text{NP}}$ OF THOSE 'It provoked the **indignation** even of those.'
	Arg1	**Theme (tem)**	(39)[***zaderžka*** [*na rabote*] $_{\text{PP-NC-Arg1-tem}}$]$_{\text{NP}}$ **DELAY** IN WORK 'A **delay** in the work'
		Patient (pat)	(40) [***izveščenije*** [*o vnezapnoj smerti* **NOTIFICATION** ABOUT SUDDEN DEATH *Kiseleva*] $_{\text{PP-NC-Arg1-pat}}$]$_{\text{NP}}$ KISELEV 'a **notification** about the sudden death of Kiselev'
	Arg2	**Beneficiary (ben)**	(41) [***Vydači*** *gematogena* [*dlja* **ADMINISTRATIONS** HEMATOGEN FOR *bol'nykh*] $_{\text{PP-NC-Arg2-ben}}$]$_{\text{NP}}$ *ne byla novost'ju* SICK NOT WAS NOVELTY '**The administration** of hematogen for sick people was not a novelty.'
		Coexperiencer (coexp)	(42)[***Znakomstv*** [*s nami*]$_{\text{PP-NC-Arg2-coexp}}$]$_{\text{NP}}$ *ne* **ACQUAINTAINCES** WITH US NOT *podderžival* STAND 'He didn't hold any relationship with us'.
		Coagent (coagt)	(43) *Ošibka est' v* [***igre*** [*so mnoj*]$_{\text{PP-NC-Arg2-coagt}}$]$_{\text{NP}}$ MISTAKE IS IN **GAME** WITH ME 'There is a mistake in the **game** with me.'

	Arg3	**Source (src)**	(44) *prava* [***vyezda*** [*s* *Kolymy*] PP-NC-Arg3-src]NP RIGHT **DEPARTURE** FROM KOLYMA 'the right to leave Kolyma'
	Arg4	**Destination (des)**	(45)[***vyezd*** [*na* *materik*] PP-NC-Arg4-dest *kak* *invalidu*]NP **DEPARTURE** TO CONTINENT AS DISABLED 'the **departure** to the continent as a disabled'
	ArgM	**Location (loc)**	(46) *ot* [***vzryvov*** [*v* *zaboje*] PP-NC-ArgM-loc]NP BECAUSE **BURSTS** IN MINE_FACE 'because of **bursts** in the mine face'
		Temporal (tmp)	(47) *dlja* [***kolonizacii*** *kolymskogo* *kraja* *v* *1932* *Godu*] PP-NC-ArgM-tmp]NP FOR **COLONIZATION** KOLYMEAN COUNTRY IN 1932 YEAR 'for the **colonization** of Kolyma's region in 1932'
		Manner (mnr)	(48) [***vyezd*** *na* *materik* *kak* [*invalidu*]PP-NC-ArgM-mnr]NP **DEPARTURE** TO CONTINENT AS INVALID 'the **departure** for the continent as an invalid'
		Goal (goal)	(49) [*strašnye* ***igry*** [*na* *živca*] PP-NC-ArgM-goal]NP TERRIBLE **GAMES** FOR LIFE_BAIT 'terrible **games** for the life bait'
		Final state (ef)	(50) *priznak* [*skorykh* ***peremen*** [*k* *luščemu*] PP-NC-ArgM-ef]NP SIGN FAST **CHANGE** INTO BETTER 'a fast **improvement**'

Table 4: PP and its arguments and thematic roles

PPs have special preferences for being Arg1s (28.81%) and ArgMs (37.28%). Other core arguments can be realized by means of a PP: Arg2 (16.94%) and with much less frequency, Arg3 (3.38%), Arg4 (10.16%) and Arg0 (1.69%).

ArgM can be associated with location (63.63%), temporal (4.54%), manner (13.63%), goal (9.09%) and final state (9.09%). Arg1 is associated with the theme (47.05%) and the patient (52.94%). Arg0 is usually associated with the experiencer; Arg2 is usually associated with the beneficiary (50%), the coexperiencer (40%) and the coagent (10%); an Arg3 is associated with the source; and, finally, an Arg4 is associated with the destination.

As we have said before, the semantic interpretation of the deverbal noun depends on the preposition introducing the NP headed by the deverbal noun. In table 5, we present different prepositions in relation with their denotations.

Denotation	Preposition
Location	*vokrug* 'around', *meždu* 'between', *mimo* 'by', *poperek* 'across', *u* 'by', *v* 'in', *za* 'behind', *na* 'on, upon; in; at', *čerez* 'over; through; via', *pod* 'under; near', *nad* 'above', *pred* 'in front of', *pri* 'near'
Source	*iz 'from', iz-za* 'from behind', *ot* 'from', *s* 'from',
Destination	*do* 'to, up to', *o* 'against, upon', *za* 'behind', *na* 'onto; to', *pod* 'under', *k* 'to', *po* 'along'
Time	*do* 'until; before', *s* 'at the same time of; after; about', *v* 'in, on, at; within; every', *za* 'at; during, over; past', *na* 'during, in; at; for', *čerez* 'in', *pod* 'near', *k* 'by', *po* 'in; up to; just after', *p(e)red* 'before', *pri* 'under', *meždu* 'between', *ot* 'from', *posle* 'after'
Cause	*iz* 'because of', *iz-za* 'because of', *ot* 'because of', *za* 'because of', *po* 'by', *blagodarja* 'because of', *s* 'because of', *po slučaju* 'in the occasion of', *v svjazi* 'in connection', *vvidu* 'in view of', *vsledstvie* 'as a result of', *v silu* 'on the strength'
Goal	*iz-za* 'for', *c* 'with', *za* 'for', *na* 'for', *k* 'for', *dlja* 'for', *radi* 'for the sake of',
Manner	*s* 'with; by', *v* 'in', *v* 'exactly like', *na* 'on', *čerez* 'by way of', *pod* 'in imitation of', *po* 'by'
Coagent/Coexperiencer	*s* 'with'
Theme	*o* 'about, on'
Final State	*v* 'into'
Beneficiary	*k* 'for', *dlja* 'for'
Instrument	*pri pomošči* 'with the help of', *s pomoščju* 'with the help of'

Table 5: Prepositions and thematic roles

Subordinate Clauses

In table 6, we list the different correspondences between the type of argument and the thematic role in which we have found the SubC.

SubC	**Substantive**	**Declarative**	**Arg1**	**Patient (pat)**	(51)[*ot* *skorbnogo* ***soznania,*** [*čto* *ikh* FROM PAINFUL **KNOWLEDGE** THAT THEIR *žizn'* *tak* *pečal'no* *složilas*]SubC-NC-Arg1-pat]NP LIFE SO SAD WAS 'from the painful **belief** that their life was so disgraceful'
		Infinitive	**Arg1**	**Patient (pat)**	(52) [*pobreždena* ***želaniem*** [*utverdit'*] SubC-NC-Arg1-pat]NP DAMAGED **WISH** TO_CONFIRM 'damaged by the **wish** to confirm'
			Arg2	**Attribute (atr)**	(53) *Ja* *ne* *v* [***sostojanii*** I NOT IN **POSITION** [*žertvovat'*]SubC-NC-Arg2-atr]NP TO_SACRIFICE 'I am not in a **position** to sacrifice.'
	Relative		**Arg0**	**Experiencer (exp)**	(54) *Viktor* *Ivanič*, *o* [***sušestvovanii*** [VIKTOR IVANIČ ABOUT **EXISTENCE** *kotorogo*]SubC-NC-Arg0-exp *ja* *uže* *stal* WHOSE I ALREADY BEGAN *zabyvat*]NP TO_FORGET 'Viktor Ivanič, whose **existence** I already began to forget'

Table 6: SubC and its arguments and thematic roles

The Subordinate Clause (SubC) (3.95%) complementing a deverbal noun can be substantive, relative or adverbial. When it is substantive, it can be declarative, interrogative or infinitive. In general terms, SubC can be Arg0 (7.69%), Arg1 (84.61%) and Arg2 (15.38%). They have a special tendency to be Arg1.

In the case of declarative substantive subordinate clauses, they realize Arg1s which act as patients. In the case of the infinitive substantive subordinate clauses, they can be Arg1 (90%) acting as patients or Arg2 (10%) acting as attributes. In the case of the relative subordinate clauses, the relative pronoun can be an Arg0 acting as experiencer.

This type of noun complement is not very frequent. However, it seems that subordinate clauses have a special tendency to act as core arguments. Specially, infinitives, situated after a deverbal noun, are considered argumental with a preference to act as patients or attributes. In this respect, we do not agree with Paducheva (2009) who considers that the infinitives are not arguments of the deverbal noun since the deverbal noun saturates the object. According to her, the deverbal noun and the infinitive are connected with each other by a kind of appositive relationship expressing co-reference.

(55) [***predloženie*** [*vystupit'*]$_{\text{SubC-NC-Arg1-tem}}$]$_{\text{NP}}$
SUGGESTION TO_COME_OUT
'the **suggestion** to make a speech'

(56)[***sovet*** [*uekhat'*]$_{\text{SubC-NC-Arg1-tem}}$]$_{\text{NP}}$
ADVICE TO_LEAVE
'the **advice** to leave'

In (55) and (56), the deverbal nouns *predloženie* 'suggestion' and *sovet* 'advice' are complemented by infinitives which are argumental and in both examples are Arg1 denoting theme, in the same way as their verbal counterparts (*predložit'* 'to suggest' and *sovetovat'* 'to advice').

The possessive determiner

In table 7, we list the different correspondences between the type of argument and the thematic role in which we have found the possessive determiner.

Possessive determiner	**Arg0**	**Agent (agt)**	(57) *Izorval* [*moi* Poss-Spec-Arg0-agt ***zapisi***]$_{\text{NP}}$ TORE MY **NOTES** 'He tore my **notes**.'
		Experiencer (exp)	(58) *dlja* [*moikh* Poss-Spec-Arg0-exp ***znakomstv***]$_{\text{NP}}$ FOR MY **ACQUAINTANCES** 'for my **acquaintances'** (59) [***zimovka*** *naša* Poss-Spec-Arg0-exp]$_{\text{NP}}$ **WINTERING** OUR 'our **wintering**'
	Arg1	**Theme (tem)**	(60) *segodnja den'* [*eë*$_{Poss}$-Spec-Arg1-tem ***roždenija***]$_{\text{NP}}$ TODAY DAY HER **BIRTH** 'Today it's her **birthday'**
		Patient (pat)	(61) *Na mesto* [*ego* Poss-Spec-Arg1-pat ***naznačenia***]$_{\text{NP}}$ TO PLACE HIS **DESTINATION** 'to the place of his **destination'**

Table 7: Possessive determiner and its arguments and thematic roles

The possessive determiner (10.33%) can only be core arguments: either the Arg0 (61.76%) -acting as agent (61.90%) or experiencer (38.04%) - or the Arg1 (38.23%) - acting as patient (23.07%) or theme (76.92%)-, being the Arg0 the special preference.

A possessive determiner will not be an argument if it denotes the possessor of an entity. In example (62), the possessive determiner has two different interpretations: If it is interpreted as the possessor of the book translated, then it is not argumental (62.i). On the other hand, if it is interpreted as the translator of the book, then it is argumental, that is, Arg0-agt (62.ii).

(62)

(i) *Ego*	***perevod***	*Gamleta*	*na*	*stole*	
HIS	**TRANSLATION**	HAMLET	ON	TABLE	

'His **translation** of Hamlet is on the table.'

(ii) *Su*	***traducción***	*de*	*Hamlet*	*está*	*sobre*	*la*	*mesa*
HIS	**TRANSLATION**	OF	HAMLET	IS	ON	THE	TABLE

'His **translation** of Hamlet is on the table.'

4.3.2 Argument patterns

In this section we present the patterns or combinations of arguments (and how they are syntactically realized) of deverbal nouns in Russian along with the figures of the frequency of their realization. Firstly, we will focus on the NP realizing one argument. Secondly, we will present those NPs with two or more arguments expressed.

One argument expressed inside the NP headed by the deverbal noun

We present the frequency of occurrence of each argumental constituent, of each type of argument and, finally, of correspondences between argumental constituents and its type of argument. In figure 4, we present the type of constituents that can be argumental in NPs with one argument expressed.

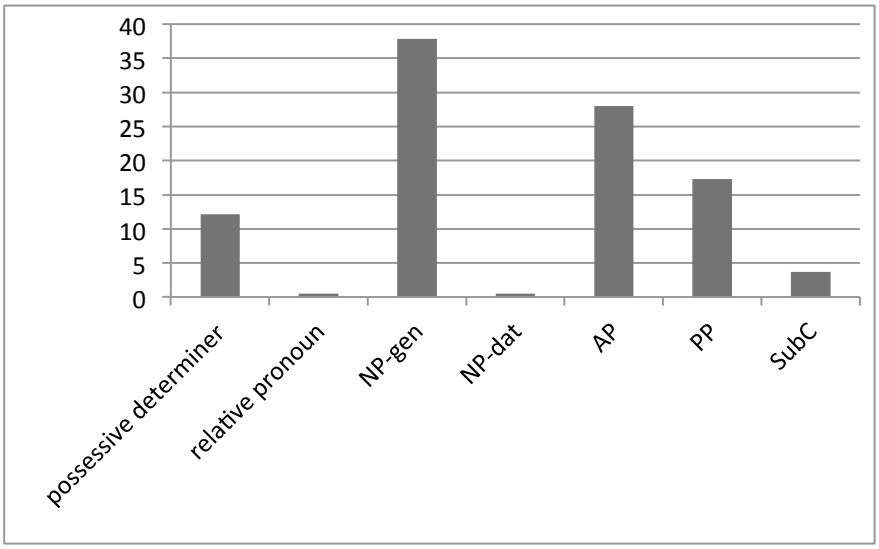

Figure 4: Type of constituent in AS with one argument syntactically realized

The constituents expressing the arguments of deverbal nouns more frequent are NPs in genitive (37.85%) or APs (28.03%), followed by PP (17.28%), the possessive determiner (12.14%) and SubC (3.73%). The least common are the NPs in dative (0.46%) and the relative pronoun (0.46%).

In figure 5, we present the type of arguments expressed in an NP with only one argument syntactically realized. As we will see, it is not strange to have the NP in genitive as the most frequent constituent expressing an argument and to have the Arg1 as the most frequent argument. It is also quite normal to have as the second most common constituent the AP and to have as the second most frequent argument the ArgM.

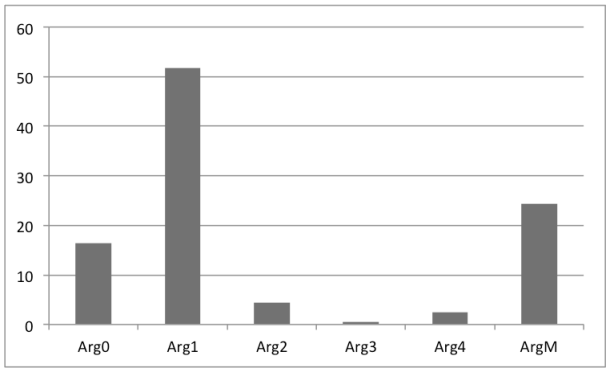

Figure 5: Type of argument in AS with 1 argument syntactically realized

In table 8, we show the correspondences between the constituent and the type of argument.

Constituent	Argument	%
NP	Arg0	12.5
	Arg1	82.5
	Arg2	1.25
	ArgM	3.75
AP	Arg0	13.55
	Arg1	18.64
	Arg2	3.38
	ArgM	64.4
PP	Arg0	3.12
	Arg1	31.25
	Arg2	18,75
	Arg3	3.125
	Arg4	15.62
	ArgM	28.12
Poss	Arg0	72.72
	Arg1	27.27

Table 8: Constituents and arguments in AS with one argument syntactically realized

Regarding deverbal nouns with one argument syntactically expressed, the usual argument is Arg1 (51.74%), followed by the ArgM (24.37%) and the Arg0 (16.41%). Arg2 (4.47%), Arg4 (2.48%) and Arg3 (0.49%) are less frequent.

NPs and Possessive determiners realize Arg0 and Arg1 almost exclusively. In MiniRuSp, 82.5% of NPs expressed an Arg1, whereas 72.72% of Possessive determiners realize Arg0. APs have a great tendency to realize ArgM (64.4%). Finally, PPs are more flexible and can realize Arg1 (31.25%), ArgM (28.12%), Arg2 (18.75), Arg4 (15.62%).

Two or three arguments expressed inside a NP headed by a deverbal noun

In this section, we present the combinations of constituents and arguments in NPs found in our sample with two or three arguments expressed. We have found very few cases of three or more arguments syntactically realized, in fact, they correspond to only 1.31% of cases.

In tables 9 and 11, we present these combinations along with an example for each one. We also present the frequencies of these combinations, the frequency of the

expressed arguments and, finally, the frequencies of constituents in relation to the argument.

Constituents	Example
AP+AP	(63)*pol'zujas'* [[*krasočnym*]AP-NC-ArgM-mnr [*lagernym*]AP-NC-ArgM-loc USING COLOURFUL CAMP ***vyraženiem***]NP **EXPRESSION** 'using a colourful **expression** of the camp (of prisoners)'
PP+PP	(64) *Odnomu Klokov ustraival* [***vyezd*** [*na* TO_ONE KLOKOV ORGANIZED **DEPARTURE** TO *materik*]PP-NC-Arg4-dest [*kak invalidu*]PP-NC-ArgM-mnr]NP CONTINENT AS INVALID 'Klokov organized for someone the departure for the continent as an invalid.'
SPEC+AP	(65) *rasskazat' o* [[*svoem*]Poss-Spec-Arg0-agt [*polnom*]AP-NC-ArgM-mnr TO_EXPLAIN ABOUT ONE'S_OWN COMPLETE ***opravdanii***]NP **JUSTIFICATION** 'to talk about his complete **justification**'
SPEC+PP	(66) *Vstrečalas' posle* [[*moego*]Poss-Spec-Arg0-exp ***znakomstva*** [*s* MET AFTER MY **ACQUAINTANCE** WITH *nim*]PP-NC-Arg1-coexp]NP HIM 'She met (him) after my **acquaintance** with him.'
SPEC+NP	(67) [[*ego*]Poss-Spec-Arg0-exp ***znanie*** [*jazykov*]NP-NC-Arg1-tem]NP HIS **KNOWLEDGE** LANGUAGES 'his **knowledge** of languages'
NP+PP	(68) [***Vydači*** [*gematogena*]NP-NC-Arg1-tem [*dlja* **ADMINISTRATIONS** HEMATOGENE FOR *bol'nykh*]PP-NC-Arg2-ben]NP *ne byli novostiju* SICK NOT WAS NEWS '**The administration** of hematogene for sick people was not a novelty'
AP+NP	(69) *trebuja* [[*skrupuleznogo*]AP-NC-ArgM-mnr ***vypolnenija*** DEMANDING ACCURATE **ACCOMPLISHMENT** [*trebovanij*]NP-NC-Arg1-tem]NP DEMANDS 'demanding an accurate **accomplishment** of the demands'
AP+PP	(70) *Načalis'* [[*strašnye*]AP-NC-ArgM-mnr ***igry*** [*na živtsa*]PP-NC-ArgM-goal]NP BEGAN TERRIBLE **GAMES** FOR LIVE_BAIT 'Terrible **games** for the live bait began'

AP+InfC	(71) *vyzvav* [[*strastnoe*]AP-NC-ArgM-mnr ***želanie*** [*izbavit'sja* HAVING_PROVOKED PASSIONATE **DESIRE** TO_GET_RID *klejma*]SubC-NC-Arg1-tem]NP STIGMA 'having provoked a passionate **desire** to get rid of the stigma'
AP+SubC	(72) *ot* [[*skorbnogo*]AP-NC-ArgM-mnr ***soznania,*** [*čto ikh žizn'* FROM PAINFUL **KNOWLEDGE** THAT THEIR LIFE *tak pečal'no složilas'*]SubC-NC-Arg1-tem]NP SO SAD WAS 'from the painful **belief** that their life was so disgraceful'

Table 9: Combination of two argumental constituents

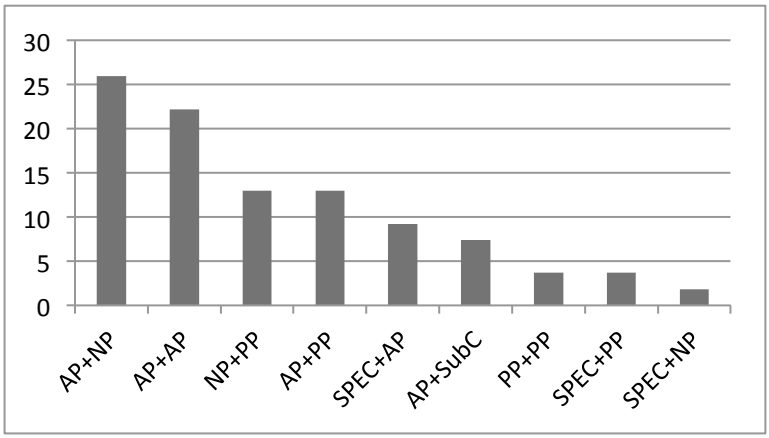

Figure 6: Type of constituent combinations in AS with two arguments syntactically realized

The commonest combination is AP+NP (25.92%) and AP+AP (22.22%). It is noticeable that these two combinations together represent 50% of cases found in MiniRuSp regarding nouns with two arguments syntactically realized. The following combinations are in the middle: NP+PP and AP+PP (12.95%); Spec+AP (9.25%) and AP+SubC (7.4%). The least common combinations are PP+PP (3.7%) and Spec+PP (3.7%) Spec+NP (1.85%).

Figure 7 presents which type of arguments are expressed in NPs with two arguments realized.

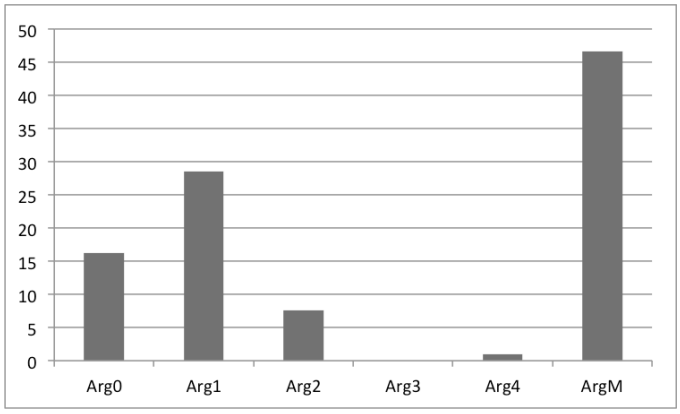

Figure 7: Type of arguments

The commonest argument, which is expressed in argument structures with two arguments syntactically realized, is the ArgM (46.66%) followed by the two core arguments, Arg1 (28.57%) and Arg0 (16.19%). The least common arguments are Arg2 (7.61%) and Arg4 (0.95%). In MiniRuSp we have not found any combination with an Arg3.

In table 10, we present the frequencies of constituents put in relation with the arguments.

Constituent	**Argument**	**%**
NP	Arg0	25.92
	Arg1	55.55
	Arg2	3.7
	Arg4	14.81
AP	Arg0	13.51
	Arg1	24.32
	Arg2	5.4
	ArgM	56.75
PP	Arg1	30
	Arg2	20
	Arg4	5
	ArgM	45
Poss	Arg0	100

Table 10: Constituent and argument in AS with two arguments syntactically realized

As we have already seen, NPs and Possessive determiners only realize Arg0 and Arg1. Concretely, NPs tend to express Arg1 (55.55%) and Possessive determiners tend to express Arg0 (100%). ArgM also preferred to be expressed by means of APs

(56.75%). PPs are more flexible, but they seem to have a preference to realize ArgMs (45%) in MiniRuSp.

The commonest combination of arguments is Arg1+ArgM (34%), followed by Arg0+ArgM (20%) and ArgM+ArgM (20%). Less common are the following combinations: Arg1+Arg2 (12%), Arg0+Arg1 (8%) and Arg0+Arg2 (2%) Arg2+ArgM (2%) and Arg4+ArgM (2%).

In table 11, we present the patterns consisting of NPs with three arguments syntactically realized.

Constituents	Example
AP+AP+AP	(73) [[*Metodičeskije*]$_{\text{AP-NC-ArgM-mnr}}$ [*ežednevnyje*] $_{\text{AP-NC-ArgM-tmp}}$ [*smertnye*] $_{\text{AP-NC-ArgM-goal}}$ METHODIC DAILY MORTAL ***izbienia***]$_{\text{NP}}$ *kazalis'* *užaznym* **BEATINGS** SEEMED HORRIBLE 'Mortal, methodic and diary **beatings** seemed horrible.'
AP+NP+PP	(74) *Končilsja* [[*obratny*]$_{\text{AP-NC-ArgM-mnr}}$ ***khod*** [*ryby*]$_{\text{NP-NC-Arg0-agt}}$ [*iz* ENDED RETURN **WAY** FISH FROM *ruč'ev*]$_{\text{PP-NC-Arg3-src}}$]$_{\text{NP}}$ STREAMS 'The return **way** of the fish from the streams ended.'
AP+AP+NP	(75) *Sygralo* *rol'* [[*ličnoe*]$_{\text{AP-NC-ArgM-mnr}}$ [*slučajnoe*] $_{\text{AP-NC-ArgM-mnr}}$ PLAYED ROLE PERSONAL ACCIDENTAL ***znakomstvo*** [Romanova]$_{\text{NP-NC-Arg0-exp}}$]$_{\text{NP}}$ **ACQUAINTANCE** ROMANOVA 'A personal accidental **acquaintance** of Romanova played a role.'
NP+PP+PP	(76) *prevyše* *vsjakikh* *čudes* [***okončanija*** [*sroka*]$_{\text{NP-NC-Arg1-tem}}$ [*v* SUPERIOR ANY MIRACLE **ENDING** CONDEMN IN *srok*]$_{\text{ArgM}}$, [*bez* *začetov* *rabočikh* *dnej*]$_{\text{PP-NC-ArgM-mnr}}$]$_{\text{NP}}$ PERIOD WITHOUT DISCOUNT WORKED DAYS 'It was superior to any miracle the **ending** of the condem in time without discounting worked days.'
AP+NP+NP	(77) [[*vnezapnaja*]$_{\text{AP-NC-ArgM-mnr}}$ ***vypiska*** [*makhorki,* SUDDEN **ORDERING** MAKHORKA, *sakhara*]$_{\text{NP-NC-Arg1-tem}}$]$_{\text{NP}}$ SUGAR 'a sudden **ordering** of makhorka and sugar.'
AP+AP+PP	(78) *Predvidel* [[*skoruju*]$_{\text{AP-NC-ArgM-tmp}}$ *i* [*nesomnennuju*]$_{\text{AP-NC-ArgM-mnr}}$ FORESAW FAST AND DOUBTLESS ***peremenu*** [*v* *obstajatel'stvakh*]$_{\text{PP-NC-ArgM-tem}}$]$_{\text{NP}}$ **CHANGE** IN CONDITIONS 'I foresaw a fast and solid **change** in the conditions.'

Table 11: Combination of three argumental constituents

Figure 8 presents the frequency of different combinations in NPs with three arguments expressed.

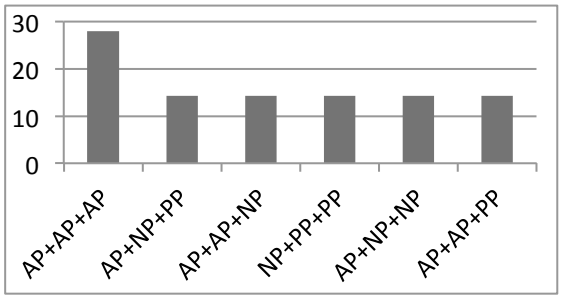

Figure 8: Frequency of combinations of constituents in AS with three arguments explicitly realized

The commonest combination is AP+AP+AP with 28% of cases, while the least common are the combinations of AP+AP+PP, AP+NP+NP or NP+PP+PP with 14.28%.

In figure 9, we see the type of arguments in NPs with three arguments realized.

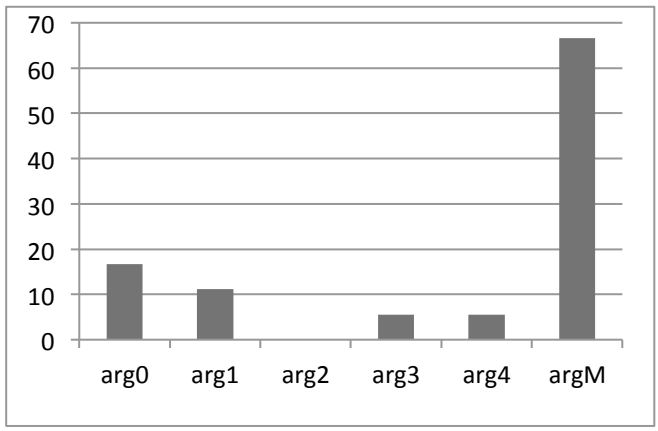

Figure 9: Type of argument

In our sample, regarding deverbal nouns' structures with three arguments syntactically realized, they usually express ArgM (63%). NPs tend to express Arg0 (40%), Arg1 (40%) and Arg3 (20%). APs tend to express ArgM (90%) and Arg4 (10%). PPs express Arg3 (25%) and ArgM (75%).

The commonest combinations of three arguments are Arg0+ArgM+ArgM (43.03%) and ArgM+ArgM+ArgM (28.57%) and less common are the following Arg0+Arg3+Arg4 (14.28%), and Arg1+ArgM+ArgM (14.28%). The number of deverbal nouns with three arguments explicitly realized inside the NP is very rare (1.31%), this fact poses a problem when it comes to generalize its behaviour.

4.4 Conclusions

In this chapter we have proposed a syntactico-semantic empirical description of Russian deverbal nouns as well as the patterns or combinations of argumental constituents in deverbal noun constructions. We assume that the argument structure of a deverbal noun is the same as its corresponding base verb but its syntactic realization can be different. The arguments of a deverbal noun can be incorporated inside the root of the deverbal noun, can be realized explicitly inside the NP headed by the deverbal noun and can be implicitly realized outside the NP headed by the nominalization. Given these three possible realizations of the arguments, it is not strange that in 43.24% of cases in MiniRuSp appeared with no arguments at all inside the NP headed by the deverbal noun. The argument structure of these nouns can be realized implicitly or incorporatedly. In 40.22% of cases in MiniRuSp appeared with only one argument explicitly expressed inside the NP headed by the deverbal noun. It is much more rare to have two or three arguments explicitly expressed (10.15% and 1.31%, repectively).

If we compare the results in Russian with those obtained in Spanish (Peris, 2012), constituents that are typically argumental in Russian are NPs, APs, PPs, SubCs and Possessive determiners. In Spanish, NPs and SubC are not argumental. What in Russian is expressed by means of a NP, in Spanish it is expressed by means of a PP introduced, most of the times, by the unmarked preposition *de* 'of'. In Russian, the arguments more usually syntactically realized are Arg1 (42.37%), ArgM (33.53%) and Arg0 (16.15%). In Spanish the more usually realized arguments are Arg0 and Arg1.

NPs and Possessive determiners tend to express Arg0 and Arg1. Concretely, NPs tend to realize Arg1, and Possessive determiners tend to realize Arg0. However, despite the preference for Arg0, they can realize Arg1. In Russian, NPs can also realize Arg2, Arg4 and ArgM. In Spanish, possessive determiners realize the argument corresponding to the subject in its verbal counterpart.

APs tend to express ArgM in Russian. In Spanish only relational adjectives can be argumental, Peris (2012) claims that APs tend to realize Arg0 (if a PP realizes the Arg1) or Arg1 and Arg0 (if there is no PP realizing the Arg1).

And, finally, PPs can express any type of argument in Russian (Arg0, Arg1, Arg2, Arg3, Arg4 and ArgM). However, its special tendency is to realize ArgM and Arg1. In Spanish they are quite flexible too, specially, those PPs introduced by the unmarked preposition *de* 'of'. However, as in Russian, when the preposition has semantic content, the type of argument realized depends on the meaning of the preposition. For instance, the Russian preposition *na* 'to' or the Spanish preposition *hacia* 'to' realize the Arg2-dest in both languages. Anyway, the main tendency of PPs in Spanish is to realize Arg1. In Russian, the equivalent constituent of a Spanish PP introduced by *de* 'of' is the NP in genitive. As in the case of the Spanish PP introduced by *de* 'of', the favourite argument is the Arg1. In both languages, this tendency can be explained because the Arg1 corresponds to the patient or the theme

roles, and these participants are crucial in order to understand the meaning of the predicate. Among the rest of the arguments, a PP in Spanish can realize an ArgM, Arg0 and Arg2 and with less frequency, an Arg3 and an Arg4.

SubC in Russian tend to realize the following core arguments: Arg0, Arg1 and Arg2. However, this is not a common complement for a deverbal noun. It is very residual in our sample. In Spanish, SubC are not considered argumental.

In the case of two arguments combined in Russian the most usual combinations are AP+NP and AP+AP with the commonest combinations of Arg1+ArgM and Arg0+ArgM. The usual constituent combinations in Spanish are AP+AP and PP+PP. In most of the cases one of these two constituents expresses the Arg1 and the other can be either an Arg0 or an Arg2. This work is the basis to study translation mismatches between Russian and Spanish in deverbal noun constructions (see chapter 5).

Chapter 5

Translation Mismatches between Russian and Spanish deverbal nouns

In this section, we present a classification of regular and productive mismatches between Russian and Spanish deverbal nouns, obtained by means of corpus-based analysis of data extracted from RuSp and UNGAR parallel corpora. Previous classifications are devoted to the verbal predicate, whereas our study is focused on the deverbal noun.

Translation mismatches are a challenging question in machine translation, specially, for those systems transfer-based but also for those statistical-based, which need parallel corpora. The detection and classification of mismatches helps to solve translation divergences occurring in the automatic process of alignment of these parallel corpora.

Our classification focuses on general deverbal noun mismatches and includes specially those linguistic phenomena, which are productive and regular.

A mismatch rarely consists of a single linguistic change, because of that we propose a description of mismatches as sets of linguistic changes. We describe and classify mismatches depending on the linguistic changes produced.

The analysis of translation mismatches serves also as a validation of our previous observations on denotation (see chapter 3). For instance, the choice of the category (that is, a verb or a noun) in the translation of a deverbal noun is related to the lexical reading of that deverbal noun. A deverbal noun translated by a verb is interpreted as an event, whereas a noun translated by a non-deverbal noun is interpreted as a result. This translation mismatch gives support to the idea about the hybrid nature of deverbal nouns, those nouns translated into verbs are more closely related to the verbal category, whereas those translated into non-deverbal nouns are close to the common nouns. This translation mismatch can be used as a criterion to distinguish between the lexical denotations of deverbal nouns.

This chapter is structured as follows. In the first section, we briefly describe the used corpora. In the second section, we introduce the properties of translated texts. In section three, we present a brief summary describing the features of previous classifications of mismatches. In section four, we propose our classification of translation mismatches. And, finally, in the last section we present our conclusions.

5.1 Brief notes on the corpora

As we have already seen in chapter 2, section 2.2, RuSp is a parallel corpus composed of translations into Spanish of Russian literary texts, and UNGAR is a multilingual corpus of administrative texts composed of the resolutions of the United Nations. In RuSp, the original texts are written in Russian (source language) and translated into Spanish (target language). In UNGAR, the source language (SL) is English while Spanish and Russian are both target languages (TL). In the case of the administrative corpus, translations tend to be more literal with respect to the source language. In fact, beyond natural typological differences, Russian and Spanish translations are also close to each other in UNGAR. On the other hand, in the case of the literary corpus, translations tend to be freer with respect to the original text, since sometimes the translator uses stylistic licenses in order to provide the reader with the taste of the source text.

Most of the examples used in our study are extracted from a subsample of RuSp, MiniRuSp, which contains 500 deverbal nouns contextualized corresponding to 114 different lemmas. MiniRuSp has been analyzed syntactically with constituent and syntactic function structures, and semantically with the argument structure and the denotation of deverbal nouns. The syntactico-semantic analysis of deverbal nouns is the base upon which we analyze translation mismatches (see chapter 4, section 4.2). UNGAR has not been annotated and it is mostly used to check particular words and constructions. Moreover, UNGAR is not the most appropriate source of our analysis since its SL is not Russian but English.

5.2 Properties of translated texts

Translation is a complex task, which combines an act of interpretation and a technique to transform an input into an output. In a translation, translator decodifies the meaning of a text in one language (that is, the SL) and subsequently generates a more or less equivalent text in another language (that is, the TL). Translated texts have special properties, some of them universal, irrespective of the SL and holding for any TL. Baker (1995) proposes four properties:

1. **Simplification**: Translations tend to use simpler language than non-translated texts in the same TL.

2. **Explicitation**: Translated texts show a tendency to spell things out rather than leave them implicit.

3. **Leveling out:** In a collection of translations compared to a collection of comparable original texts in the same language as the TL, the individual texts in the translations are more similar to each other than the individual texts in the set of original texts.

4. **Normalization**: Translations have a tendency to conform to the typical patterns of the TL and to exaggerate the features of the TL.

As we will see in section 5.4, Baker's explicitation feature is correlated with a mismatch found in our analysis. In the process of explicitation, the translator adds information to the target text in order to make the text comprehensible to the reader. This added information can be grammatically required by the TL or needed to explain a concept or a real-world entity that only exists or has a word in the SL. For example, the Russian words *moroz* and *stuzha* 'cold below zero' are words that reflect a concept that does not exist in Spanish.

In our approach, we keep in mind that a translator must choose between different possible translations, one among a number of possibilities, as a result of a subjective interpretation. As an obvious consequence of this, it does not exist a unique translation, human language has a variety of forms to express the same information, which is an intrinsic part of the creative process. For instance, in (1), the same sentence, extracted from Flaubert's novel *Madame Bovary*, is translated in two different ways.

(1)
(i) *Emma* ***pleurait****, et il s'efforçait de la* ***consoler, enjolivant de calembours ses protestations.***
(ii) a. *Emma* ***burst into tears*** *and he tried to* ***confort*** *her,* ***saying things to make her smile.***
b. *Emma* ***cried****, and he tried to* ***console*** *her,* ***adorning his words with puns****.*

[Barzilay and McKeown, 2001:50]

The wide range of possible translations is specially relevant when we discuss translation mismatches, since we may present as an example of mismatch a piece of text that could have been translated without provoking that mismatch, that is, the translator could have chosen a closer translation to the source text. In (1), a closer translation to the source text is *(ii.b)*.

The variety of translations is due to the fact that a translation is the result of a creative process in which the transfer of a text into another language takes place. Translation is a complex process, not a straightforward transmission of semantic and lexicogrammatical features into TL, which, according to Steiner (1991), involves two subprocesses: (i) de-metaphorization and (ii) rewording. The process of de-metaphorization involves relating "meaningful grammatical units to some of their less metaphorical variants". After the de-metaphorization, translator rewords the understood source text into the target text. It is in this sense that Steiner considers

the translation as a semantic paraphrase. It is in this rewording process in which translation disagreements take place. Steiner (1991) claims that in example (2) the German translation is non-metaphorical relative to its English counterpart.

(2)
(i) The fifth day saw them in the summit.

(ii) *Am fünften Tag kamen sie am Gipfel an.*
'At the fifth day came they at the summit.'

We study these types of disagreement to establish the different types of mismatches between Russian and Spanish regarding deverbal nouns. In this sense, translation can be understood as a type of paraphrase, that is, as a way to convey the same information in a different form. Following Mel'chuk & Wanner (2006), we can distinguish between intralinguistic and interlinguistic mismatches. In the case of intralinguistic mismatches, paraphrase takes place in the same language. For instance, in Spanish there are two constructions to express the action of stabbing, that is, the single verb *apuñalar* and the light verb construction *dar una puñalada*. While in interlinguistic mismatches, paraphrase takes place between different languages. For instance, in Russian 'fishing' is expressed by two words, that is, *rybnaja lovlja* 'fish catching', whereas in Spanish the same concept is expressed by the single word *pesca*. Both intra- and interlinguistic mismatches can be seen as semantic paraphrases (Steiner, 2001), since the same information is conveyed in different ways. We understand an event deverbal noun construction as an intralinguistic paraphrase, since the same meaning could have been conveyed by means of a verbal predicate, and mismatches between Russian and Spanish constructions as interlinguistic paraphrases. It is relevant to notice that paraphrasing is not conveying the exactly same meaning. As Vila, Martí & Rodríguez (2013:9) point out "paraphrasing must be situated in the field of approximation, opening the path to different semantic similarity or paraphrasality degrees". Following this path, we conceive translation mismatches as different degrees of semantic similarity between SL and TL. At one edge of the continuum, we would have a word-for-word translation of a sentence, while at the opposite edge we would have a freer translation, that is, a continuum from more to less literal translation.

5.3 Previous classifications of translation mismatches

Most of the classifications presented in this section are oriented to the systematization and representation of translation mismatches between different languages in order to treat them in a Machine Translation (MT) system. In fact, one of the major challenges to be solved in this type of Natural Language Processing application are translation mismatches. These proposals followed different linguistic

theoretical approaches: Dorr (1990, 1994) proposal is influenced by the Lexical Conceptual Structure (Jackendoff, 1990); Mel'chuk & Wanner (2006) proposal is based on Meaning-Text Theory (Mel'chuk, 1981); and Fernández (2000) and Vázquez, Fernández & Martí (2000) representation is a more eclectic proposal based on different lexical-semantic approaches (Jackendoff, 1990; Talmy, 1995, basically). Dorr and Mel'chuk & Wanner classifications are more focused on changes and differences in the morpho-syntactic structure, whereas Fernández (2000) and Vázquez, Fernández & Martí (2000) classifications are focused on lexical-semantic differences. These proposals are presented briefly below.

Dorr (1990, 1994) presents a classification of translation mismatches linguistically grounded, with the aim to describe them formally for their treatment in an interlingual MT system. The interlingual representation is based on lexical-conceptual structures (Jackendoff, 1983, 1990) and it is implemented in a system called UNITRAN[41]. Later on Dorr *et al.* (2004) use this classification to solve translation divergences occurring in the process of alignment of parallel corpus. Concretely, they build a word-level alignment algorithm to improve the quality of statistical MT systems. For these authors, translation divergences are "structural differences that occur when the underlying meaning of a set of words in one language is distributed over a set of words with different syntactic and lexical semantic properties in the other language. Translation mismatches are not disjoint, *i.e.* they frequently co-occur." (Dorr *et al.* 2004:2-3). In this work, they propose the following translation divergence types:

(1) Categorial Divergence: A categorial divergence is the translation of a word in one language into a word, which has a different part of speech. In (3) the English adjective *jealous* is translated into the Spanish noun *celos* 'jealousy'.

(3)
(i) *to be* ***jealous***$_{\text{adjective}}$
(ii) *tener* ***celos***$_{\text{noun}}$
'to have jealousy'

(2) Conflational/Inflational Divergence: A conflation is the translation of two or more words into one word in the other language. Inflation is the reverse image of conflation. Common forms of this divergence are light-verb constructions and manner conflations. In (4), a light-verb construction, that is, a combination of semantically "light" verb and some other meaning unit (NP or PP) is the Spanish translation of a single verb in English. In (5), we observe a manner conflation, that is,

[41] The name of UNITRAN stands for Universal Translator. This system serves as the basis for translation across a variety of languages.

the translation of a single manner verb in English into a light verb of motion and a manner in Spanish.

(4)
(i) *to* ***kick***
(ii) ***dar una patada***
'to give a kick'

(5)
(i) *to* ***float***
(ii) ***ir flotando***
'to go via floating'

(3) Structural Divergence: A structural divergence is the realization of verb arguments in different syntactic configurations in different languages. For example, in (6) the verbal object is realized as the NP *the house* in English and as a PP *en la casa* 'into the house' in Spanish.

(6)
(i) *to enter* [***the house***]$_{NP}$
(ii) *entrar* [***en la casa***]$_{PP}$
'to enter in the house'

(4) Head Swapping Divergence: Head swapping is the inversion of a structural dominance relation between two semantically equivalent words when translating from one language to another. An example is the demotion of a head verb and the promotion of its modifiers to the head position. In example (7), what in English is expressed by a verb and a preposition in Spanish corresponds to a complex verbal construction with a verb and a gerund. The gerund indicates the manner and the main verb the direction. In English the direction was expressed by means of the preposition *in* and the manner of the movement by means of the verb.

(7)
(i) *to* ***run in***
(ii) ***entrar corriendo***
'to enter running'

(5) Thematic Divergence: A thematic divergence occurs when arguments are realized in different syntactic configurations that reflect the same thematic hierarchies (thematic to syntactic mapping orders). In (8), in the English structure the subject is the experiencer and the object the theme, on the other hand, in the Spanish structure the subject is the theme and the object is the experiencer.

(8)
(i) *I like* ***grapes.***

(ii) ***Me*** *gustan* ***las uvas.***
'I_{IO} like $grapes_{Subj}$'

Mel'chuk & Wanner (2001, 2006) classification of translation mismatches is framed in Meaning-Text Theory (MTT, Mel'chuk, 1981) and it is aimed at formalizing rules of transfer component in MT systems. These authors are mainly focused on syntactic mismatches. Taking into account Dorr's proposal, they present a similar mismatch classification adapting basically the terminology to the framework of Meaning-Text Theory. Correspondences between Mel'chuk & Wanner's proposal and Dorr's proposal (in brackets) are the following:

(1) Mismatches due to part-of-speech changes (*Categorial divergence*).

(2) Mismatches due to lexeme-phrase substitution, or lexical fission/fusion (*Lexical conflational/ inflational divergence*).

(3) Mismatches due to function-word introduction/elimination (*Structural divergence*).

(4) Mismatches due to dependency inversion, or head switching (*Head swapping divergence*).

(5) Mismatches due to syntactic actant permutation, or conversion (*Thematic divergence*).

Mel'chuk & Wanner (2006:83) claim that "semantically equivalent syntactic structures within one language (paraphrases) reveal mismatches of the same kind as those identified between equivalent syntactic structures across languages". Therefore, interlinguistic mismatches are also considered paraphrases and, as in the case of other authors, they consider that paraphrases are not fully synonymous.

The classification of mismatches proposed by Fernández (2000) and Vázquez, Fernández & Martí (2000) is more semantically oriented. Their proposal aims at the lexical-semantic representation of translation mismatches between Spanish and English. This classification is based on the model for lexical description appeared in Vázquez *et al.* (2000), which takes into account the meaning components, the argument structure and the event structure. Concretely, Fernández (2000) creates a

"theoretical framework for the semantically motivated description of verbal lexical items established from a multilingual prespective". Moreover, this theoretical based model is implemented in a Lexical Knowledge Base [42], demonstrating the computational validity of the model. Their proposal of classification is the following:

(1) Mismatches based on meaning components are those that are due to different lexicalization patterns:

(i) Lexicalization of a meaning component, for instance, *to fax* is translated into *enviar un fax* 'to send a fax'. The main difference here is that in English the manner meaning component in which the action is carried out is incorporated in the word; while in Spanish is analytically expressed. Different meaning components can be incorporated: manner, affected entity, cause and instrument.

(ii) Incorporation of a meaning component by pronominalization, for instance, *ir/irse* are translated into *go/leave* respectively. English cannot incorporate the pronoun then it uses a different lexical item. Pronominal Spanish verb includes the source.

(iii) Incorporation of a meaning component by derivation, for instance, *oír mal* is translated into *mishear*. In the English verb the manner component has been incorporated by means of the prefix *mis-*, that is, by means of a derivation process.

(iv) Conflation of a meaning component through the context, for instance, *dijo sonriendo* is translated into *he laughed*. In the English verb, the manner component is incorporated along with the main action of saying. The context in which the structure *he laughed* appears tells us about the fact that he is not only laughing but also saying something.

(v) Mismatches based on lexical semantic composition: In this type of mismatch, a language has gone through a process of lexical semantic composition which ends in a construction that has no direct equivalent in the other language. For instance, *He ran out of the room* is translated into *salió corriendo de la habitación*. In English the verb *run out* is a complex element that expresses *movement, manner* and *path*. In Spanish it is needed a movement verb denoting *movement* and *path*, that is, *salir* and an adjunct denoting *manner*, that is, *corriendo*.

(2) Mismatches based on the argument structure: These mismatches result from the interaction between the verb and its arguments. In these mismatches, the meaning components are differently distributed in the source language and in the target language. Authors divide these mismatches in two types:

[42] "Knowledge Bases are tools used to store information in a structured manner. The information is stored by means of a symbolic representation or formalization of the objects that are conceived as conceptual entities." (Fernández, 2000:280)

(i) Simple mismatches:

a. Different distribution of meaning components: The resulting sentences in both languages are semantically equivalent but syntactically they realize the participants in switched positions. In (9), what in English is the participant that is the subject, in Spanish is the object, and the other way round.

(9)
(i) [***He***]$_{\text{NP-Subj}}$ *likes reading* [*mystery* ***books***]$_{\text{NP-DO}}$

(ii) [***Le***] *gusta leer* [***novelas de misterio***]$_{\text{NP-Subj}}$

b. Different realization of argument-2: In this mismatch, the structure in one language is transitive but in the other is intransitive. They have the same number of arguments, but they are realized syntactically in a different way. In fact, the intransitivity depends on the realization of one argument as a prepositional phrase. In English the object is expressed by means of a NP (10.i), whereas in Spanish it is expressed by means of a PP (10.ii).

(10)
(i) *He entered* [***the room***]$_{\text{NP}}$

(ii) *Entró* [***en la habitación***]$_{\text{PP}}$

c. Different expression of possession: This type of mismatch can be found in constructions that express possession. An object or person can be viewed as a possessor of properties, qualities or characteristics. In (11), we observe that in Spanish the possessor and the possessed object are expressed in two syntactic constituents, whereas in English the possessor and the possessed object are expressed in the same constituent.

(11)
i. ***Me*** *lo he dejado* ***en el coche.***

ii. *I left it* ***in my car.***

(ii) Mismatches corresponding to diathesis alternations:

a. Mismatches regarding causativity: In this mismatch a verb is characterized by the presence of two alternative argument structures that denote the opposition cause-anticausative (12 and 13, respectively) and "the way in

which these verbs express the different structures can vary in each language" (Vázquez, Fernández & Martí, 2000:7). For instance, in Spanish the verb uses a pronoun to express the anticausativity (13.i), whereas in English a periphrastic construction is required (13.ii).

(12)
(i) *Mirar* *la* *televisión* *me* *aburre.*
TO_WATCH THE TELEVISIÓN ME BORES
'Watching TV bores me.'

(ii) *Watching TV bores me.*

(13)
(i) *Se* *aburrió.*
-- BORED
'He got bored.'

(ii) *He got bored.*

b. Mismatches based on different order: In this mismatch, one language permits a change in the order to relocate the focus, whereas the other language does not change the order but requires a more complex syntactic construction. For instance:

(14)
(i) *Sus* *compañeros* *han* *regalado* *un* *reloj* *a* *Juan*
HIS COLLEAGUES HAVE GIVE A WATCH TO JUAN
por *su* *cumpleaños.*
FOR HIS BIRTHDAY
'His colleagues have given a watch to John for his birthday.'

(ii) *His colleagues have given John a watch for his birthday.*

(15)
(i) *A* *Juan* *le* *han* *regalado* *un* *reloj* *por* *su*
TO JUAN -- HAVE GIVEN A WATCH FOR HIS
cumpleaños.
BIRTHDAY
'John has been given a watch for his birthday.'

(ii) *He has been given a watch for his birthday.*

(3) Mismatches due to a different event structure:

i. Mismatches based on aspectual information: In this mismatch, there is "a divergence in the temporal information reflected by the verb" (Vázquez, Fernández & Martí, 2000:9). For instance, in English we can express iteration by means of the verbs *to strike* or *to beat* in (16), whereas we are expressing a single action by means of the verb *to hit*, in (17). In Spanish there is no such variety of verbs to express these aspectual differences, but this aspectual information is expressed by means of different constructions with the same verb. Therefore, the verb *golpear* is used to translate *hit, strike* and *beat*, however, to express the aspectual meaning of iteration is necessary to use a periphrastic construction, such as (16.iii) while the single action can be expressed with the following constructions (17.ii and 17.iii).

(16)
(i) *He struck John*

(ii) *He beat John*

(iii) *Estuvo golpeando a John*

(17)
(i) *He hit John*

(ii) *Dio un golpe a John*

(iii) *Golpeó a John*

5.4 A classification of Russian-Spanish deverbal nouns mismatches

This section focuses on the linguistic characterization of deverbal noun mismatches. The main difference between our proposal and the previous ones is that they are devoted to the verbal predicate, whereas our study is focused on the deverbal noun, with the goal of establishing the different types of interlinguistic mismatches between Russian and Spanish. Although our approach is based on these previous works, it is worthy to note that not all the phenomena described for verbs take place in deverbal nouns. Our classification focuses on general deverbal noun mismatches and includes specially those linguistic phenomena, which are productive and regular. We describe and classify mismatches depending on the linguistic changes produced rather than on the reasons behind these linguistic changes -typological, pragmatic, cultural or subjective.

In the sample analyzed we have confirmed Dorr's claim about the co-occurrence of mismatches, since we have observed that a mismatch rarely consists of a single linguistic change, the commonest is to have more than one linguistic change at

different levels (morphological, syntactic, lexical-semantic and pragmatic). We describe mismatches as sets of linguistic changes and in a further analysis, based on the present approach they could be represented as a feature structure, where each linguistic change would correspond to a different feature[43].

The proposed classification has been empirically validated by means of the analysis of 500 pairs of Russian-Spanish deverbal contextualized nominalizations (which correspond to 114 different lemmas) extracted from the parallel corpus MiniRuSp. The analysis has followed two phases: in a first stage we have analyzed 100 Russian deverbal nouns and their Spanish translations to propose a preliminary list of linguistic changes. In the second stage, we have analyzed the whole sample, that is, 400 deverbal nouns have been analyzed in order to complete the list of linguistic changes and to propose the final classification.

In the following section, firstly, we describe linguistic changes individually (see section 5.4.1); secondly, we present the classification of mismatches depending on the number and type of linguistic changes involved (see section 5.4.2); and finally we discuss the results obtained (see section 5.4.3).

5.4.1 Linguistic changes in mismatches

We define linguistic changes considering previous works on translation mismatches applicable to Russian and Spanish particularities.

From the analyzed sample we have organized linguistic changes in 8 different types, taking into account the level of the linguistic structure involved: (1) determiner change, (2) PoS change, (3) structural change, (4) argument structure change, (5) lexical change, (6) head swapping change, (7) number change, and (8) discourse change. In figure 1, we present the list of linguistic changes and after that, we describe them in more detail.

[43] A feature structure is a matrix of a pair attribute-value. Features are pairs of the form <A=B>, where A is the name of an attribute and B the assigned value. Attributes are labels that describe the type of information described by the value.

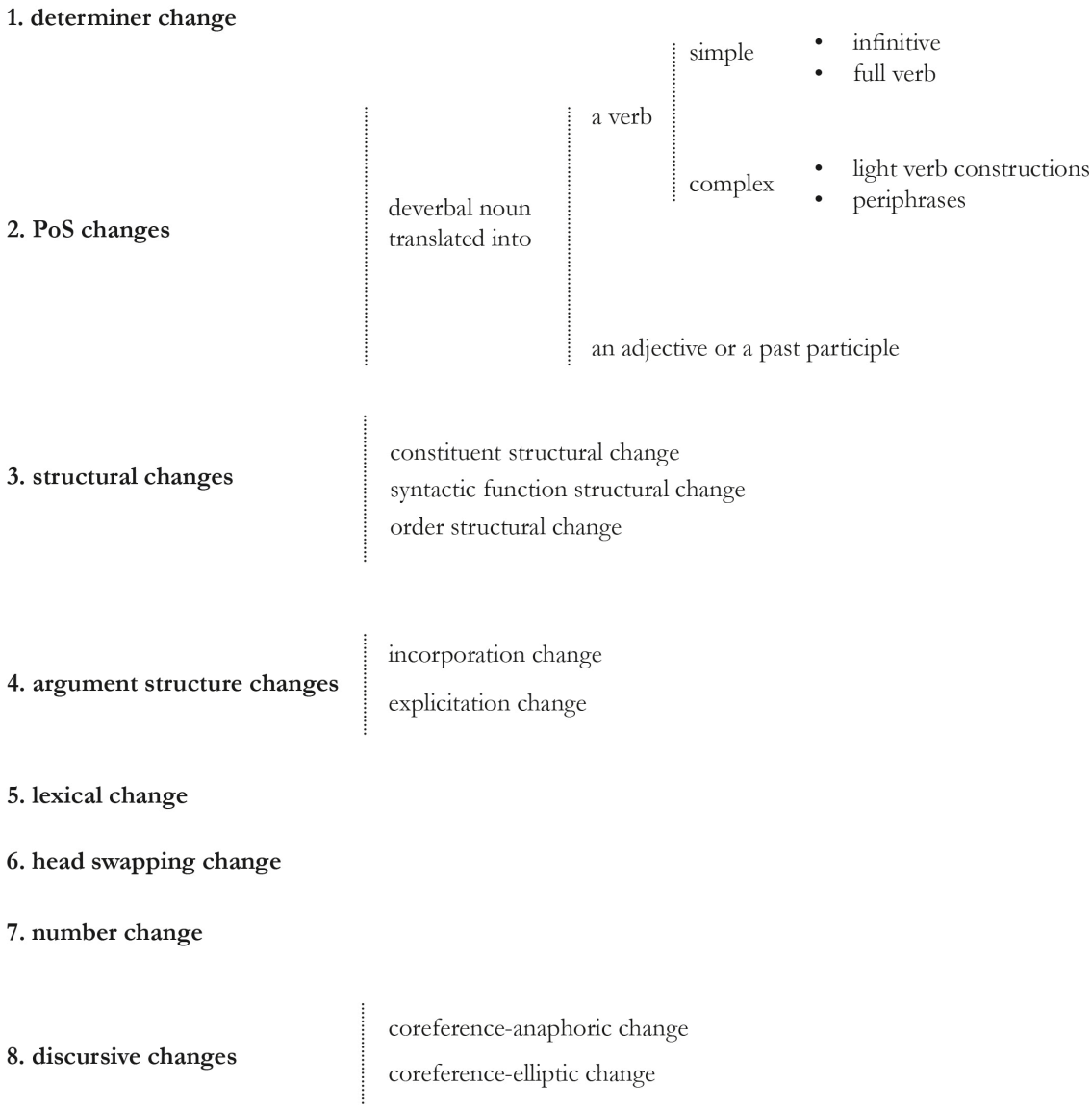

Figure 1: Linguistic changes

(1) Determiner change: The determiner change is a change motivated by typological reasons, due to the fact that Russian does not have articles, whereas Spanish does. The addition of an article is the commonest type of determiner change, though other types of determiner changes can take place, for instance, the addition of a possessive, an undefined or a demonstrative determiner, among others.

(18)

(i) *prosit'*	[***blagoslovenija***]$_{NP}$	*u*	*mužika*
TO_ASK	**BLESSING**	TO	MUZHIK

'to ask a muzhik for the **blessing**'

(ii) *pedir* [[*la*]$_{det}$ ***bendición***]$_{NP}$ *a* *un* *mujik*
TO_ASK THE **BLESSING** TO A MUZHIK
'to ask a muzhik for the **blessing**'

(19)
(i) *Ja* *ne* *khoču* *davat'* [***razrešenie*** [*na* *brak*]$_{PP}$]$_{NP}$
I NOT WANT TO_GIVE **PERMISSION** TO MARRIAGE
'I don't want to give **permission** to the marriage.'

(ii) *Yo* *no* *quiero* *dar* [[*mi*]$_{Poss}$ ***consentimiento*** [*para* *que*
I NOT WANT TO_GIVE MY **BLESSING** FOR THAT
se *casen*]$_{PP}$]$_{NP}$
– MARRY
'I don't want to give them my **blessing** to marry.'

(20)
(i) *vydači* [***znanij***]$_{NP}$
DELIVERY **KNOWLEDGE**
'**knowledge** transmission'

(ii)*exponer* [[*algún*]$_{Indef}$ ***conocimiento***] $_{NP}$
TO_PRESENT SOME **KNOWLEDGE**
'to present some **knowledge**'

In (18), the deverbal noun *bendición* 'blessing' is specified with the article *la* 'the'. In (19), the deverbal noun *razrešenie* 'permission' is specified with the possessive determiner *mi* 'my'. In (20), the deverbal noun's complement is specified with the undefined determiner *algún* 'some'.

Cases with no determiner change are due to the fact that the Russian deverbal noun is already specified by a demonstrative, a possessive determiner (21) or because Spanish nouns can appear as bare nouns (22).

(21)
(i) [[*Naši*]$_{Poss}$ ***znakomstva***]$_{NP}$ *ukrepljalis'*
OUR **ACQUAINTANCES** GOT_STRONGER
'Our **relations** got stronger.'

(ii) [[*Nuestras*]$_{Poss}$	***relaciones***]$_{NP}$	*se*	*fortalecían*
OUR	**RELATIONS**	--	STRENGTHEN

'Our **relations** got stronger.'

(22)

(i) *Doma*	*-ne*	*samyj*	*nadežnyj*	*vid*	[*vklada*]$_{NP}$
HOUSES	-NOT	MOST	SAFE	TYPE	**INVESTMENT**

'Houses are not the safest type of **investment.**'

(ii) *Las casas*	*no*	*eran*	*el*	*mejor*	*tipo*	*de*	[***inversión***]$_{NP}$
THE HOUSES	NOT	WERE	THE	BEST	TYPE	OF	**INVESTMENT**

'Houses were not the best type of **investment.**'

In a very residual way, we have found some cases in which the specifier is simplified in the translation, that is, while in the original we have a quantifier and a determiner in the target text we only have the determiner (24).

(24)

(i) *Izorval*	*v*	*kločja*	[[*vse*]$_{Indef}$	[*moi*]$_{Poss}$]$_{SPEC}$	***zapisi***]$_{NP}$
TORE	IN	PIECES	ALL	MY	**NOTES**

'He tore into pieces all my **notes**.'

(ii) *Hizo*	*pedazos*	[[*mis*]$_{Poss}$	***apuntes***]$_{NP}$
MADE	PIECES	MY	**NOTES**

'He tore into pieces my **notes**.'

In MiniRuSp, the determiner change has a frequency of 34.17%.

(2) PoS change (Dorr, 2004: *Categorial Divergence*): This feature indicates whether there is a morphosyntactic categorial change. If there is no PoS change, it means that a noun (either deverbal or not) is used in both languages. If there is a PoS change, it means that a different morphological category has been used in the target language. According to the sample analyzed, there are two types of PoS change: (a) nouns that are translated into verbs, (b) nouns that are translated into adjectives or past participles.

a. Deverbal noun to a verb: The verbal form used in the transaltion of a deverbal noun can be of two types: simple -infinitive and inflected forms (25 and 26, respectively) – and complex -light verb constructions and periphrases (27 and 28, respectively).

(25)

(i) *sposobnost'*	[***vydači***$_{\text{noun}}$	[*znanij*]$_{\text{NP}}$]$_{\text{NP}}$	
SKILL	**DELIVERY**	KNOWLEDGES	

'the skill of knowledge **transmission**'

(ii) *la*	*facultad*	*de*	[***exponer***$_{\text{inf}}$	[*algún*	*conocimiento*]$_{\text{NP}}$]$_{\text{InfC}}$
THE	SKILL	OF	**TO_PRESENT**	SOME	KNOWLEDGE

'the skill **to present** some knowledge'

In example (25), the Russian deverbal noun *vydača* 'delivery' is translated into the Spanish infinitive *exponer* 'to present'.

(26)

(i)*Ždal*	[[*eë*]$_{\text{Poss}}$	***vozvraščenija***$_{\text{noun}}$]$_{\text{NP}}$
WAITED	HER	**RETURN**

'He waited for her **return**.'

(ii) *Esperaba*	*a*	*que*	[[*ésta*]$_{\text{NP}}$	***apareciera***$_{\text{verb}}$]$_{\text{VP}}$
WAITED	FOR	THAT	SHE	**APPEAR**

'He waited for her to **appear**.'

In example (26), the Russian deverbal noun *vozvraščenije* 'return' is translated into the inflected form of the verb *aparecer* 'to appear'.

(27)

(i) *Ja*	*prognal*	*nabegajušij*	*son*	[***naprjaženiem***$_{\text{noun}}$	[*vsego*	*tela*]$_{\text{NP}}$]$_{\text{NP}}$
I	BANISHED	INVADING	SLEEP	**TENSION**	ALL	BODY

'I shook off the sleep that was overcoming me by **tensing** my whole body.'

(ii) *Me*	*deshice*	*del*	*sueño*	*que*	*me*	*invadía*	[***poniendo***
ME	BANISHED	THE	SLEEP	THAT	ME	INVADED	**PUTTING**
en tensión$_{verb}$		[*todo*	*el*	*cuerpo*]$_{NP}$]$_{VP}$			
IN TENSION		ALL	THE	BODY			

'I shook off the sleep that was overcoming me by **tensing** my whole body.'

In example (27), we see that the Russian deverbal noun *naprjaženie* 'tension' is translated into a light verb construction *poniendo en tensión* 'putting in tension' in Spanish.

(28)

(i) [***vozvraščenie***$_{noun}$	[*k*	*razdeleniju*]$_{PP}$]$_{NP}$	
RETURN	TO	DIVISION	

'a **return** to the division'

(ii) [***volver***	a	establecer$_{verb}$	[*la*	*diferencia*]$_{NP}$] $_{VP}$	
RETURN	TO	ESTABLISH	THE	DIFFERENCE	

'to establish the difference **again**'

In example (28), we see that the deverbal noun *vozvraščenie* 'return' is translated into a verbal periphrasis *volver a establecer* 'to establish again' in Spanish.

b. Deverbal noun to an adjective or to a past participle: In this type of mismatch, deverbal nouns are translated into adjectives (29 and 30), usually deverbal adjectives (29), or past participles (31).

(29)

(i) *Eto*	*bylo*	*nebol'šoe*	*i*	*nestrašnoe*		*sobytie*	*po*
THIS	WAS	LITTLE	AND	NOT_TERRIBLE		FACT	IN
[***sravnenininju***$_{noun}$	[*s*	*tem,*	*što*	*on*	*videl*]$_{PP}$]$_{NP}$		
COMPARISON	WITH	THAT,	THAT	HE	SAW		

'This was not a terrible fact in **comparison** with the one that he saw.'

(ii) *Y*	*por*	*irrelevante*	*que*	*hubiera sido*		[*el*	*hecho,*	***incomparable***$_{adj}$
AND	FOR	IRRELEVANT	THAT	HAD	BEEN	THE	FACT	**INCOMPARABLE**

[*con*	*los*	*horrores*	*que*	*había*	*presenciado*]$_{PP}$]$_{NP}$
WITH THE		HORRORS	THAT	HAD	SEEN

'and even if the fact was irrelevant and **incomparable** with the horrors that he had seen'

In (29) the Russian deverbal noun *sravnenije* 'comparision' is translated into the deverbal adjective *incomparable* 'incomparable'.

(30)

(i) *Krist*	*zamer*	*v*	[*ožidanii*$_{noun}$]$_{NP}$
KRIST	TO_COME_TO_STANDSTILL	IN	WAITING

'Krist froze in expectance.'

(ii) *Krist*	*se*	*quedó*	*quieto,*	[*expectante*$_{adj}$]$_{AP}$
KRIST	-	REMAIN	STILL	EXPECTANT

'Krist remained still, expectant.'

In (30), the Russian deverbal noun *ožidanie* 'waiting' is translated into the adjective *expectante* 'expectant'[44], in Spanish.

(31)

(i)[*posle*	***okončanija***$_{noun}$	[*kursov*] $_{NP}$]$_{NP}$
AFTER	**END**	COURSES

'after **ending** the courses'

(ii)[***acabados***$_{past\ part}$	[*éstos*]$_{NP}$]$_{PartC}$
ENDED	THESE

'**ended** them'

In (31), the Russian deverbal noun *okončanie* 'end' is translated into the participle *acabados* 'finished'.

[44] The adjective *expectante* 'expectant' has its origin in the Latin *expectans*, *-antis* which is the active participle of *exspectāre* 'to observe'.

Having studied all the cases of PoS change found in our sample, we observed the following facts.

1) The majority of Russian deverbal nouns (88.78%) are translated into Spanish nouns (most of the times a deverbal noun). When translated into a different category, the preferred category is the verb (83.69%). Regarding the rest of categories, we have obtained a past participle in 10.86% of cases and an adjective in 5.43% of cases.

2) In those cases in which the noun has been translated into a verb, the commonest form is the infinitive (53.24%), followed by full verbs (29.87%), light verb constructions (14.28%), and finally, periphrases (2.59%).

3) The choice of a verb or a noun in the translation of deverbal nouns is related to the lexical reading of the deverbal noun. As we have observed, deverbal nouns translated into verbs were interpreted as events, whereas those translated into non-deverbal nouns were interpreted as results.

(3) Structural change (Dorr, 2004: *Structural Divergence*): In a structural change the number of explicit arguments realized in the NP headed by the deverbal noun is the same but they are realized in different syntactic configurations or with different syntactic functions. We distinguish between (a) constituent structural change, (b) functional structural change and (c) order change.

a. Constituent structural change: In this subtype of structural change, the same number of explicit arguments is realized but in different syntactic constituents. This change may be due to a typological reason: Russian has 6 grammatical cases (nominative, accusative, genitive, dative, instrumental and prepositional) and morphological case can be expressed by means of inflectional suffixes and prepositions. On the other hand, Spanish has not grammatical case and it uses prepositions instead. For this reason what is expressed by means of a complement NP in Russian it is usually translated into a PP in Spanish (32). It also can be translated into an AP (33) or, even, into a RelC (34).

(32)

(i) *rasširenija*	[***vklada***	[*Agentstva*]$_{\text{NP-GEN-NC-Arg0-agt}}$	[*v*	*delo*	
INCREASING	**CONTRIBUTION**	AGENCY	IN	DEAL	
ulučšenija	*uslovij*	*dlja*	*bežentsev*]$_{\text{PP}}$]$_{\text{NP}}$		
IMPROVEMENT	CONDITIONS	FOR	REFUGEES		

'the increasing of **contributions** of the Agency regarding the task of improving the conditions of refugees'

(ii) *incrementar* [*las* ***contribuciones*** [*del* *Organismo*]$_{\text{PP-NC-Arg0-agt}}$
TO_INCREASE THE **CONTRIBUTIONS** OF_THE ORGANISM
[*al* *mejoramiento* *de* *las* *condiciones* *de* *los* *refugiados*]$_{\text{PP}}$]$_{\text{NP}}$
TO_THE IMPROVEMENT OF THE CONDITIONS OF THE REFUGEES
'to increase the **contributions** of the organism to the improvement of refugees' conditions'

In example (32), the argument of Russian deverbal noun *vklad* 'contribution' which is realized by means of a NP in genitive, that is, *agentsva* 'agency' is translated into a PP introduced by the non-marked preposition *de* 'of', that is, *del organismo* 'of the organization', in Spanish.

(33)
(i) [***delenija*** [*khromosom*]$_{\text{NP-GEN-NC-Arg1-pat}}$]$_{\text{NP}}$
DIVISION CHROMOSOMES
'the **division** of chromosomes'

(ii)[*la* ***división*** [*cromosómica*] $_{\text{AP-NC-Arg1-pat}}$] $_{\text{NP}}$
THE **DIVISION** CHROMOSOMATIC
'the chromosomatic **division**'

(iii) [*la* ***división*** [*de* *los* *cromosomas*]$_{\text{PP- NC-Arg1-pat}}$]$_{\text{NP}}$
THE **DIVISION** OF THE CHROMOSOMES
'the **division** of chromosomes'

In (33), the argument realized by means of a NP in genitive is translated into an AP, that is, from *delenie khromosom* 'division of chromosomes' into *la división cromosómica* 'chromosomic division'. However, we have found another occurrence of this construction translated into a PP (iii).

(34)
(i)[*beznadežnym* ***sostojaniem*** [*otca*]$_{\text{NP-GEN-NC-Arg0-exp}}$]$_{\text{NP}}$
PITIFUL **CONDITION** FATHER
'the pitiful **condition** of the father'

(ii)[la lamentable **situación** [en que se hallaba su
THE PITIFUL **SITUATION** IN WHICH -- WAS HIS
padre]$_{RelC\text{-}NC\text{-}Arg0\text{-}exp}$]$_{NP}$
FATHER

'the pitiful **condition** in which his father was'

In (34), the argument of the deverbal noun *sostojanie* 'condition' has been translated into a RelC, that is, *en que se hallaba su padre* 'in which his father was'.

(35)
(i) [***Zavedovanie*** [*otdeleniem*] $_{NP\text{-}INS\text{-}NC\text{-}Arg1\text{-}pat}$] $_{NP\text{-}Subj}$ *pereš̆lo* *k* *doktoru*
DIRECTORSHIP SECTION PASSED TO DOCTOR
Zaderu
ZADER

'The **directorship** of the section was passed to doctor Zader.'

(ii)[*La* ***jefatura*** [*de* *la* *sección*]$_{PP\text{-}NC}$]$_{NP\text{-}Subj}$ *pasó* *al*
THE **DIRECTORSHIP** OF THE SECTION PASSED TO-THE
doctor *Zader*
DOCTOR ZADER

'The **directorship** of the section was passed to doctor Zader.'

Regarding the NP in dative or in instrumental, it is translated into a PP (35). In (35), the argument of the deverbal noun, that is, the NP in instrumental *otdeleniem* 'of the section' is translated into a PP introduced by the non-marked preposition *de* 'of', that is, *de la sección* 'of the section'.

(36)
(i) [[*blatnye*]$_{AP\text{-}NC\text{-}Arg0\text{-}agt}$ ***vyraženia***]$_{NP}$
CRIMINAL **EXPRESSIONS**

'crime slang **expressions**'

(ii) [*las* ***expresiones*** [*de* *los* *criminales*]$_{PP\text{-}NC\text{-}Arg0\text{-}agt}$]$_{NP}$
THE **EXPRESSIONS** OF THE CRIMINALS

'the **expressions** of criminals'

(iii) *las **expresiones** criminales
THE **EXPRESSIONS** CRIMINALS
'the criminal **expression**'

In Russian, the AP can be translated into PP, given the fact that it is not always possible to translate them into the corresponding AP. In example (36), the argument of Russian deverbal noun *vyraženie* 'expression', realized as the AP *blatnye* 'criminal', is translated into a PP, that is, *de los criminales* 'of criminals' in Spanish. In this example, it is not possible to translate it into the corresponding AP since the meaning would change (36.iii).

(37)
(i) [***Peremena*** [*v* *eë* *lice*]$_{\text{PP-NC-ArgM-loc}}$]$_{\text{NP}}$ *porazila* *menja*
CHANGE IN HER FACE STRUCK ME
'The **change** in her face struck me.'

(ii) [*El* ***cambio*** [*que* *se* *había* *producido* *en* *su* *cara*]$_{\text{RelC-NC}}$]$_{\text{NP}}$
THE **CHANGE** THAT -- HAVE PRODUCED IN HER FACE
me *asombró*
ME STRUCK
'The **change** that had been produced in her face struck me.'

In the case of a noun complement realized as a PP, it can be translated into a RelC (37). In (37), Russian PP *v eë litse* 'in her face' is translated into a RelC *que se había producido en su cara* 'which had been produced in her face', in Spanish.

(38)
(i) *pobreždena* [***želaniem*** [*utverdit'* *sebja*]$_{\text{SubC-NC-Arg1-tem}}$] $_{\text{NP}}$
DAMAGED **WISH** TO_CONFIRM ONESELF
'damaged by the **wish** to confirm oneself'

(ii) *trastornada* *por* [*el* ***deseo*** [*de* *afirmarse*]$_{\text{PP-NC-Arg1-tem}}$]$_{\text{NP}}$
DISTURBED BY THE **WISH** OF TO_CONFIRM_ONESELF
'disturbed by the **wish** to confirm oneself'

(39)

(i) *Ona* *plakala* *ot* [*skorbnogo* ***soznania,*** [*čto* *ikh* *žizn'* *tak*
SHE CRIED FROM PAINFUL **BELIEF** THAT THEIR LIFE SO
pečal'no *složilas'*]$_{\text{SubC-NC-Arg1-tem}}$]$_{\text{NP}}$
DISGRACEFUL WAS

'She cried because of the painful **belief** that their life was so disgraceful.'

(ii) Lloraba *por* [*la* *dolorosa* ***evidencia*** [*de* *que* *sus* *vidas*
CRIED FOR THE PAINFUL **EVIDENCE** OF THAT THEIR LIVES
tenían *un* *destino* *tan* *aciago*]$_{\text{PP-NC-Arg1-tem}}$] $_{\text{NP}}$
HAVE A FATE SO DISGRACEFUL

'She cried because of the painful **evidence** that their lives had a fate so disgraceful.'

Finally, noun complements realized as infinitives (38) or subordinate clauses (39) are translated into PPs introduced by the non-marked preposition *de* 'of'. In examples (38) and (39), the argument of the deverbal nouns *želanie* 'wish' and *soznanie* 'belief', realized as the infinitive *utverdit'* 'to confirm' and as the relative clause *čto ikh žizn' tak pečal'no složilas'* 'that their life was so disgraceful', are translated into a PP that includes the corresponding infinitive or RelC, that is, *de afirmarse* 'of to confirm oneself' and *de que sus vidas tenían un destino tan aciago* 'that their lives had a destiny so disgraceful', in Spanish.

When translating the NP in Russian into a PP in Spanish the most usual preposition is the unmarked *de* 'of', however, other prepositions can be used as in example (40): the PP in Spanish is introduced by the preposition *contra* 'against'. This preposition is not the only possibility, since the NP in genitive could also be translated into a PP, introduced by the preposition *de* 'of' as in example (40.iii).

(40)

(i) [***lečenie*** [*impotencii*]$_{\text{NP-GEN-NC-Arg1-pat}}$]$_{\text{NP}}$
TREATMENT IMPOTENCY

'the **treatment** of the impotency'

(ii)[*el* ***tratamiento*** [*contra* *la* *impotencia*]$_{\text{PP-NC-Arg1-pat}}$]$_{\text{NP}}$
THE **TREATMENT** AGAINST THE IMPOTENCY

'the **treatment** against the impotency'

(iii) [*el* ***tratamiento*** [*de* *la* *impotencia*]$_{\text{PP-NC-Arg1-pat}}$]$_{\text{NP}}$
THE **TREATMENT** OF THE IMPOTENCY
'the **treatment** of the impotency'

See appendix B to consult the different correspondences between Russian and Spanish argumental constituents.

b. Syntactic function structural change: In this type of structural change the number of arguments explicitly realized in the NP headed by the deverbal noun is the same, but they have different syntactic functions. It is worthy to note that syntactic function change does not necessarily involves constituency changes. The syntactic function change is necessarily triggered by a PoS change.

(41)
(i) *Okhota,* [***dobyča*** [*Savčenko*]$_{\text{NP-NC-Arg0-agt}}$]$_{\text{NP}}$
HUNTING **PREY** SAVCHENKO
'The prey that **was caught** by Savchenko.'

(ii) *Presa* [*que* ***había*** ***cazado*** [*Savchenko*]$_{\text{NP-Subj-Arg0-agt}}$]$_{\text{SubC}}$
PREY THAT **HAD** **HUNTED** SAVCHENKO
'The prey that Savchenko **had hunted**.'

(42)
(i) *sposobnost'* [***vydači*** [*znanij*]$_{\text{NP-GEN-NC-Arg1-pat}}$]$_{\text{NP}}$
SKILL **DELIVERY** KNOWLEDGE
'the skill of knowledge **transmission'**

(ii) la facultad de [***exponer*** [*algún* *conocimiento*]$_{\text{NP-DO-Arg1-pat}}$] $_{\text{InfC}}$
THE SKILL OF **TO_PRESENT** SOME KNOWLEDGE
'the skill **to present** some knowledge'

In (41), the Arg0 realized as a NP, that is, *Savčenko* 'Savchenko', has different syntactic functions in each language: in Russian it is a noun complement and in Spanish it is the subject of the Spanish verb *había cazado* 'had hunted'. In (42), Arg1 is realized, in both languages, as a NP, however, they have different syntactic functions. Russian NP *znanij* 'knowledge' is in genitive and it acts as a noun complement,

whereas the Spanish NP *algún conocimiento* 'some knowledge' acts as the direct object of the infinitive.

(43)

(i) *bez*	*prava*	[***vyezda***	[*s*	*Kolymy*] $_{\text{PP-NC-Arg3-src}}$]$_{\text{NP}}$
WITHOUT	RIGHT	**DEPARTURE**	FROM	KOLYMA

'without the right to leave Kolyma'

(ii) *sin*	*derecho*	[*a*	***abandonar***	[*Kolimá*]$_{\text{NP-DO-Arg3-src}}$]$_{\text{InfC}}$
WITHOUT	RIGHT	TO	**ABANDON**	KOLYMA

'without the right to **abandon** Kolyma'

In example (43), the argument of the Russian deverbal noun *vyezd* 'departure' is realized as a different constituent in a different syntactic function with respect to the original in Spanish. In Russian the argument is realized as a PP acting as a noun complement, on the other hand, in Spanish, the argument is realized as a NP acting as a direct object.

c. Order structural change: In this structural change, there is a change in the distribution of the possessive determiner and noun complements. Adjectives in Russian tend to appear before the noun (44.i), whereas in Spanish, most of the times, they appear after the noun (44.ii).

(44)

(i) *načalis'*	[[*strašnye*]$_{\text{AP}}$	***igry***	[*na*	*živtsa*]$_{\text{PP}}$]$_{\text{NP}}$
BEGAN	TERRIBLE	**GAMES**	FOR	LIVE_BAIT

'terrible **games** for the life bait began'

(ii) *Se*	*inició*	[*el*	***juego***	[*pavoroso*]$_{\text{AP}}$	[*del*	*cebo*]$_{\text{PP}}$]$_{\text{NP}}$
--	BEGAN	THE	**GAME**	TERRIFYING	OF_THE	LIVE_BAIT

'The terryfying **game** for the live bait began'

(45)

(i) [***Razmyšlenia***	[*moi*]$_{\text{Poss}}$]$_{\text{NP}}$	*prervany*
REFLECTIONS	MY	INTERRUPTED

'My **reflections** were interrupted'

(ii) [[*Mis*]$_{\text{Poss}}$ ***reflexiones***]$_{\text{NP}}$ *se* *vieron* *interrumpidas*
MY **REFLECTIONS** -- SEEN INTERRUPTED
'My **reflections** were interrupted'

In example (45), the Russian possessive pronoun appears after the deverbal noun *razmyšlenija* 'reflections', while in Spanish the possessive determiner appears before the deverbal noun.

So far, we observe the following facts:

1. In our sample, 31.26% of cases present a structural change. Among them, 16.77% of cases present constituent change, whereas 7.69% of cases present syntactic function change and 6.80% of cases present an order structural change.
2. PoS change triggers a syntactic functional change necessarily 34.07%, whereas only 9.49% of cases with a PoS change present constituent change. The relation between the categorial change and the order change is not relevant, since, there always will be an order change when there is a change in the category of the deverbal noun.
3. Constituent function and syntactic function changes take place together in 8.37% of cases.

(4) Argument structure change: It involves a change in the number of arguments syntactically realized inside the NP. This change involves either an incorporation of one argument to the root of the deverbal noun, or the explicitation by adding an argument in the NP. Regarding thematic roles, there are no changes.

a. Incorporation change (Dorr, 2004: *Inflational-conflational divergence*): As we do not consider the direction of the incorporation, we do not distinguish between conflation and inflation, in both cases we assume that in one of the two languages an incorporation has taken place. Therefore, both (46) and (47) are treated as incorporations in Spanish and in Russian, respectively.

(46)
(i)[*rybnaja*]$_{\text{AP-NC-Arg1-pat}}$ ***lovlja***
FISH **CATCHING**
'**fishing**'

(ii) ***pesca***
FISHING
'**fishing**'

(47)
(i)***zazemlenie***
EARTHING
'**earthing'**

(ii) *toma* [*de* **tierra**]$_{\text{PP-NC-Arg1-pat}}$
TAKING OF **EARTH**
'**earthing'**

The commonest is to have an argument (core argument) incorporated in the deverbal noun, as in (46) and (47) where the Arg1 is incorporated. However, in (48) the element incorporated is an adjunct (non-core argument).

(48)
(i) [***zimovka***]$_{\text{NP}}$
WINTERHOME
'**winter home'**

(ii) [***residencia*** [*de* *invierno*]$_{\text{PP-NC-ArgM-temp}}$]$_{\text{NP}}$
RESIDENCE OF WINTER
'winter **home'**

A special case of incorporation concerns to light verb constructions. Both Russian and Spanish can express the same action by means of a construction formed by a light verb and a deverbal noun (49.i, 50.i) and by a simple verb form (49.ii, 50.ii). In the former case, the base verb of the deverbal noun is the same as the simple verb form.

(49)
(i) *vnesti* ***vklad***
TO_DO **CONTRIBUTION**
'to do a **contribution**'

(ii) *vkladyvat'*
TO_CONTRIBUTE
'to contribute'

(50)
(i) *hacer* *una* ***contribución***
TO_DO A **CONTRIBUTION**
'to do a **contribution'**

(ii) *contribuir*
TO_CONTRIBUTE
'to contribute'

Therefore, a light verb construction can be translated either into the corresponding light verb construction (a more literal translation) or into the corresponding simple verb form.

(51)
(i) *tot* *umer,* *ne* [*prokhodja* *v* [***soznanie***]$_{NP}$]$_{VP}$
THAT_ONE DIED, NOT PASSING IN **CONSCIOUSNESS**
'The man died without regaining **consciousness**.'

(ii) *El* *hombre* *murió* *sin* [*recobrar* [*el*
THE MAN DIED WITHOUT TO_RECOVER THE
conocimiento]$_{NP}$]$_{VP}$
CONSCIOUSNESS
'The man died without regaining **consciousness**.'

(52)
(i) [[*Lemkusa*]$_{NP}$ ***priveli*** ***v*** [***soznanie***]$_{NP}$]$_{VP}$
LEMKUS TOOK TO **CONSCIOUSNESS**
'They **revived** Lemkus.'

(ii) [***Reanimaron*** [*a* *Lemkus*]$_{NP}$]$_{VP}$
REVIVED TO LEMKUS
'They **revived** Lemkus.'

Examples (51) and (52) show two light verb constructions with the Russian deverbal noun *soznanie* 'consciousness' which are translated into a light verb construction (51) and into a single verb in (52).

In incorporation cases where the deverbal noun is translated into a light verb construction, the argument incorporated can be core (53) or non-core (54).

(53)

(i) *prosja*	[[*roditel'skogo*]$_{\text{AP-NC-Arg0-agt}}$	***blagoslovenja***]$_{\text{NP}}$
ASKING	PARENTAL	**BLESSING**

'asking for parental **blessing**'

(ii) *rogándole*	*que*	[[*me*]$_{\text{NP-IO-Arg2-dest}}$	***diera***	[*su*]$_{\text{Poss-SPEC-Arg0-agt}}$	***bendición***]$_{\text{VP}}$	
BEGGING	THAT	TO_ME	**GIVE**	THEIR	**BLESSING**	

'begging that he **gave** me their **blessing'**

(54)

(i) *Ja*	*prognal*	*nabegajušij*	*son*	[***naprjaženiem***	[*vsego*	
I	BANISHED	INVADING	SLEEP	**TENSION**	WHOLE	

tela]$_{\text{NP-NC-Arg1-tem}}$]$_{\text{NP}}$
BODY

'I shook off the sleep that was overcoming me by **tensing** my whole body.'

(ii) *Me*	*deshice*	*del*	*sueño*	*que*	*me*	*invadía*
ME	BANISHED	OF_THE	SLEEP	THAT	ME	INVADED

[***poniendo***	***en***	***tensión***	[*todo*	*mi*	*cuerpo*]$_{\text{NP-NC-Arg1-tem}}$]$_{\text{NP}}$
PUTTING	**IN**	**TENSION**	WHOLE	MY	BODY

'I shook off the sleep that was overcoming me by **tensing** my whole body.'

b. Explicitation change: It involves the syntactic realization of an argument inside the NP, which is implicit from the context. In an explicitation, an argument, which is not expressed neither syntactically in the NP nor incorporated into the root of the deverbal noun, can be recovered from the context and added to the structure of the deverbal noun. In example (55), the Russian deverbal noun is translated into a verb in Spanish, which expresses the Arg0-agent in its verbal ending and realizes syntactically the Arg1-patient by means of a personal pronoun. In Russian, the Arg0 and the Arg1 are not realized syntactically inside the NP headed by the deverbal

noun *vstreča* 'meeting', since they are already expressed syntactically in the previous context.

(55)

(i) *Ty* $_{\text{NP-Subj-Arg0-agt}}$ *mne* $_{\text{NP-IO-Arg2-ben}}$[45] *ne* *poklonilsja* *pri* [***vstreče***]$_{\text{NP}}$
YOU TO_ME NOT SAY_HELLO DURING **MEETING**

'You didn't say hello to me at the **meeting**.'

(ii) *No* *me* *has* *saludado* [*cuando* [[*me*]$_{\text{NP-DO-Arg1-pat}}$ ***has*** ***visto***]$_{\text{VP}}$]$_{\text{SubC}}$
NOT TO_ME HAVE SAY_HELLO WHEN ME **HAVE** **SEEN**

'You didn't say hello when you **have seen** me.'

Another type of explicitation change is related to the explicitation of the possessive determiner, expressing a core argument, an Arg1-pat as in (56). In example (56), the type of determiner (in this case a possessive determiner) has to be inferred from the context since it is not present in Russian.

(56)

(i) *Eto* *tebe* *nužno* *dlja* [***lečenija***]$_{\text{NP}}$
THIS YOU IS_NECESSARY FOR **TREATMENT**

'This is necessary for you for the **treatment**.'

(ii) *te* *conviene* *para* [[*tu*]$_{\text{Poss-SPEC-Arg1-pat}}$ ***tratamiento***]$_{\text{NP}}$
YOU NEED FOR YOUR **TREATMENT**

'You need this for your **treatment**.'

So far, we summarize the following observations related to the incorporation change.

1) The incorporation change takes place 2.77% in our sample. The commonest is to have an argument (core argument) incorporated in the deverbal noun (68.18%), but the element incorporated can also be an adjunct (non-core argument) in 31.88% of cases.

2) In 45.45% of incorporation cases are light verb constructions. Among them, 90% of cases have a core argument incorporated, whereas in 10% of cases the incorporated argument is non-core.

[45] The thematic role can be different that the one that would be expressed if it was explicitly expressed in the NP containing the deverbal noun.

3) In the sample, the explicitation change takes place in 7.31% of cases. Out of them, 81.81% of cases the argument introduced is a core argument, whereas in 10.97% of cases it is a non-core argument.

4) The explicitation of the internal argument of the deverbal noun can be an indicator used by the language to disambiguate the denotation of the deverbal noun. As we have seen in chapter 3 section 3.6.2, the expression of the internal argument is tightly related to the event reading.

(5) Lexical change (Dorr, 1994: *Lexical divergence*)**:** A lexical change is the translation of the deverbal noun into a word, which does not convey completely the meaning of the source noun, since the noun (or the verb, the adjective or, even, the past participle) chosen for the translation is either more general or more nuanced.

(57)
(i) [***vyezda*** [*s* *Kolymy*]$_{PP}$]$_{PP}$
DEPARTURE FROM KOLYMA
'the **departure** from Kolyma'

(ii) [***abandonar*** [*Kolimá*]$_{NP}$]$_{InfC}$
TO_ABANDON KOLYMA
'to **abandon** Kolyma'

In example (57), the word used does not correspond in content directly to the original, that is, instead of *salida* o *salir* 'exit or to exit', he has chosen *abandonar* 'to abandone', which is much more nuanced than the Russian one.

Lexical change takes place in the sample 8.82% of cases.

(6) Head swapping change (Dorr, 2004: *Head Swapping Divergence*): A head swapping mismatch is "the inversion of a structural dominance" as was already defined in Dorr (2004:16). See section 3.

(58)
(i)[*rebjačeskij* ***vyrez***$_{\mathbf{noun}}$ *gub*]$_{NP}$
INFANTILE **CONTOUR** LIPS
'infantile **contour** of lips'

(ii)[*la* ***perfilada***$_{adj}$ *boca* *de* *un* *niño*]$_{NP}$
THE **OUTLINED** MOUTH OF A CHILD

'the **outlined** mouth of a child'

(59)
(i) [*tverjak* *po* ***rožden****iju*$_{noun}$]$_{NP}$
TVERIAN ACCORDING **BIRTH**

'Tverian according to **birth'**

(ii) [*originario*$_{adj}$ [*de* *Tver*]$_{PP}$]$_{AP}$
NATIVE OF TVER

'**Native** from Tver'

In example (58), Russian deverbal noun *vyrez* 'contour' is translated into the adjective *perfilada* 'outlined' in a modifier position in Spanish. Another example is (59), Russian deverbal noun *rožden**ie*** 'birth', which is a complement of an AP, is translated into the adjective *originario* 'native' in a head position, in Spanish.

This linguistic change is not very representative in our sample. In fact, it only takes place in 0.63% of cases.

(7) Number change: The number change is the translation of the deverbal noun into a noun in a different number. For instance:

(60)
(i) ***Vydači***$_{noun-PL}$ *gematogena* *dlja* *bol'nykh* *ne*
ADMINISTRATIONS HEMATOGENE FOR SICK_PEOPLE NOT
byli *novost'ju*
WAS NEW

'The **administration** of hematogene for the sick people was not a novelty.'

(ii) *La* ***administración***$_{noun-SG}$ *de* *hematógeno* *a* *los* *enfermos* *no*
THE **ADMINISTRATION** OF HEMATOGENE TO THE SICK NO
era *algo* *nuevo*
WAS SOMETHING NEW

'The **administration** of hematogene to the sick people was nothing new.'

In (60), the plural Russian deverbal noun *vydači* 'administrations' is translated into the singular deverbal noun *administración* 'administration' in Spanish.

In the sample analyzed the number change takes place 3.02% of cases.

(8) Discursive changes: A discursive change takes place when in Russian the same discursive entity is named through two NPs, whereas in Spanish, the second mention of that discursive entity is not referred by means of a NP, but by means of a pronoun or an elliptical element which has as an antecedent a NP realized in the previous sentence.

a. Coreference-anaphoric change: This change takes place when the deverbal noun is translated into a pronoun, whose reference is a discursive entity appeared in the previous sentence. Therefore, the use of an anaphoric element is to avoid redundancy.

(61)

(i) *Ja*	*ne*	*vyjdu*	*za*	*tebja*	*bez*	***blagoslovenia***
I	NOT	WILL_GO	BEHIND	YOU	WITHOUT	**BLESSING**

tvojkh	*roditelej.*	[[*Bez*	[[*ikh*]$_{\text{Poss-SPEC-Arg0-agt}}$	***blagoslovenia***]$_{\text{NP}}$]$_{\text{PP}}$
YOUR	PARENTS.	[[WITHOUT	THEIR	**BLESSING**

ne	*budet*	*tebe*	*sšastia.*
NOT	WILL_BE	TO-YOU	HAPPINESS

'I won't marry you without your parent's blessing. Without their **blessing** you won't be happy'

(ii) *Sin*	*la*	***bendición***	*de*	*tus*	*padres*	*no*	*me*
WITHOUT	THE	**BLESSING**	OF	YOUR	PARENTS	NO	ME

casaré	*contigo.*	[*Sin*	[***ella***]$_{\text{NP}}$]$_{\text{PP}}$	*no*	*serás*
WILL_MARRY	WITH_YOU.	WITHOUT	**SHE**	NO	WILL_BE

feliz.
HAPPY.

'Without the **blessing** of your parents, I won't marry you. Without **it**, you won't be happy.'

In (61), Russian deverbal noun *blagoslovenie* 'blessing' can be translated into the personal pronoun *ella* SHE 'it' in Spanish because its antecedent is realized in the previous sentence.

This type of change is not very representative in our sample since it only takes place 0.37% of cases.

In those cases in which the Russian deverbal noun was translated into an anaphoric element, they were personal pronouns in 75% of cases, whereas in the rest of the cases they were indefinite pronouns.

b. Coreference-elliptic change: This change corresponds to the case when in Russian we have two NPs referring to the same discursive entity, but in Spanish we have a NP, which is in coreference with an elliptical element.

(62)

(i) [***Poverki***]$_{NP}$*? Net,*	[***poverok***]$_{NP}$	*zdes'*	*ne*	*bylo*	
CHECKING? NO,	**CHECKING**	HERE	NOT	WAS	

'Checking? No, here there wasn't any **checking**.'

(ii)*¿Alguna*	[***comprobación***]$_{NP}$*?*	*No,*	*aquí*	*no*	*se*	*hacían*
ANY	**CHECKING**?	NO,	HERE	NO	--	DID

'Any checking? No, here there weren't any.'

In (62) the second deverbal noun *poverka* 'checking' is an elliptical element in Spanish, while in Russian the discursive entity of the 'checking' is named twiced by means of two coreferred NPs.

This type of change is not very representative in our sample it only takes place 0.37% of cases.

5.4.2 Interrelation of translation changes

As we have already said above, a mismatch is rarely constituted by a single linguistic change. In fact, a translation mismatch is usually composed of different interrelated linguistic changes. In this section we present a classification of mismatches based on the number of linguistic changes and according to the type of the linguistic changes involved. We understand translation mismatches as different degrees of semantic similarity between the SL and the TL. This is seen as a continuum where, at one edge we would find the literal translation, that is, the word-for-word translation of a sentence (63), while at the opposite edge we would find constructions that convey the same meaning by means of completely different structures (64).

(63)

(i) [[*Naš*]$_{\text{Poss}}$	***znakomstva***]$_{\text{NP}}$	*ukreplialis'*	
OUR	**ACQUAINTANCES**	STRENGTHENED	

'Our **relations** strengthened.'

(ii) [[*Nuestras*]$_{\text{Poss}}$	***relaciones***]$_{\text{NP}}$	*se*	*fortalecían*
OUR	**RELATIONS**	--	STRENGTHEN

'Our **relations** strengthened.'

(64)

(i) *vidja*	[[*blagoprijatnuju*]$_{\text{AP}}$	***peremenu***	[*pogody*]$_{\text{NP}}$]$_{\text{NP}}$
SEEING	GENTLE	**CHANGE**	WEATHER

'seeing the gentle **change** of weather'

(ii) *viendo*	*que*	[*escampaban*	[*los*	*nubarrones*]$_{\text{NP}}$]$_{\text{VP}}$
SEEING	THAT	WENT_AWAY	THE	CLOUDS

'seeing that the clouds were going away'

In (63), the source and the target languages are literal correspondences, that is, the target sentence is a word-for-word translation of the source sentence. On the other hand, in (64), the source and the target sentences are completely different, even if they preserve some part of the source meaning, the whole sentence changes dramatically.

The focus of this study is not situated upon the two extreme edges, but on those cases, where a mismatch is composed of one, two or three systematic linguistic changes. Cases with more than three co-occurring mismatches are rare but possible. However, given the fact that the number of cases with four or more linguistic changes was so low and the possible combinations are so large in our sample, it was not possible to systematize them, and for this reason they have been considered free translations.

Those rare cases in which we have found four or more linguistic changes were considered as free translations. Some of these cases result in completely free translations, but, in other occasions, they do not seem so distant (65).

(65)

(i) *bez*	[***objasnenia***$_{\text{noun}}$	[*pričin*	*otkaza*]$_{\text{NP-NC}}$]$_{\text{NP}}$
WITHOUT	**EXPLANATION**	CAUSES	REFUSAL

'without any **explanation** about the causes of the refusal'

(ii) *sin* [[***dar*** [***motivo***]$_{noun}$]$_{verb}$ *alguno* [*a* *los* *rechazos*]$_{PP\text{-}CREG}$]$_{VP}$
WITHOUT **TO_GIVE** **REASON** ANY TO THE REFUSALS
'without **giving any reason** to the refusals'

Example (65) combines four linguistic changes, that is, PoS, constituent structure, syntactic function structure and the incorporation argument structure changes. The Russian deverbal noun *objasnenie* 'explanation' is translated into the light verb construction *dar motivo* 'to give an explanation'. The Russian deverbal noun has incorporated the Arg1 into its root, while in Spanish a construction that realizes its complements is used analytically. In Spanish the verb requires different functions for the complements, the complement *pričin otkaza* 'causes of the refusals' is no longer a noun complement but a prepositional object (CREG). Moreover, the type of constituent is no longer a NP but a PP *a los rechazos* 'to the refusals'.

In our sample, 16% of cases are literal translations, that is, there are no linguistic changes. On the other edge of the translation continuum, we would have 3.2% of translations with more than three translation mismatches.

In the following sections we introduce the different mismatches found in our sample. They have been classified in three groups according to the number and type of linguistic changes. Tables (in sections a, b and c) are organized in the following way: the first column contains the linguistic change, the second column contains an example of the mismatch and the third column presents the frequency of each type of mismatch in each group (a, b, c).

a. Mismatches composed of one linguistic change

In table 1, we present the 8 different types of translation mismatches composed of one linguistic change that have been found in MiniRuSp, which represent 25.8% of examples in the sample.

	Example	%
Determiner	(66) (i) *Vyzyvala* [***vozmuščenie*** *daže* [*u* *tekh*]$_{PP}$]$_{NP}$ (…) PROVOKED **INDIGNATION** EVEN OF THOSE 'It provoked the **indignation** even of those (…).' (ii) *Suscitaba* [*la* ***indignación*** *incluso* [*de* *aquellos*]$_{PP}$]$_{NP}$ (…) PROVOKED THE **INDIGNATION** EVEN OF THOSE 'It provoked the **indignation** even of those (…).'	50.38

PoS	(67) (i) *Sekundy* *tratila* *ona* *na* [***otdykh***]$_{NP}$ SECONDS SPENT SHE IN **REST** 'She spent some seconds to **rest**.' (ii) *Consumió* *su* *tiempo* *en* [***descansar***]$_{VP}$ SPENT HER TIME IN **TO_REST** 'She spent her time **resting**.'	9.30
Constituent structure	(68) (i) *po* [***vyraženiju*** [*Dostoievskogo*]$_{NP}$]$_{NP}$ ACCORDING **EXPRESSION** DOSTOIEVSKY 'according to the **expression** of Dostoyevsky' (ii) *en* [***expresión*** [*de* *Dostoyevski*]$_{PP}$]$_{NP}$ IN **EXPRESSION** OF DOSTOIEVSKY 'according to the **expression** of Dostoyevsky'	10.07
Order structure	(69) (i) [***Razmyšlenija*** [*moi*]$_{Poss}$]$_{NP}$ *prerveny* **REFLECTIONS** MY INTERRUPTED 'My **reflections** were interrupted.' (ii) [[*Mis*]$_{Poss}$ ***reflexiones***]$_{NP}$ *se* *vieron* *interrumpidas* MY **REFLECTIONS** -- SEEN INTERRUPTED 'My **reflections** were interrupted.' (70) (i) [[*živyje*]$_{AP}$ ***illjustratsii***]$_{NP}$ LIVELY **ILLUSTRATIONS** 'lively **illustrations'** (ii) [***ilustraciones*** [*vivas*]$_{AP}$]$_{NP}$ **ILLUSTRATIONS** LIVELY 'lively **illustrations'**	7.75

Incorporation	(71) (i) *vidja* [[*vaše*]$_{\text{Poss}}$ [*nervnoe*]$_{\text{AP-NC-Arg2-atr}}$ ***sostojanie***]$_{\text{NP}}$ SEEING YOUR NERVOUS **STATE** 'seeing your nervous **state'** (ii) *viendo* [[*su*]$_{\text{Poss}}$ ***nerviosismo***]$_{\text{NP}}$ SEEING HIS **NERVOUSNESS** 'seeing his **nervousness'**	2.32
Explicitation	(72) (i) *Ja ne mog najti* [*nikakogo* ***rešenia***]$_{\text{NP}}$ I NOT COULD TO_FIND NO **SOLUTION** 'I couldn't find a **solution**.' (ii) *Y no podía encontrar* [*ninguna* ***solución*** [*al* AND NOT COULD FIND NO **SOLUTION** TO *problema*]$_{\text{PP-NC-Arg1-tem}}$]$_{\text{NP}}$ PROBLEM 'And I couldn't find any **solution** to the problem.'	6.20
Lexical	(73) (i) *Na* [[*ego*]$_{\text{Poss}}$ ***zov***]$_{\text{NP}}$ *otklikalas'* TO HIS **CALL** RESPONDED 'She responded to his **call**.' (ii) *Reaccionaba a* [[*su*]$_{\text{Poss}}$ ***voz***]$_{\text{NP}}$ REACTED TO HIS **VOICE** 'She reacted to his **voice**.'	7.75
Number	(74) (i) *Oni na* [*etoj* ***službe***]$_{\text{NP-SG}}$ *menjajut svoi familii* THEY AT THIS **SERVICE** CHANGE THEIR SURNAMES 'They in this **position** change their surnames.' (ii) *En* [*estos* ***cargos***]$_{\text{NP-PL}}$, *esta gente cambia de* IN THESE **POSITIONS** THESE PEOPLE CHANGE OF *apellido* SURNAME 'In these **positions**, these people change their surnames.'	1.55

Coreference-anaphoric	(75) (i) *Pobežhal k Šejninu (…) otprazdnovat' [**posylku**]$_{NP}$* RAN TO SHEININ (…) TO_CELEBRATE **PACKAGE** 'I ran to Sheinin to celebrate the **package**.' (ii) *Corrí por Sheinin para celebrar[**lo**]$_{NP}$* RAN FOR SHEININ FOR TO_CELEBRATE_[**IT**]$_{NP}$ 'I ran to Sheinin in order to celebrate **it.**'	2.32
Coreference-elliptic	(76) (i) *On postupil [na [**službu** [v Dal'stroj]$_{PP}$]$_{NP}$]$_{PP}$* HE ENTERED INTO **SERVICE** IN DALSTROI 'He entered into the **service** in Dalstroi' (ii) *Ingresó [en [el Dalstroi]$_{NP}$]$_{PP}$* ENTERED INTO THE DALSTROI 'He began the service at Dalstroi'	2.32

Table 1: One linguistic change mismatches

The commonest type of mismatch composed of one linguistic change is clearly the determiner change (50.38%). This is not strange since in the overall corpus this is the linguistic change with the highest frequency of appearance (54.2%) due to the typological differences between the two languages. The least frequent is the mismatch composed of a number change (1.55%).

The constituent structural mismatch (10.07%) follows in frequency the determiner mismatch, however the frequency is quite low when compared with the determiner mismatch (50.38%). This linguistic change is also explained by typological reasons. In Spanish it is not possible to have two NPs, one acting as the head and the other as the complement. Therefore, Russian NPs, acting as noun complements, are translated into PPs.

The PoS mismatch is found 9.30% of cases. As we will see in sections (b) and (c), a PoS mismatch usually can involve other linguistic changes such as a syntactic function structural change or the addition of a determiner specifying an infinitive.

The order structural mismatch and the lexical mismatch are equally frequent in our sample (7.75%). They are followed in frequency by the explicitation mismatch (6.20%). Finally, the incorporation mismatch, the coreference-anaphoric mismatch

and the coreference-elliptic mismatch have a low frequency (2.32%). These three are quite anecdotic, but they show possible mismatches.

b. Mismatches composed of two linguistic changes

Mismatches included in this group involve two linguistic changes. This mismatch has an overall frequency of 41.2%. As we will see mismatches composed of two linguistic changes are the commonest. We have classified them into three subgroups, according to the linguistic change that they have in common: the determiner group, the PoS group and the order structure group.

(i) The determiner group: This first group includes those mismatches that have in common the determiner change. The determiner change, as we have already seen, is a typological change, which can appear with a wide range of other linguistic changes.

		Example	%
Determiner	PoS	(77) (i) *V* [***dojke***]$_{NP}$ *on* *často* *ošibalsja* IN **MILKING** HE OFTEN MISTOOK 'He often got confused at the **milking**.' (ii) A[l ***ordeñar***]$_{VP}$ *a_menudo* *se* *equivocaba* IN_THE **TO_MILK** OFTEN - MISTOOK 'He often got confused at the **milking**.'	0.48

	Constituent structure	(78) (i) *nabljudat'* *za* [***vypolneniem*** [*plana*]$_{NP\text{-}gen}$]$_{NP}$ TO_OBSERVE FOR **ACCOMPLISHMENT** PLAN 'to observe the **accomplishment** of the plan' (ii) *vigilar* [*el* ***cumplimiento*** [*del* *plan*]$_{PP}$]$_{NP}$ TO_CONTROL THE **ACCOMPLISHMENT** OF_THE PLAN 'to control the **accomplishment** of the plan' (79) (i) [***Zavedovanie*** [*otdeleniem*] $_{NP\text{-}INS}$] $_{NP}$ *peršlo* *k* *doktoru* **DIRECTORSHIP** SECTION PASSED TO DOCTOR *Zaderu* ZADER 'The **directorship** of the section was passed to doctor Zader.' *(ii)* [*La* ***jefatura*** [*de* *la* *sección*]$_{PP}$]$_{NP}$ *pasó* THE **DIRECTORSHIP** OF THE SECTION PASSED *al* *doctor* *Zader* TO_THE DOCTOR ZADER 'The **directorship** of the section was passed to doctor Zader.' (80) (i) *vygovarivaja* *v* *nadležaščej* *intonacii* [[*blatnye*]$_{AP}$ PRONOUNCING IN CORRECT INTONATION CRIMINAL ***vyraženija***]$_{NP}$ **EXPRESSIONS** 'pronouncing in the correct intonation criminal **expressions**' (ii) *pronunciando* *con* *la* *entonación* *adecuada* [*las* PRONOUNCING WITH THE INTONATION CORRECT THE ***expresiones*** [*de* *los* *criminales*]$_{PP}$]$_{NP}$ **EXPRESSIONS** OF THE CRIMINALS 'pronouncing in the correct intonation the **expressions** of criminals' (81) (i) [[*strastnoe*]$_{AP}$ ***želanie*** [*izbavit'sja* *klejma*]$_{SubC}$]$_{NP}$ PASSIONATE **DESIRE** TO_GET_RID STIGMA 'the passionate **desire** to get rid of the stigma' (ii) [*aquel* [*ardiente*]$_{AP}$ ***deseo*** [*de* *librarse* THAT PASSIONATE **DESIRE** OF TO_GET_FREE *del* *estigma*]$_{PP}$]$_{NP}$ OF_THE STIGMA 'that passionate **desire** to get rid of the stigma'	34.46

	Order structure	(82) (i) *Eto* [[*organizovannoe*]$_{AP}$ ***vystuplenie***]$_{NP}$ THIS ORGANIZED **ACTION** 'This is an organized **action**.' (ii) *Esto es* [*una* ***protesta*** [*organizada*]$_{AP}$]$_{NP}$ THIS IS A **PROTEST** ORGANIZED 'This is an organized **protest**.'	11.65
	Incorporation	(83) (i) *Perepisyval načisto* [[*trebovatel'nye*]$_{AP\text{-}NC\text{-}Arg1\text{-}tem}$ ***zapisi***]$_{NP}$ REWRITE CLEAN COMPLAINT **NOTES** 'He writes a clean version of the complaint **notes**.' (ii) *Pasaba a limpio* [*las* ***reclamaciones***]$_{NP}$ PASSED TO CLEAN THE **COMPLAINTS** 'He writes a clean version of the complaint **notes**.'	2.91
	Explicitation	(84) (i) *Vaša doč' prosit* [***rešenia*** [*na* YOUR DAUGHTER ASKS **PERMISSION** TO *brak*]$_{PP}$]$_{NP}$ MARRIAGE 'Your daughter is asking **permission** to get married.' (ii) *Su hija pide* [[*su*]$_{Poss\text{-}SPEC\text{-}Arg1\text{-}agt}$ ***consentimiento*** YOUR DAUGHTER ASKS YOUR **PERMISSION** [*para casarse*]$_{PP}$]$_{NP}$ FOR TO_GET_MARRIED 'Your daughter is asking your **permission** to get married.'	13
	Lexical	(85) (i) *Ona našla* [***opory***]$_{NP}$ SHE FOUND **SUPPORT** 'She found **support**.' (ii) *Halló* [*un* ***punto de apoyo***]$_{NP}$ FOUND A **POINT OF SUPPORT** 'She found a **point of support**.'	10.67

	Head swapping	(86) (i) [[*Procentnye*]$_{AP}$ ***nadbavki***]$_{NP}$ *nado* *bylo* *vyslužívat'* PERCENTAGE **COMPLEMENT** NEED WAS TO_EARN *zanovo* AGAIN 'The percentage **complements** needed to be earned again.' (ii) *El* *tanto_por_ciento* *de* [*los* ***complementos***]$_{NP}$ *tuvo* THE PERCENTAGE OF THE **COMPLEMENTS** HAD *que* *ganárselos* *de_nuevo* THAT EARN AGAIN 'He had to earn again the percentage of the **complements**.'	0.48
	Number	(87) (i) *Byli* *dopuščeny* *k* [***ispytanijam***]$_{NP-PL}$ WERE ACCEPTED IN **EXPERIMENTS** 'They were accepted in the **experiments**.' (ii) *Se* *permitió* *participar* *en* [*el* -- PERMITTED PARTICIPATE IN THE ***experimento***]$_{NP-SG}$ **EXPERIMENT** 'To participate in the experiments was permitted.'	2.91

Table 2: Two linguistic changes mismatch

In the determiner group, the commonest mismatch is the one composed of the determiner change and the constituent structure change (34.46%). This is followed by the mismatch that combines the determiner and the explicitation changes (13%). The determiner change combined with either an order or a lexical change is quite common (11.65% and 10.67%, respectively).

Mismatches with a very low frequency, such as the combination of a determiner change and a PoS change (0.48%), are presented not as productive and representative mismatches, but as possible combinations of linguistic changes.

However, the determiner change does not appear with a syntactic function structural change since we need a PoS change in order to have a syntactic function. It does not appear with the coreference anaphoric and coreference ellyptical changes, since in the first case, the deverbal noun has been translated into a pronoun and in the second case it has not been translated.

(ii) The PoS group: This second group of mismatches includes those that have in common the PoS change.

		Example	%
PoS	Syntactic function	(88) (i) *Pri* [***izobretenii*** [*elektronnogo* *mikroskopa*]$_{NP\_NC}$]$_{NP}$ (…) AT **INVENTION** ELECTRONIC MICROSCOPE 'With the **invention** of the electronic microscope (…).' (ii) *Cuando* [*se* ***inventó*** [*el* *microscopio* *electrónico*]$_{NP\_DO}$]$_{VP}$ WHEN -- **INVENTED** THE MICROSCOPE ELECTRONIC 'When the electronic microscope **was invented** (…).'	16.50
	Incorporation	(89) (i) *Na* [***objasnenia***]$_{NP}$ *četyrnadcatyj* *ne* *rešilsja* TO **EXPLANATIONS** FOURTEENTH NOT DECIDED 'The fourteenth didn't decide to ask for **explanations**.' (ii) *El* *decimocuarto* *no* *se* *decidió* *a* [***pedir*** ***explicaciones***]$_{VP}$ THE FOURTEENTH NOT -- DECIDED TO **ASK** **EXPLANATIONS** 'The fourteenth didn't decide **to ask for explanations.**' (90) (i) *Lemkusa* [*priveli* *v* [***soznanie***]$_{NP}$]$_{VP}$ LEMKUS TOOK INTO **CONSCIOUSNESS** 'They **revived** Lemkus.' (ii) [**Reanimaron**]$_{VP}$ *a* *Lemkusa* **REVIVED** TO LEMKUS 'They **revived** Lemkus.'	1.45

	Explicitation	(91) (i) *Skazat'* *vslukh,* *čto* *rabota* *tiažela,-* *dostatočno* *dlja* TO_SAY ALOUD THAT WORK TOUGH ENOUGH TO [***rasstrela***]NP **SHOOTING** 'Saying aloud that work is tough is enough **to get shooted**.' (ii) *Decir* *en* *voz* *alta* *que* *el* *trabajo* *es* *duro* *es* TO_SAY IN VOICE HIGH THAT THE WORK IS HARD IS *suficiente* *para* *que* [[*te*]NP-DO-Arg1-tem ***fusilen***]VP ENOUGH TO THAT YOU **SHOOT** 'To say aloud that work was tough is enough **to get shooted**.'	2.91

Table 3: Two linguistic changes mismatch

The commonest mismatch of this group is the one that combines the PoS change and the syntactic function change (16.50%). The correlation of these two linguistic changes is previsible since, as we have already observed, it is not possible to have a functional syntactic change without a prior PoS change. Moreover, in some occasions it can require the explicitation of an argument not overtly expressed in the source sentence (2.91%). The incorporation of an argument combined with a categorial change is not very usual (1.45%).

(iii) The order group: This third group includes those mismatches that have in common the order change. The order change is a change that has to do with nominal complements and not with the noun itself. This linguistic change can be accompanied by the following linguistic changes.

		Example	%
Order	Constituent structure	(92) (i) [[*Dorožnye*]$_{AP}$ ***razmyšlenia*** [*moi*]$_{Poss}$]$_{NP}$ *byli* *ne* *ochen'* TRAVELLING **REFLECTION** MY WERE NOT VERY *prijatny* PLEASANT 'My travelling **reflections** were not very pleasant.' (ii)[[*Mis*]$_{Poss}$ ***reflexiones*** [*durante* *el* *viaje*]$_{PP}$]$_{NP}$ *no* MY **REFLECTIONS** DURING THE TRIP NOT *fueron* *muy* *agradables* WERE VERY PLEASANT 'My **reflections** during the trip were not very pleasant.'	1.45
	Incorporation	(93) (i) [***Zimovka*** [*naša*]$_{Poss}$]$_{NP}$ *byla* *gotova* **WINTER_RESIDENCE** OUR WAS READY 'Our **winter residence** was ready.' (ii)[[*Nuestra*]$_{Poss}$ ***residencia*** ***de*** ***invierno***]$_{NP}$ *estuvo* *lista* OUR **RESIDENCE** **OF** **WINTER** WAS READY 'Our **winter residence** was ready.'	0.48
	Explicitation	(94) (i) *Sygralo* *rol'* [*kakoe-to* [*ličnoe*]$_{AP}$ [*slučajnoe*]$_{AP}$ PLAYED ROLE CERTAIN PERSONAL CASUAL ***znakomstvo*** [*Romanova*]$_{NP\text{-}NC\text{-}Arg0\text{-}agt}$]$_{NP}$ **ACQUAINTANCE** ROMANOV 'Some casual and personal **acquaintance** of Romanov played a special role.' (ii)*Había* *jugado* *algún* *papel* [*cierta* ***relación*** HAD PLAYED SOME ROLE CERTAIN **RELATION** [*personal*]$_{AP}$ [*fortuita*]$_{AP}$]$_{NP}$ PERSONAL CASUAL 'A personal and casual **relation** had played a certain role.'	0.48

	Number	(95) (i) [***Znakomstvo***]$_{NP\text{-}SG}$ *u* *menja* *bylo* *bol'šoje* **ACQUAINTANCE** CLOSE_TO ME WAS BIG 'I had a lot of **acquaintances**.' (ii) *Tenía* *yo* [*muchos* ***conocidos***]$_{NP\text{-}PL}$ HAD I A_LOT_OF **ACQUAINTANCES** 'I had a lot of **acquaintances**.'	0.48

Table 4: Two linguistic changes mismatch

In our corpus, these mismatches are not very common since their frequency ranges from 1.45% (which corresponds to the combination of the order structure change and the constituent change) to 0.48% (which corresponds to the combinations of an order change with an incorporation, an explicitation or a number change).

c. Mismatches composed of three linguistic changes

Mismatches included in this class involve three linguistic changes. The overall frequency of these mismatches is quite low, that is, 13.6%. We have organized this mismatch in three groups: the determiner group, the PoS group and the lexical group.

(i) The determiner group: This group is subdivided in six subgroups depending on the two linguistic changes that are combined with the determiner change.

			Example	%
Determiner	PoS	Syntactic function	(96) (i) [***vozvraščenie*** [*k* *tomu* *vremeni*]$_{PP\text{-}NC}$]$_{NP}$ **RETURN** TO THAT TIME 'a **return** to that time (…)' (ii)[*un* ***retornar*** [*al* *tiempo*]$_{PP\_CREG}$]$_{NP}$ A **TO_RETURN** TO TIME 'to **return** to that time (…)'	2.89

	Constituent structure	Order structure	(97) (i)[[*zgučee*]$_{AP}$ ***želanie*** [*očutit'sja* *k* POIGNANT **DESIRE** FIND_ONSELF IN *lesu*]$_{SubC}$]$_{NP}$ FOREST 'the poignant **desire** to find oneself in the forest' (ii) [*el* ***deseo*** [*ardiente*]$_{AP}$ [*de* *encontrarte* THE **DESIRE** POIGNANT OF FINDING_YOU *en* *el* *bosque*]$_{PP}$]$_{NP}$ IN THE FOREST 'the poignant **desire** of finding oneself in the forest'	7.24
		Incorporation	(98) (i) [***peremena*** [*žizni*]$_{NP}$ [*k* *luščemu*]$_{PP}$]$_{NP}$ **CHANGE** LIFE INTO BETTER 'an **improvement** in life' (ii) [*toda* ***mejora*** [*en* *su* *vida*]$_{PP}$]$_{NP}$ EVERY **IMPROVEMENT** IN HIS LIFE 'any **improvement** in his life'	2.89
		Explicitation	(99) (i) *po* [***želaniju*** [*luščim* *obrazom* *poslužit'* ACCORDING **DESIRE** BETTER WAY SERVE *rodnomu* *kraju*]$_{SubC}$]$_{NP}$ NATIVE LAND 'according to the **desire** of a better way to serve your motherland' (ii) *por* [[*su*]$_{Poss}$ ***deseo*** [*de* *servir* ACCORDING HIS **DESIRE** OF TO_SERVE *del* *mejor* *modo* *a* *su* *tierra* OF_THE BETTER WAY TO HIS LAND *natal*]$_{PP}$]$_{NP}$ NATIVE 'according to his **desire** to serve in a better way his own motherland'	1.44

		Lexical	(100) (i) *Javlialos'* [***soznaniem*** [*neudobletvorennoj,* EXISTED **REALIZATION** DISSATISFACTORY *isporčennoj* *žizni*]$_{NP}$]$_{NP}$ SPOILED LIFE 'There was the **realization** of a dissatisfactory and spoiled life.' (ii) *Se* *debía* *a* [*la* ***evidencia*** [*de* *una* -- DUE TO THE **EVIDENCE** OF A *vida* *insatisfecha,* *echada_a_ perder*]$_{PP}$]$_{NP}$ LIFE DISSATISFIED SPOILED 'It was due to the **evidence** of a dissatisfactory and spoiled life.'	21.73
		Number	(101) (i) [***vydači*** [*gematogena*]$_{NP}$ [*dlja* **ADMINISTRATIONS** HEMATOGENE FOR *bol'nykh*]$_{PP}$]$_{NP\text{-}PL}$ (…) SICK_PEOPLE 'the **administration** of hematogene for the sick people (…)' (ii) [*la* ***administración*** [*de* *hematogeno*]$_{PP}$ THE **ADMINISTRATION** OF HEMATOGENE [*a* *los* *enfermos*]$_{PP}$]$_{NP\text{-}SG}$ TO THE SICK_PEOPLE 'the **administration** of hematogene for the sick people'	8.69
	Order structure	Explicitation	(102) (i) *polučit'* [[*medicinskie*]$_{AP}$ ***znania***]$_{NP}$ TO_RECEIVE MEDICAL **KNOWLEDGES** 'to receive medical **knowledge**' (ii) *recibir* [[*sus*]$_{Poss}$ ***conocimientos*** [*médicos*]$_{AP}$]$_{NP}$ TO_RECEIVE HIS **KNOWLEDGES** MEDICAL 'to receive his medical **knowledge**'	2.89

		Lexical	(103) (i) *Na* [[*utrennykh*] AP *i* [*večernykh*]AP ***poverkakh***]NP IN MORNING AND EVENING **CHECKING'S** *čitalis'* *vesčislennyje* *rasstrelenie* *prikazy* WERE_READ UNCOUNTABLE SHOOTING ORDERS 'In the morning and evening's **checkings** a lot of shooting orders were read.' (ii) *En* [*los* ***recuentos*** [*matinales*]AP]NP *se* *leían* IN THE **RECOUNTS** MORNING -- READ *innumerables* *órdenes* *de* *fusilamiento* UNCOUNTABLE ORDERS OF SHOOTINGS 'In the morning **recounts** a lot of shooting orders were read.'	4.34
		Number	(104) (i) *Osvobožden* *ot* *sdači* [[*ustnykh*]AP FREED FROM PASSING ORAL ***ispytannij***]NP-PL **EXAM** 'He was freed from passing the oral **exam**.' (ii) *Me* *libraron* *de* [*la* ***prueba*** [*oral*]AP]NP-SG Me freed of the **exam** oral 'They freed me from passing the oral **exam**.'	2.89
	Explicitation	Lexical	(105) (i) *Prisutstvovavšaja* *pri* [***vypravlenii*** [*vyvikha*]NP]NP PRESENT AT **REPOSITION** DISLOCATION 'She was present in the **reposition** of the dislocation.' (ii) *presente* *en* [*la* ***operación***]NP PRESENT IN THE **OPERATION** 'She was present in the **operation**.'	5.79
		Number	(106) (i) *Vyderživala* [***sravnenia*** [*s* *neju*]PP]NP-PL STOOD **COMPARISONS** WITH HER 'She stood **comparisons** with her.' (ii) *Resistía* [*la* ***comparación***]NP-SG STAND THE **COMPARISON** 'She stood the **comparison**.'	1.44

	Lexical	Number	(107) (i) *Nevstavanii na* [***poverku***]$_{\text{NP-SG}}$ NOT_RISING IN **CHECKING** 'They didn't stand up for the **control recounts**.' (ii) *No se levantaban para* [*los* ***recuentos*** NO -- STOOD_UP FOR THE **RECOUNTS** ***de control***]$_{\text{NP-PL}}$ **OF CONTROL** 'They didn't stand up for the **control recounts**.'	1.44
	Head swapping	Number	(108) (i) *Traktuetsja kak* [[*samoe čudovishchnoe*]$_{\text{AP}}$ IS_CONSIDERED AS MOST HORRIBLE ***prestuplenie***]$_{\text{NP-SG}}$ **CRIME** 'It was considered the most horrible **crime**.' (ii) *Se consideraba como* [*el peor* [*de los* -- CONSIDERED AS THE WORST OF THE ***delitos***$_{\text{noun-PL}}$]$_{\text{PP}}$]$_{\text{AP}}$ **CRIMES** 'It was considered the worst **crime**.'	1.44

Table 5: Three linguistic changes mismatch

The commonest mismatch is composed of the determiner change, the constituent change and the lexicalization change (21.73%). This is followed by the combination of a determiner and a constituent with a number change or an order change (8.69% and 7.24%, respectively). The rest of mismatches belonging to this type have frequencies that range from 5.79% to 1.44%.

(ii) The PoS group: This group is subdived in three subgroups depending on the two linguistic changes that they share.

			Example	%
PoS	Syntactic function	Constituent structure	(109) (i) *Dostatočnykh dlja* [***prokormlenia*** ENOUGH FOR **ALIMENTATION** [*Gullivera*]NP-NC]NP GULLIVER 'This is enough to **feed** Gulliver.' (ii) *Sufficiente como para* [***alimentar*** [*a* ENOUGH LIKE TO **FEED** TO *Gulliver*]PP-CREG]NP GULLIVER 'This is enough to **feed** Gulliver.'	13.04
		Incorporation	(110) (i) [***napriaženiem*** [*vsego tela*]NP-NC]NP **TENSION** ALL BODY '**tensing** my whole body' (ii) [***poniendo en tensión*** [*todo mi* **PUTTING IN TENSION** ALL MY *cuerpo*]NP-DO]VP BODY '**tensing** my whole body'	1.44
		Explicitation	(111) (i) *vo vremia* [***znakomstva*** [*s Isaem* IN TIME **ACQUAINTANCE** WITH ISAIAH *Rabinovičem*]PP-NC]NP RABINOVICH 'at the time of **acquaintance** with Isaiah Rabinovich' (ii) *en el tiempo en que* [[*me*]NP-DO IN THE TIME IN WHICH I ***relacioné*** [*con Rabinovich*]PP-ADJ]VP **RELATED** WITH RABINOVICH 'at the time I **was in touch** with Rabinovich'	2.89

	Incorporation	Explicitation	(112) (i) *kozy, ukhod za nimi,* [***kormlenie***]$_{NP}$, GOATS GO AFTER THEM **FEEDING** *uborka, dojka* CLEANING MILKING 'going after the goats, **feeding**, cleaning, milking' (ii) *cuidar de las cabras,* LOOK_AFTER OF THE GOATS, [***darles_de_comer,***]$_{VP}$ *limpiarlas, y* **TO_GIVE_THEM_TO_EAT** TO_CLEAN_THEM AND *ordeñarlas* TO_MILK_THEM 'look after the goats, to **give them the feeding**, to clean them and to milk them'	4.34
	Lexical	Syntactic function	(113) (i) *Zhdal* [[*eë*]$_{Poss-Spec}$ ***vozvraščenia***]$_{NP}$ WAITED HER **RETURN** 'He waited for her **return**.' (ii) *Esperaba a que* [[*éste*]$_{NP-Subj}$ ***apareciera***]$_{NP}$ WAITED TO THAT THIS **APPEARED** 'He waited for him to **appear**.'	7.24
		Explicitation	(114) (i) *Ty mne ne poklonilsja pri* [***vstreče***]$_{NP}$ YOU TO_ME NOT SAID_HELLO AT **ENCOUNTER** 'You didn't say hello to me at the **encounter**.' (ii) *No me has saludado cuando* NOT TO_ME HAVE SAID_HELLO WHEN [[*me*]$_{NP}$ ***has visto***]$_{VP}$ ME **HAVE SEEN** 'You didn't say hello to me when you **have seen** me.'	2.89
		Head swapping	(115) (i) *tveriak po* [***roždeniu***]$_{NP}$ TVERIAN ACCORDING **BIRTH** 'Tverian according to **birth**' (ii)[***originario*** [*de tver*]$_{PP}$]$_{NP}$ **NATIVE** OF TVER '**native** from Tver'	1.44

Table 6: Three linguistic changes mismatch

In this second group the commonest subtype is the one that combines the following linguistic changes: PoS, syntactic function and constituent changes (13.04%). It is followed by the subtype composed of the following linguistic changes: PoS, lexical and syntactic function (7.24%). The rest of subtypes have frequencies that range from 4.34% to 1.44%.

(iii) The lexical group: This last type is composed of a lexical change, the constituent change and the number change and it is not very frequent (1.44%).

			Example	%
Lexical	Constituent structure	Number	(116) (i) *podležaščimi ispravleniju, a ne* [[*karatel'nomu*]$_{AP}$ SUBJECT_TO CORRECTION BUT NOT PUNISHING ***vozdejstviju***]$_{NP-SG}$ **INFLUENCE** 'subject to the correction but not to the punishing **influence**' (ii) *No se sometía a* [***penas*** [*de* NO -- SUBJUGATE TO **PENALTIES** OF *castigo*]$_{PP}$]$_{NP-SG}$ PUNISHMENT 'He was not subjected to punishing **penalties**'	1.44

Table 7: Three linguistic changes mismatch

5.5 Conclusions

In this chapter, we have analyzed and described mismatches between Russian and Spanish deverbal nouns. We have obtained a classification of mismatches based on the number and type of linguistic changes found (morphologic, syntactic, semantic and pragmatic), concretely 8 linguistic changes. Our classification has been empirically validated in a sample of 500 sentences including a deverbal noun. The study is focused on those linguistic changes and mismatches that are systematic and productive, although some of them have obtained a low representation. Despite this fact, we have taken them into consideration because they are examples either of the different linguistic changes or of the possible combinations of linguistic changes found.

The usual translation of a Russian deverbal noun is its corresponding deverbal noun or a noun in Spanish (81.44%). The 18.56% left have been translated into other morphological categories such as verbs, deverbal adjectives or past participles.

The translation into a noun (deverbal or not) does not necessarily involve a word-for-word translation. In MiniRuSp, only 16% of cases are literal translations. On the opposite side of the translation mismatches continuum, we have a 3.2% of free translations, that is, mismatches with four or more linguistic changes. Therefore, the general tendency is to have translation mismatches with one, two or three linguistic changes.

In table 8, we summarize the obtained results concerning the frequency of each linguistic change in the overall linguistic changes found.

Linguistic change		**%**
Determiner change		34.17
PoS change		11.22
Structural change		31.26
	Constituent change	16.77
	Syntactic Function change	7.69
	Order change	6.80
Argument Structure change		10.08
	Incorporation change	2.77
	Explicitation change	7.31
Lexical change		8.82
Head swapping change		0.63
Number change		3.02
Discursive change		0.74
	Coreference-anaphoric change	0.37
	Correference-elliptical change	0.37

Table 8: Total frequency of linguistic changes

The commonest linguistic changes are the determiner change (34.17%) and the constituent structural change (16.77%). These two linguistic changes together account for more than the half of all the changes found in our sample (50.94%). It is important to note that both changes are due to typological reasons: on the one hand, Spanish has article determiners, whereas Russian does not; a bare NP in Russian is translated, most of the times, into a specified NP. On the other hand, Russian has morphological case, whereas Spanish does not. A deverbal noun complemented by a

noun phrase in Russian can be realized as an NP complemented by a prepositional phrase. Among preferred prepositions, the usual one is the unmarked *de* 'of'. Regarding the other structural changes, that is, syntactic function and order change, they are also quite frequent. In fact, all the structural changes together represent 31.26% of linguistic changes found in our sample. It is worth noting that the syntactic function change represents 7.69% of all changes and it is triggered by the PoS change (34.07%), since there is no syntactic function change without a PoS change. Constituent function and syntactic function changes take place together in 8.37% of cases. The order change takes place in the overall of MiniRuSp with a frequency of 6.80%.

PoS change takes place in 11.22% of cases. Verbs, adjectives and past participles are the morphological categories used to translate deverbal nouns. Verbs are more common than adjectives or past participles (83.69%, 10.86% and 5.43%, respectively) being the infinitive the preferred category (53.24%). PoS change can be related to the findings presented in our previous study on event structure (see chapter 3). We observed that deverbal nouns translated into verbs were interpreted as events, whereas those translated into non-deverbal nouns were interpreted as results. It seems reasonable because events have predicative character; whereas non-deverbal nouns clearly show the loss of predicate properties. As we have already said in chapter 3, the reason for translating a deverbal noun into a verb or a noun can be the natural tendency to express processes and actions by means of verbal forms, and results (objects) by means of nouns.

Changes related with Argument Structure are also quite frequent (10.08%), concretely the explicitation change, which occurs with a frequency of 7.31% in MiniRuSp. This change can be caused by a PoS change, and this usually entails changes in the number of arguments and consequently syntactic changes too. In other occasions, this change has nothing to do with PoS change and the explicitation of an argument is done simply to be clearer.

The lexical change takes place in 8.82% of cases. The election of a word, which does not correspond totally with the meaning of the original, may end in free translations. However, even though this linguistic change does not necessarily trigger other changes, it can be accompanied by other changes such as the determiner change, the constituent change, the order change, the explicitation change, among others (see 5.4.2).

The rest of linguistic changes: the number change (3.02%), the incorporation change (2.77%), the head swapping change (0.63%), the coreferential anaphoric (0.37%) and the coreferential elision change (0.37%) have low frequencies in our corpus, specially, the last two.

Regarding the classification of mismatches, we have organized them in three groups depending on the number and type of linguistic changes involved. Linguistic changes tend to co-occur. The commonest are those mismatches with two linguistic changes

(41.2%), followed in frequency by mismatches composed of one linguistic change (25.8%) and mismatches composed of three linguistic changes (13.6%).

As we have already noticed, determiner change and constituent change together amount to the half of the overall of linguistic changes found in MiniRuSp (50.94%). Therefore, they are the most frequent in mismatches composed of one, two and three linguistic changes. Regarding mismatches of one linguistic change, the one composed of a determiner change represents 50.38% being clearly the most frequent, and the one composed of a constituent change represents 10.07%. Regarding mismatches including two linguistic changes, the combination of a determiner change and a constituent change is the commonest (34.46%). Finally, among mismatches composed of three linguistic changes, the combination of these two linguistic changes –determiner and constituent- with the lexical change is the commonest mismatch (21.73%).

It is clear that the determiner change and the constituent change are the most systematic of all the changes because they are typologically motivated, if we also include the order change, motivated by typological reasons too, they represent 57.76% of all the linguistic changes found in MiniRuSp.

As future lines of research, we are interested in analyzing to what extent typological changes, that is, changes which are not dependent on translator's choice, can be predicted and systematized. If that is possible, it will suppose that more than the half of linguistic changes (at least in MiniRuSp) can be treated automatically.

On the other hand, this systematization and forecast is probably not possible with the rest of linguistic changes, since they tend to be motivated by stylistic reasons.

Chapter 6

Conclusions and further research

This thesis presents an empirical descriptive study of Russian deverbal nouns following a corpus-based approach. In this chapter we first bring together the results obtained in the three main aspects of deverbal nouns that have been analyzed: (a) the relationship between the lexical denotation of a deverbal noun and the aspect of the base verb, (b) the argument structure of deverbal nouns, and (c) translation mismatches between Spanish and Russian deverbal noun structures (section 6.1). Finally, we highlight the main contributions to the research of Russian deverbal nouns (section 6.2), and future work (section 6.3).

6.1 Concluding remarks

a. *The relationship between the lexical denotation of a deverbal noun and the aspect of the base verb*

The first issue we were concerned about was the relationship between morphological and lexical aspect of the base verb and the lexical denotation of the deverbal nominalization derived.

On the basis of traditional verbal classifications, we distinguish among symmetric, neutralized, biaspectual and uniaspectual nominalizations taking into account the type of the base verb and the corresponding inherited aspectual marks according to its verbal origin, that is, nominalizations derived from both members of the aspectual pair, from only one member of the aspectual pair, and, finally, from verbs that do not have aspectual pair. The aim of this classification is to define the possible correspondences between the aspect of the base verb and that of the nominalization. As it has been claimed by authors such as Vinogradov (1972), Schoorlemmer (1995), Zimmermann (2002), Tatevosov (2003) and Spencer & Zaretskaya (2011), we also assume that these aspectual inherited marks have no grammatical function in deverbal nouns. Assuming this, the experiment carried out to see to what extent the verb base morphological aspect had an influence on the lexical denotation of the

deverbal noun showed that there is a certain influence but not determining: there is a tendency to express an event by means of a nominalization derived from an imperfective base verb, and a tendency to express a result by means of a nominalization derived from a perfective base verb. This is consistent with the fact that imperfective forms tend to denote ongoing processes, whereas perfective forms tend to denote finished actions or actions with a culmination point. However, this fact is not systematic since deverbal nouns derived from imperfective verbs can also occur with a result reading and nouns derived from perfective verbs with an event reading, as seen in the conducted experiment. In fact, there is a loss of aspectual properties which leads to a preference for deriving nouns from imperfective verb forms. The lack of its perfective pair shows that this aspectual opposition is no longer crucial. As it was also claimed by Comrie (1976), in our experiment we also observed that boundaries between predicative readings (that is, between states and events) are better established than between the action and the outcome of that action (that is, between events and results).

Regarding the lexical aspect of the base verb in relation to the lexical denotation of the deverbal noun, we have observed, as in the case of the morphological aspect, that the lexical class of the verb influences the lexical denotation of its corresponding deverbal noun. In Russian, except for the case of activity base verbs which derive event nouns, our observations regarding the verbal lexical classes are in line with the claims of Alexiadou (2001), Picallo (1999), Peris & Taulé (2009) and Fábregas & Marín (2011) for English and Spanish (among other languages). In general, state verbs derive state deverbal nouns, accomplishment and achievement verbs derive both event and result nouns although in the case of achivements, they usually derive nouns with result readings. Nevertheless, in their approaches, activity base verbs derive result deverbal nouns, which is contrary to our findings for Russian and also to Jezek & Melloni's (2009) findings for Italian.

Regarding nouns derived from accomplishments, we have observed that they can express both events and results, but they can be specialized in one particular reading. This can be explained by the tendency of languages to avoid ambiguity by using words that clearly express one denotation. Then the internal argument of the deverbal noun is used to denote the result reading and a verb is used to denote the event reading. Moreover, sometimes the same meaning can be expressed by means of two different deverbal nouns: one deverbal noun is specialized in an event reading, whereas the other is specialized in the result reading, and this responds to the context surrounding the deverbal noun and the frequency of the deverbal noun in the language.

Therefore, these analyses of the relation between the aspects of the base verb and the lexical denotation of the deverbal noun confirmed our initial hypothesis: neither the morphological nor the lexical aspect of the base verb determines straightforwardly the lexical denotation of the deverbal noun, although they have a significant influence on it, and they can be considered as important clues for distinguishing the

denotative interpretation of deverbal nouns in Russian.

We have also analyzed well-known criteria to distinguish between the denotation of deverbal nouns: the pluralization, the expression of the internal argument and the presence of specifiers. Regarding the ability to pluralize we have observed that deverbal nouns with an event reading usually appear in singular while result nouns can appear either in singular or in plural. Therefore, it seems that event nouns are much more restricted in this respect than result nouns. However, as for Spanish and English, in Russian we can find nouns in plural denoting an event which refer to actions that have a culmination point (that is, which are telic) and are repeated a number of times or ordered in time. Regarding the expression of the internal arguments, nouns denoting an event tend to express the internal argument, whereas nouns denoting a result tend not to do so. Regarding specifiers, they are not informative enough in Russian since most of deverbal nouns appear as bare nouns without a specifier.

Therefore, deverbal nouns are influenced by the morphological and lexical aspects of their corresponding base verbs, but the context, in which the deverbal noun appears, is more determinant.

b. The argument structure of deverbal nouns

We presented a descriptive study of the argument structure of Russian deverbal nouns denoting events, states and results. The mapping between syntactic constituents and semantinc arguments basically depends on the syntactico-semantic structure of the corresponding base verb. Despite having the same argument structure of their corresponding base verbs, the arguments of a deverbal noun can be syntactically realized in three different ways: they can be incorporated inside the root of the deverbal noun; they can be realized explicitly in the NP headed by the deverbal noun or they can be implicitly realized outside the NP headed by the deverbal noun. Our corpus-based approach shows that in 43.24% of cases in MiniRuSp, deverbal nouns appeared with no arguments inside the NP headed by the deverbal noun, whereas 40.22% appeared with one argument and 10.15% and 1.31% with two or three arguments, repectively). These figures support the claim about the fact that deverbal nouns focus on the denoted action/result, that is, on the predicate meaning, rather than on the arguments taking part in the action. The information provided by these arguments is usually recoverable from the linguistic and extralinguistic context by means of implicit arguments.

We have observed that constituents that are typically argumental are: NPs, APs, PPs, which act syntactically as noun complements, and Possessive determiners, which function syntactically as specifiers.

The arguments that are syntactically more often realized are Arg1 (42.37%), ArgMs (33.53%) and Arg0 (16.15%). Arg0 and Arg1 are core arguments and, for this reason, they are either more often expressed or can be left implicit, so that they must be

inferred from the linguistic/extralinguistic context. On the other hand, ArgM may not be recovered from the context if it is not realized explicitly and, consequently, it could be lost.

NPs and Possessive determiners tend to express Arg0 and Arg1 core arguments. Concretely, a noun complement syntactically realized by means of an NP expresses mainly an Arg1 and Arg0. The NP appears in genitive most of the times (98%). Possessive determiners only express core arguments either Arg0 or Arg1 being the Arg0 the preferred one. Regarding the rest of determiners (such as demonstrative or negative), they cannot be argumental in Russian. Moreover, NPs in Russian have a strong tendency to appear unspecified.

APs mostly express ArgM, but they can also express core arguments such as Arg0 or Arg1. However, the strong tendency is to be realized as an adjunct, that is, location, manner or time in which the action named by the deverbal noun is carried out.

Finally, PPs are more flexible since we have found this constituent expressing any type of argument core as well as adjunct arguments.

Subordinate clauses have a strong tendency to realize Arg1, despite the fact that they are not very common. They are usually infinitive clauses acting as Arg1.

In the case of NPs with only one argument syntactically realized, the commonest constituents are NPs, APs, PPs and possessive determiners. In the case when two arguments are explicitly realized the most usual combinations are the AP + NP and the AP + AP (26% and 22%, respectively). When three arguments are realized, the usual combination is three APs. Other combinations are also possible but very uncommon and most of the times include the Arg1.

c. Translation mismatches between Spanish and Russian deverbal noun structures

The analysis and observations of the argument structure of deverbal nouns in Russian have been crucial to the study of systematic and productive translation mismatches between Russian and Spanish deverbal noun constructions. We have presented a classification of mismatches based on the number and type of linguistic changes involved (morphologic, syntactic, semantic and pragmatic). The proposed classification has been empirically validated in a sample of 500 translated sentences.

The obtained observations from the corpus-based analysis are the following. The usual translation of a Russian deverbal noun is its corresponding deverbal noun in Spanish, when not, the other morphological categories involved are verbs, deverbal adjectives and past participles. The translation into a noun (deverbal or not) does not necessarily involve a word-for-word translation. In MiniRuSp, only 16% of cases are literal translations, that is, translations without linguistic changes, and 3.2% of cases are free translations, that is, mismatches with four or more linguistic changes.

Therefore, the most common is to find translation mismatches with one, two or three linguistic changes.

The most important linguistic changes are the determiner change and the constituent structural change. These two linguistic changes together account for more than the half of all the changes found in the sample. It is important to note that both changes are related to typological reasons (the inexistence of article determiners and the existence of morphological case, in Russian). Regarding the other structural changes, syntactic function and order change are also quite frequent. In fact, all the structural changes together represent 31.26% of the total, and the most common change is the syntactic constituent one. It is worth noting that the syntactic function change represents 7.69% of all changes and it is triggered by the PoS change (34.07%), since there is no syntactic function change without a PoS change.

PoS change takes place in 11.22% of cases. Verbs, adjectives and past participles are the morphological categories used to translate deverbal nouns. Among the possible categories into which a deverbal noun can be translated, verbs are more common than adjectives or past participles, being the infinitive the preferred category. PoS change can be related to the findings presented in our previous study on event structure (see chapter 3). We observed that deverbal nouns translated into verbs were more often interpreted as events, whereas those translated into non-deverbal nouns were mostly interpreted as results. This is consistent with the fact that events have a predicative character, whereas non-deverbal nouns clearly show the loss of predicate properties. As we have already seen in chapter 3, the usual translation of a deverbal noun into a verb is related to the fact that the verb is the specialized category in expressing processes and actions, while nouns are specialized in denoting objects and results. PoS change can cause changes related to argument structure, which are also quite frequent, concretely the explicitation change. Argument structure changes entail changes in the number of arguments and consequently syntactic changes too. In other occasions, this change has nothing to do with a PoS change and then the explicitation of an argument clarifies the meaning of the sentence.

The lexical change takes place in 8.82% of cases. The election of a word, which does not correspond totally to the meaning of the original deverbal noun, can end in free translations. However, even though this linguistic change does not necessarily trigger other changes, it can be accompanied by changes such as the determiner, constituent, order, and explicitation changes, among others.

The rest of linguistic changes, that is, the number, incorporation, head swapping, coreferential anaphoric and the coreferential elision changes have very low frequencies in our corpus, specially, the last two.

Regarding the classification of mismatches, we have organized them in three groups depending on the number and type of linguistic changes involved. Linguistic changes

tend to co-occur, as we have seen with the PoS change and the syntactic function change. The commonest are those mismatches with two linguistic changes (41.2%), followed by mismatches with one linguistic change (25.8%) and with three linguistic changes (13.6%).

Regarding mismatches of one linguistic change, the more frequent are the determiner and the constituent changes. As for mismatches with two linguistic changes, the combination of determiner and constituent changes is the commonest. Finally, mismatches with three linguistic changes are the result of the combination of the two above mentioned with the lexical change.

Determiner and constituent change are the most systematic of all the linguistic changes because they are typologically motivated. The order change with the other two typological changes represents 57.76% of all the linguistic changes found in MiniRuSp.

6.2 Main contributions

This thesis contributes to the semantic analysis of the Russian deverbal noun from a theoretical point of view, giving rise to a detailed description of both the semantic denotation and the argument structure of nominalizations, and a comparative interlinguistic study between Russian and Spanish deverbal nouns. The linguistic findings can be useful for models of semantic representation and for the development of resources and tools of language technology for the automatic treatment of Russian. Moreover, it can be useful for the development of resources for the teaching/learning of Russian as a second language. Next, we briefly summarize the main contributions of this thesis:

(1) A classification of Russian deverbal nouns in four types: symmetric, neutralized, biaspectual and uniaspectual based on the type of base verb and the morphological aspectual marks inherited. We have also analyzed the lexical denotation of the deverbal noun and its relation with the base verb morphological and lexical aspect. We have also reviewed the widely used criteria to distinguish between the lexical denotation of deverbal nouns, and the detection of denotative selectors to discriminate between event and result nouns in Russian.

(2) A detailed description of the typical patterns of the argument structure of Russian deverbal nouns accompanied by an analysis of the syntactic realization of the semantic arguments.

(3) A classification of the more regular and systematic translation mismatches between Russian and Spanish deverbal nouns.

(4) The creation of a parallel Russian-Spanish corpus, named RuSp, aligned at the paragraph level and consisting of 710,622 tokens.

(5) The syntactico-semantic analysis of MiniRuSp, a subsample of RuSp,

consisting of 84,375 tokens containing 230 different deverbal nouns. We annotated at a syntactico-semantic level a group of 114 different deverbal nouns. For this study we have proposed a series of annotation guidelines that can be the base on which annotate the whole RuSp at a syntactico-semantic level.

6.3 Further research

One of the possible future research lines is the increase of the amount of corpus analyzed. The present thesis is the descriptive analysis upon which we can ground the annotation guidelines to complete the annotation of RuSp and the creation of nominal lexicons with this formalized information. The complete annotation of a bigger corpus tagged with the denotative type of deverbal nouns and also with the semantic annotation of its arguments and thematic roles would be an important resource for a deeper linguistic analysis of Russian deverbal nouns. Moreover a corpus with this annotated information can be a useful resource for NLP applications and for developing didactic material for Russian learning as a second language.

A second line of future research will be the linguistic study of incorporated and implicit arguments of deverbal nouns since from our study we have observed that in 43.24% of cases the NP headed by the deverbal noun has no syntactically realized explicit arguments. Therefore, it would be interesting to analyze these structures with either implicit or incorporated arguments and its relation with the lexical denotation of the deverbal noun.

Regarding translation mismatches, as future lines of research, we are interested in analyzing to what extent typological changes can be predicted and systematized which would lead to the automatic treatment of linguistic changes. It would become an important tool for automatic translation systems allowing the derivation of heuristic rules for the implementation of automatic alignment. Moreover, the detection of these mismatches can also help translators, teachers and learners of Russian as a second language.

Literature

Alexiadou, A. 2001. *Functional Structure in Nominals. Nominalization and Ergativity*. Linguistik Aktuell/Linguistics Today 42. Amsterdam: John Benjamins.

Andersen, Ø. 2007. 'Deverbal nouns as hybrids: Some textual aspects.' Ahmad, Khursnid & Rogers, Margaret (eds.), *Evidence-based LSP. Translation, Text and Terminology*. Bern: Peter Lang, pp. 155-171.

Aparicio, J., Taulé, M. Martí, M. A. 2008. 'AnCora-Verb: A Lexical Resource for the Semantic Annotation of Corpora'. Proceedings of 6th International Conference on Language Resources and Evaluation. Marrakesh (Morocco).

Apresian, J. 1971. 'Regular polysemy'. Izvestija akademii nauk SSSR: Serija literatury i jazyka XXX, pp. 509–523.

Apresian Ju. D., Boguslavsky I. M. & Iomdin B. L. 2005. 'Sintaksičeski i semantičeski annotirovannyj korpus russkogo jazyka: sovremennoe sostojanie i perspektivy'. Nacional'nyj korpus russkogo jazyka: 2003-2005. M.: Indrik. pp. 193-214.

Baker, M. 1995. 'Corpus-based translation studies: The challenges that lie ahead'. Somers, H. (ed.), *Terminology, LSP, and translation: studies in language engineering in honour of Juan C. Sager*. Amsterdam: John Benjamins, pp. 175-186.

Baker, C. F., Fillmore Ch. J. & Cronin B. 2003. 'The Structure of the FrameNet Database'. International journal of lexicography. Vol. 16, pp. 281-296.

Balvet, A., Barque, L. & Marín, R. 2010. 'Building a lexicon of French deverbal nouns from semantically annotated corpus'. Proceedings of the Seventh conference on International Language 22. European Language Resources Association (ELRA).

Balvet, A., Barque, L., Condette, M. H., Haas, P., Huyghe, R., Marín, R. & Merlo, A. 2011. 'Nomage: an electronic lexicon of French deverbal nouns based on a semantically annotated corpus.' Proceedings of the International Workshop on Lexical Resources (WoLeR) at European Summer School in Logic, Language and Information (ESSLLI 2011).

Braginsky P. & Rothstein S. 2008. 'Vendlerian Classess and the Russian Aspectual System'. Journal of Slavic Linguistics, Vol. 16, Number 1, pp. 3-55.

Brandtner, R. & Heusinger K. 2009. 'Nominalization in Context - Conflicting Readings and Predicate Transfer'. Artemis Alexiadou & Monika Rathert (eds.), *Nominalizations across Languages and Frameworks*. Berlin: de Gruyter (Interface explorations), pp. 25-50.

Brazilay, R. & McKeown, K. R. 2001. 'Extracting Paraphrases from a Parallel Corpus'. ACL annual meetings.

Burchardt, A., Erk, K., Frank, A., Kowalski, A., Padó, S., & Pinkal, M. 2009. 'FrameNet for the semantic analysis of German: Annotation, representation and automation'. Boas, H. C., (ed.), *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. Mouton de Gruyer.

Comrie, B. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge University Press.

Chomsky N. 1970. 'Remarks on Nominalization'. Jacobs, R. & Rosenbaum P. (eds.) *Readings in English Transformational Grammar*. Waltham, Mass.: Ginn and Co, pp. 184-221.

de Valdivia, G., Castellví, J. & Taulé, M. 2013. 'Morphological and lexical aspect in Russian deverbal nominalizations'. Kor Chahine, I. (ed.), *Current Studies in Slavic Linguistics*. xi, Studies in Language Companion Series, John Benjamins, pp. 267–280.

Dorr, B. 1990. 'Solving Thematic Divergences in Machine translation'. Proceedings of the 28th Annual Conference of the Association for Computational Linguistics, University of Pittsburgh, Pittsburgh, PA, pp. 127-134.

Dorr, B., Lee J. & Suh, S. 1994. 'Development of Cross-Linguistic Syntactic and Semantic Parameters and Techniques for Machine Translation'. UMIACS-TR-94-26, CS-TR-3233, University of Maryland, College Park, MD.

Dorr, B., Ayan, N. F. & Habash, N. 2004. 'Divergence unraveling word alignment of parallel corpus'. Natural language Engineering 1(i).

Eberle, K., Faaß, G. & Heid, U. 2009. 'Corpus-based identification and disambiguation of reading indicators in German nominalizations'. *Language Resources and Evaluation Conference,* Liverpool, UK.

Fernández, A. M. 2000. *The analysis of verbal lexical items and translation mismatches in English and Spanish: A theoretical framework*. Universitat Autònoma de Barcelona.

Fábregas, A. & Marín, R. 2011. 'The role of Aktionsart in deverbal nouns: State nominalizations across languages'. J. Linguistics 48, Cambridge University Press, pp. 35–70.

Fillmore, C. J. 1968. 'The case for case'. Bach, E. W. & Harms, R. T., (eds), *Universals in Linguistic Theory,* Holt, Rinehart & Winston, New York, pp 1–88.

Fleiss, J. L. 1971. 'Measuring nominal scale agreement among many raters'. *Psychological Bulletin* 76(5), pp. 378–382.

Galambos, A. 2007. 'Primary and secondary imperfectives in Russian: A cumulativity analysis'. LSO Working Papers in Linguistics: Proceedings of WIGL.

Goldberg, A. 1995. *Construction: a construction grammar approach to argument structure.* University of Chicago Press.

Grimshaw, J. 1990. *Argument Structure.* Cambridge MA: The MIT Press.

Høeg Müller, H. 2011. 'The Copenhagen Dependency Treebank (CDT) Extending syntactic annotation to morphology and semantics'. International Conference on Dependency Linguistics.

Jackendoff, R. 1983. *Semantics and Cognition.* Cambridge, MA: MIT Press.

Jackendoff, R. 1990. *Semantic Structures.* Cambridge, MA: MIT Press.

Ježek, E. & Melloni, Ch. 2009. 'Complex type in the (morphologically) complex lexicon', N. Calzolari & A. Rumshisky (eds), Proceedings of the 5th International Conference on Generative Approaches to the Lexicon, (Sept. 17-19, 2009), Pisa, pp. 59-67.

Ježek, E. & Melloni, Ch. 2011. 'Nominals, polysemy and co-predication'. *Journal of Cognitive Science* 12, pp. 1–31.

Komur, G. 2005. 'Le transfert de l'aspect verbal sur le substantif en polonais'. *Cahiers Chronos 12: Temporalité et attitude: Structuration du discours et expression de la modalité*, Amsterdam-NewYork, pp.143-158.

Landis, J. R. & Koch, G. G. 1977. 'The measurement of observer agreement for categorical data'. *Biometrics* 33, pp. 159–174.

Levin B. 1993. *English Verb and Alternations: A Preliminary Investigation*. University of Chicago Press.

Liashevskaya O. N., Plungian V. A. & Sichianava D. B. 2005. 'O morfologičeskom standarte nacional'nogo korpusa russkogo jazyka'. Rezultaty i perspektivy, pp. 111-135.

Marín R. 2002. *El componente aspectual de la predicación*. PPU, Barcelona.

Martí M. A., Taulé M., Bertran M. & Márquez Ll. 2008. 'AnCora: Multilingual and Multilevel Annotated Corpora' (*pending to be published*).

Mel'chuk, I. A. 1981. 'Meaning-Text Models: A recent trend in Soviet linguistics'. Annual Review of Anthropology 10, pp. 27–62.

Mel'chuk, I. A. 1974. 'Opyt teorii lingvisticheskikh modelej *smysl'-tekst*'. Moskva.

Mel'chuk, I. & Wanner, L. 2001. 'Towards a Lexicographic Approach to Lexical Transfer in Machine Translation (Illustrated by the German-Russian Language Pair)'. Machine Translation 16, pp. 21-87.

Mel'chuk, I. & Wanner, L. 2006. 'Syntactic Mismatches in Machine Translation'. Machine Translation, 20: 2, pp. 81-138.

Meyers, A., Reeves, R., Macleod, C., Szekeley, R., Zielinska, V., & Young, B. 2004. 'The Cross-Breeding of Dictionaries'. Proceedings of Fourth International Conference on Language Resources and Evaluation (LREC-04), Lisbon, Portugal.

Meyers, A. 2007. 'Annotation Guidelines for NomBank Noun Argument Structure for PropBank'. Technical report, University of New York. Molendijk & Co Vet (eds), Amsterdam: Rodopi, pp. 143–158.

Miller, G. A. 1995. 'WordNet: A lexical Database for English'. Communications of the ACM. Vol. 38, No. 11, pp. 39-41.

Mourelatos, A. 1978. 'Events, Processes and States'. Linguistics and Philosophy 2, pp. 415-434.

Offord D. 1993. *Modern Russian. An advanced grammar course*. Bristol Classical Press.

Ohara, K. 2009. 'Frame-based contrastive lexical semantics in Japanese FrameNet: The case of risk and kakeru'. Boas, H. C., (ed), *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. Mouton de Gruyer.

Ozhegov S. & Shvedova N. 1992. *Tolkovyj slovar' russkogo jazyka.* Rossiskaja Akademia Nauk. Institut russiskogo jazyka im. V. V. Vinogradov.

Paducheva E., Rakhilina E. & Filipenko M. 1992. 'Semantic dictionary as a lexical database'. COLING, pp. 1295-1299.

Paducheva, E. V. 1996. *Semantičeskie issledovanija*. Moscow: Jazyki Russkoi Kul'tury.

Palmer, M, Gildea, D. & Kingsbury P. 2005. 'The Proposition Bank: An Annotated Corpus of Semantic Roles.' Computational Linguistics, 31:1, pp. 71-105.

Pazelskaya, A. & Tatevosov, S. 2003. 'Nominalization in Russian: Eventuality types and aspectual properties'. Paul Dekker (ed), *Proceedings of the XVII Amsterdam Colloquium*, Amsterdam: Institute of Logic, Language and Computation, pp. 169–175.

Pazelskaya, A. 2009. 'Modeli derivacii otglagol'nyx suščestvitel'nyx: vzgljad iz korpusa'. K. L. Kiseleva, V.A. Plungjan, E.V. Raxilina, S.T. Tatevosov (eds). *Korpusnaya grammatika russkogo jazyka*. Moskva: IMLI RAN.

Peris, A. & Taulé, M. 2009. 'Evaluación de los criterios lingüísticos para la distinción evento y resultado en los sustantivos deverbales'. Proceedings of the 1st International Conference on Corpus Linguistics (CILC-09), Murcia, Spain.

Peris, A., Taulé, M., Rodríguez, H. 2010. 'Semantic Annotation of Deverbal Nominalizations in the Spanish corpus AnCora'. Proceedings of The Ninth International Workshop on Treebanks and Linguistic Theories (TLT9), University of Tartu, Estonia, pp. 187-198.

Peris A. & Taulé M. 2011. 'AnCora-Nom: A Spanish Lexicon of Deverbal Nominalizations', Procesamiento del Lenguaje Natural, Revista nº 46.

Peris, A., Taulé, M. & Rodríguez, H. 2012. Empirical methods for the study of denotation in nominalizations in Spanish. *Computational Linguistics* 38(4).

Peris, A., Taulé, M., Rodríguez, H., Bertran, M. 2013. 'LIARc: Labeling Implicit Arguments in Spanish deverbal nominalizations'. Computational Linguistics and Intelligent Text Processing - 14th International Conference, CICLing 2013, Samos, Greece, March 24-30, Proceedings, Part I. Springer, Lecture Notes in Computer Science, Berlín, Alemania. Vol. 7816, pp. 423-434.

Pesetsky, D. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.

Plungjan V. & Sičinava D. 2004. 'Nacional'nyj korpus russkogo jazyka: opyt sozdania korpusov tekstov sovremennogo russkogo jazyka'. L. N. Beljaeva *et al.* (eds), *Trudy mežd. Konferencii 'korpusnaja lingvistika'*, pp. 216-238.

Picallo, C. 1999. 'Nominals and Nominalizations in Catalan'. *Probus* 3, pp. 279–316.

Pustejovsky, J. 1995. *Generative Lexicon*. Cambridge MA: The MIT Press.

Rafalovich A. & Dale R. 2009. 'United Nations General Assembly Resolutions: A six-language parallel corpus'. Proceedings of the MT Summit XII, Ottawa, Canada, August, pp. 292-299.

Recasens, M. & Martí, M. A. 2010. 'AnCora-CO: Coreferentially annotated corpora for Spanish and Catalan'. Language Resources and Evaluation, Springer Science, pp. 315-345.

Roy I. & Soare E. 2011. 'Nominalizations: New insights and theoretical implications'. Recherches linguistiques de Vincennes, 40, pp. 7-23.

Scott, W. A. 1955. 'Reliability of content analysis: The case of nominal scale coding'. *Public Opinion Quarterly* 19(3), pp. 321–325.

Schoorlemmer, M. 1995. *Participal Passive and Aspect in Russian*. Ph.D. dissertation, University of Utrecht.

Spencer, A. & Zaretskaya, M. 2010. 'The Essex Database of Russian Verbs and their Nominalizations'. http://privatewww.essex.ac.uk/~spena/res_interests.html.

Steiner, E. 2001. 'Translations English-German: Investigating the relative importance of systemic contrasts and of the text type translation'. SPRIK reports 7, pp. 1-49.

Subirats, C. 2009. 'Spanish FrameNet: A frame semantic analysis of the Spanish Lexicon'. Boas, H. C., (ed.), *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. Mouton de Gruyer.

Švedova, N. Ju. 1982. *Russkaja grammatika*. Moskva: Nauka.

Taulé, M., Martí M. A. & Recasens, M. 2008. 'Ancora: Multilevel Annotated Corpora for Catalan and Spanish'. Proceedings of 6th International Conference on Language Resources and Evaluation. Marrakesh (Morocco).

Tiedemann, J. 2009. 'News from OPUS - A Collection of Multilingual Parallel Corpora with Tools and Interfaces'. N. Nicolov, K. Bontcheva, G. Angelova & R. Mitkov (eds), *Recent Advances in Natural Language Processing*. Vol. V, John Benjamins, Amsterdam/Philadelphia, pp. 237-248.

Vázquez, G., Fernández, A. & Martí M. A. 2000. 'Dealing with lexical semantic mismatches between English and Spanish'. Proceedings on the International Conference on Knowledge Based Computer Systems, pp. 308-319.

Vendler, Z. 1957. *Linguistics in Philosophy*. Ithaca NY: Cornell University Press.

Verkuyl, H. 1989. 'Aspectual classes and aspectual composition'. Linguistics and Philosophy, 12, pp. 39-94.

Verkuyl, H. 1993. *A theory of Aspect*. Cambridge: Cambridge University Press.

Vinogradov, Viktor V. 1972. *Russkij jazyk*. Moskva: Vyssšaja Škola.

Vila, M,. Martí, M. A. & Rodríguez, H. 2013. 'Is This a Paraphrase? What Kind? Paraphrase Boundaries and Typology.' Open Journal of Modern Linguitics, 4.

Wade T. 1992. *A Comprehensive Russian Grammar*. Blackwell Publishing.

Xue, N. 2006. 'Annotating the Predicate-Argument Structure of Chinese Nominalizations'. Proceedings of the International Conference on Language Resources and Evaluation (LREC).

Xue N. 2008. 'Labeling Chinese Predicates with Semantic Roles'. Computational Linguistics. Vol. 34, Number 2.

Zalizniak, A. 1977. *Grammatičeskij slovar' russkogo jazyka*. Izdatel'stvo Russkij jazyk.

Zimmermann, I. 2002. 'Process Nominalizations in Russian'. *ZAS Papers in Linguistics* 27, pp. 125–140.

Zubizarreta, M. L. 1987. *Levels of Representation in the Lexicon and in the Syntax*. Foris. Dordrecht.

RNC: Russian National Corpus – www.ruscorpora.ru

Appendices

A List of deverbal nouns analyzed in MiniRuSp

In table 1 we present the number of occurrences of the deverbal nouns analyzed both in RuSp and in MiniRuSp corpora.

DEVERBAL NOUN	**Occurrences in RuSp**	**Occurrences in MiniRuSp**
Vyraženie 'expression'	81	10
Služba 'service'	76	12
Upravlenie 'direction', *Okhrana* 'guarding'[46]	57	10
Otdelenie 'department'	55	10
Igra 'game'	53	10
Komandirovka 'bussiness trip'	42	10
Želanie 'desire'	39	10
Posylka 'sending', *Sostojanie* 'state'	37	10
Zabor 'fence'	33	10
Rešenie 'decision', *Khod* 'move', *Vstreča* 'meeting'	32	10
Prestuplenie 'crime', *Znakomstvo* 'acquaintance'	29	10
Znanie 'knowledge'	28	10
Smena 'change'	27	10
Prisutstvie 'presence'	26	10
Soznanie 'consciousness'	25	13
Rasstrel 'shooting'	24	10
Kraska 'paint'	24	1
Otdykh 'break'	23	11
Roždenie 'birth'	23	10
Peresylka 'sending'	21	10
Napravlenie 'direction'	21	3
Molčanie 'silence'	20	14
Razrešenie 'decision'	19	10
Primer 'example', *Opravdanie* 'justification', *Spasenie*	18	10

[46] In every cell with more than one deverbal noun we must consider the number of occurrences in the second and third cell is for each deverbal noun.

'salvation', *Objasnenie* 'explanation', *Okončanie* 'end', *Vozvrašenie* 'return', *Peremena* 'change'		
Zapis' 'note'	17	15
Pitanie 'nourishment', *Sušestvovanie* 'existence'	17	10
Poverka 'verification'	16	10
Vkhod 'entry'	14	2
Obeščanie 'promise', *Obman* 'lie'	14	1
Okraska 'painting'	14	2
Ožidanie 'waiting'	13	1
Vypolnenie 'accomplishment'	12	5
Organizacija 'organization'	12	3
Vyezd 'departure', *Ispytanie* 'test'	11	2
Kraža 'stealing'	11	3
Lečenie 'healing'	11	4
Zaderžka 'delay', *Zakaz 'order', Osnovanie* 'foundation'	10	1
Vozmuščenie 'indignation'	9	1
Naprjaženie 'effort'	9	2
Oborot 'turn'	8	1
Vzryv 'burst'	7	4
Vskrytie 'opening; autopsy', *Vyigryš* 'gain'	7	1
Naznačenie 'fixing', *Osmotr* 'examination'	7	2
Vozdejstvije 'influence', *Dobyča* 'prey', *Izdanie* 'publication'	6	1
Vydača 'delivery'	6	3
Korm 'feeding', *Naselenie* 'colonization'	6	2
Vypiska 'extraction', *Nadzor* 'supervision', *Opora* 'support'	5	2
Ispravlenie 'correction', *Nabor* 'admission; collection'	5	1
Vystuplenie 'performance', *Zagotovka* 'procurement', *Izveščenie* 'notification', *Nadbavka* 'markup'	4	1
Zakhoronenie 'burial', *Kvalifikacija* 'qualification'	4	2
Vklad 'investment'	3	1
Dežurstvo 'watching', *Delenie* 'division', *Zapor* 'lock', *Razdel* 'division'	3	2
Metka 'marking'	3	1
Vyrez 'cut'	2	1
Dojka 'milking'	2	2
Zamer 'measurement', *Zimovka* 'wintering', *Zov* 'call', *Izobretenie* 'invention', *Illjustracija* 'illustration', *Lovlja* 'catching', *Myšlnie* 'thinking', *Kolonizacija* 'colonization', *Obmen* 'exchange', *Oplata* 'payment'	2	1
Izmerenie 'measuring'	2	2
Vpravlenie 'reposition', *Vyplata* 'payment', *Gorenie* 'burning', *Zavedovanie* 'management', *Zazemlenie* 'grounding', *Izbienie* 'beating', *Iskaženie* 'distortion', *Kopka* 'digging', *Kormlenie* 'feeding', *Nakal* 'heat', *Napolnenie* 'filling', *Obnovlenie* 'renewing', *Prokormlenie* 'feeding'	1	1

Table 1: Deverbal nouns analyzed

B Correspondences of the Argument Structure of deverbal nouns in Russian and Spanish

In appendix B, we present translations into Spanish of the different deverbal noun's arguments found in MiniRuSp corpus. In the first column, we find the translation of the deverbal noun, that is, $N_{dev} > N_{dev}$ means that the source deverbal noun is translated into a deverbal noun in Spanish. In the second column we find the category of the argument in Russian, while in the third its counterpart in Spanish. In the fifth column we find an example of use extracted from MiniRuSp in both languages.

Syntactic realization of Arg0

The Arg0 can have the following realizations and correspondences in the two languages. In Russian, the Arg0 is mainly expressed by a possessive determiner or by an NP in genitive. It can also be expressed by an AP and by a PP.

The corresponding translation into Spanish obviously depends first on the translation of the deverbal noun. If the deverbal noun is translated into a verb ($N_{dev} > V$), the possessive determiner or the NP in genitive are translated into a pronoun or into a NP acting as subjects. If the deverbal noun is translated into a noun ($N_{dev} > N_{dev}$), then the possessive determiner is preserved, as well as the PP or the AP. In other occasions the original AP or NP is not preserved in the translation and they are translated into PPs. In some occurrences, the Arg0 is made implicit or explicit.

	Russian	**Spanish**	**Example**
$N_{dev} > N_{dev}$	PP	PP	*Vozmuščenie u tekh* *La **indignación** de aquellos*
	AP	PP	*Blatnye **vyraženija*** *Las **expresiones** de los criminales*
	AP	AP	*Na sosednej vitaminnoj **komandirovke*** *Una **expedición** vitamínica vecina*
	--	Poss. Det	*Proverjaja sistematičeski **zapisi*** *Comprobando sistemáticamente nuestros **apuntes***
	Poss. Det.	Poss. Det.	***Zimovka** naša* *Nuestra **residencia** de invierno*
	Poss. Det.	--	*Dlja moikh **znakomstv*** *Para tener todos aquellos **conocidos***
	NP	PP	*Po **vyraženiju** Dostojevskogo* *En **expresión** de Dostoyevski*
$N_{dev} > N_{no\text{-}dev}$	Poss. Det	Poss. Det.	*Na ego **zov*** *A su **voz***
	AP	AP	*Po sovstvennomu **želaniju*** *Por propia **voluntad***

N_{dev} > V	Poss. Det.	Pronoun	*Zdal eë* ***vozvraščenija*** *Esperaba a que éste* ***apareciera***
	NP-GEN	NP	***Dobyča*** *Savčenko* *Una presa que* ***había cazado*** *Savchenko*

Table 1: Arg0 syntactic realizations in Russian and Spanish

Syntactic realization of Arg1

In Russian the Arg1 can expressed mainly by means of a NP in genitive (even in instrumental), but also by means of an infinitive, a PP or an AP. In other occasions, it can be made implicit or explicit depending on discourse preferences. As in the previous case the translation of the complements depends on the translation of the deverbal noun. Therefore, if a deverbal noun is translated into a verb (an infinitive, for instance) the NP in genitive can be translated into a NP acting as direct object.
However, the main translation of a noun's complement in NP in genitive is a PP. Arg1 expressed by means of a PP is translated into a PP, most of the times. Arg1 expressed by means of AP are translated into APs or PPs. Arg1 realized by means of an infinitive are translated into a PP, which includes the infinitive. In some occurrences the Arg1 is incorporated into the deverbal noun in Russian, in these cases if the target language lacks a noun with the Arg1 incorporated, then it is added a PP expressing it.

	Russian	**Spanish**	**Examples**
N_{dev}> N_{dev}	NP	--	*Pri* ***vypravlenii*** *vyvykha* *En la* ***operación***
	NP-GEN	PP	*na* ***vypolnenie*** *plana* *en función del* ***cumplimiento*** *del plan*
	NP-GEN	AP	***delenija*** *kromosom* *la* ***division*** *cromosómica*
	Inf	PP-inf	*Strastnoe* ***želanie*** *izbavit'sja* *Aquel ardiente* ***deseo*** *de liberarse*
	PP	PP	***Zaderžka*** *na rabote* ***Un retraso*** *en el trabajo*
	Incorporated	PP	***Zazemlenie*** *Las* ***tomas*** *de tierra*
	AP	AP	*Poluchit' medicinskie* ***znanija*** *Recibir sus* ***conocimientos*** *médicos*
	AP	PP	*Lagernogo* ***nadzora*** *La* ***vigilancia*** *del campo*
	--	Poss. Det.	*Eto tebe nužno dlja* ***lečen'ja*** *Te conviene para tu* ***tratamiento***
	AP	incorporated	*Rybnaja* ***lovlja*** *La* ***pesca***
	Poss. Det.	Poss. Det.	*Vaše nervnoje* ***sostojanie*** *Viendo su* ***nerviosismo***

	NP	O. Subord. Rel.	*Ogorčennyj beznadežnym* ***sostojanijem*** *otca* *Dolido por la lamentable* ***situación*** *en que se hallaba su padre*
N_{dev}> N_{dev}	NP-GEN	PP	***Razdel*** *svojej nauki* *Las* ***partes*** *de la anatomía*
	NP-INST	PP	***Zavedovanie*** *otdeleniem* *La* ***jefatura*** *de la sección*
N_{dev} > V	--	Pron	*Pri* ***vstreče*** *Cuando me* ***has visto***
	NP-GEN	NP	*Sposobnost'* ***vydači*** *znanij* *La facultad* ***de exponer*** *algún conocimiento*
	NP	Pron.	*Dlja* ***obmena*** *na tabak pajku kleba* *Para* ***cambiar****la por tabaco*
	Inf	O. Sub. Subst.	*Isprosil* ***razreščenija*** *zamenit' s Narym na Kolymu* ***Consiguió*** *que lo trasladaran de Narim a Kolimá*

Table 2: Arg1 syntactic realizations in Spanish and Russian

Syntactic realizations of Arg2

The Arg2 can be expressed by a PP, which can be translated into a PP in the target language. It can also be expressed by means of an AP and translated into an AP or being incorporated into the root of the deverbal noun. Arg2 also can be expressed by means of a NP and translated into a AP. Finally, an Arg2 can be expressed by an infinitive and translated into a PP.

	Russian	**Spanish**	**Example**
N_{dev} > N_{dev}	PP	PP	***Vydači*** *gematogena dlja bol'nykh* *La* ***administración*** *de hematógeno a los enfermos*
	AP	AP	*Sledil za moral'nym* ***sostojaniem*** *Controlaba el* ***estado*** *moral*
	AP	incorporated	*Vaše nervnoje* ***sostojanie*** *Viendo su* ***nerviosismo***
	NP	AP	*Nakhodilsja v* ***sostojanii*** *oshelomnenija* *Se encontraba en un* ***estado*** *atolondrado*
N_{dev} > N	Inf	PP	*V* ***sostojanii*** *žertvovat'* *Estoy en* ***condiciones*** *de sacrificar*

Table 3: Arg2 syntactic realizations in Spanish and Russian

Syntactic realizations of Arg3

The Arg3 can be expressed by a PP in the original Russian and be translated into a NP in the target Spanish.

	Russian	Spanish	Example
N_{dev} > V	PP	NP	*Prava* ***vyezda*** *s Kolymy* *Derecho a* ***abandonar*** *Kolimá*

Table 4: Arg3 syntactic realizations in Spanish and Russian

Syntactic realizations of Arg4

The Arg4 can be expressed by a PP in Russian and can be translated by a PP in the target language.

	Russian	Spanish	Example
N_{dev} > V	PP	PP	***Vyezd*** *na materik* ***Viajar*** *al continente*

Table 5: Arg4 syntactic realizations in Spanish and Russian

Syntactic realizations of the ArgM

ArgMs in Russian are mainly expressed by PPs and APs and translated into the same constituents in the target Spanish. They can also be translated into adverbial locutions or into subordinate clauses. Sometimes, the AP is incorporated into the root of the deverbal noun. An ArgM can be expressed by a NP and then be translated into a AP. ArgM can be implicit in the target language and can be made explicit in the target language. As always, the translation of its components depends on the translation of the deverbal noun, then a deverbal noun translated into a verb which have an ArgM expressed by means of an AP can be translated into an adverb.

	Russian	Spanish	Example
N_{dev}> N_{dev}	PP	PP	*v slučae polnogo* ***razryva*** *s sovetskim prošlym* *en caso de* ***ruptura*** *con el pasado soviético*
	AP	AP	*skrupuloznogo* ***vypolnenija*** *trebovanij* *el* ***cumplimiento*** *escrupuloso de las normas*
	AP	PP	*zemnoe* ***suščestvovanie*** *Belikova* *el* ***tránsito*** *de Bélikov por este mundo*
	NP	AP	***upravlenija*** *lagerja* *la* ***administración*** *penitenciaria*
	--	PP	*zdanie otrjada* ***okhrany*** *el edificio de la* ***guardia*** *del campo*
	--	AP	*na k****h****odu* *de* ***andar*** *ligero*
	AP	Incorporated	*obratnyj* ***khod*** *ryby is ruch'ev* *el* ***retorno*** *de los peces de los torrente*
	AP	Loc. Adv.	*lič**naja*** ***vstreča*** *un* ***encuentro*** *cara a cara*
	AP	--	*eto lagernoje* ***prestuplenije***

			por este ***delito***
	PP	Incorporated	***peremena*** *žizni* *k luščemu* *toda* ***mejora*** *en su vida*
	PP	SubC	***peremena*** *v eje lice porazila* *el* ***cambio*** *que se había producido en su cara me asombró*
N_{dev} > N	AP	AP	*vnezapnaja* ***vypiska*** *repentino* ***paquete***
N_{dev} > V	AP	Adv	*iduščego prjamym* ***khodom*** *v ad* *llevaba directamente al infierno*

Table 6: ArgM syntactic realizations in Spanish and Russian