OSLa
Oslo Studies in Language
2 (2) / 2010

Atle Grønn
& Irena Marijanović (eds.)

Russian in Contrast. Lexicon

Oslo Studies in Language
General editors: Atle Grønn and Dag Haug

OSLa
Oslo Studies in Language
2 (2) / 2010

Atle Grønn
& Irena Marijanović (eds.)

Russian in Contrast. Lexicon

Oslo Studies in Language, 2(2), 2010.
Atle Grønn & Irena Marijanović (eds.):
Russian in Contrast. Lexicon
Oslo, University of Oslo

ISSN 1890-9639

Set in LaTeX fonts Gentium Book Basic and Linux Libertine by
Vladyslav Dorokhin, Atle Grønn and Pavel Iosad
Cover design by UniPub publishing house
Printed by UniPub from camera-ready copy supplied by the editors

http://www.journals.uio.no/osla

Contents

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 243–259. (ISSN 1890-9639)
http://www.journals.uio.no/osla

RUSSIAN WORD-FORMATION IN CONTRAST WITH CZECH AND NORWEGIAN

LAURA A. JANDA
University of Tromsø

ABSTRACT

Although some scholarship on metonymy has mentioned examples of word-formation and some scholarship on word-formation has mentioned the presence of metonymy, to date there has been no attempt to examine the systematic role of metonymy in the domain of word-formation. This article presents an empirical study of the metonymies signalled by derivational suffixes in Russian, Czech and Norwegian. This approach facilitates cross-linguistic comparison that reveals significant differences among languages, and these differences appear relevant to cultural differences.

[1] INTRODUCTION

Examples (1) and (2) illustrate lexical metonymy with parallels to word-formation in Russian and Czech.

(1) PART FOR WHOLE

a. We need *a good head* for this project.
b. (Russian) *brjux*an (lit. 'belly'-an) 'person with a large belly'
c. (Czech) *břich*áč (lit. 'belly'-áč) 'person with a large belly'

(2) CONTAINED FOR CONTAINER

a. *The milk* tipped over. (cf. Peirsman & Geeraerts 2006: 281)
b. (Russian) *saxar*nica (lit. 'sugar'-nica) 'sugar-bowl'[1]
c. (Czech) *květin*áč (lit. 'flower'-áč) 'flower-pot'

In keeping with current scholarly convention, the metonymy designations in (1) and (2) are identified in terms of VEHICLE FOR TARGET (Kövecses & Radden 1998; Radden & Kövecses 1999; Panther & Thornburg 1999, 2007; Peirsman & Geeraerts 2006). In (1) a body PART (head or belly) serves as the VEHICLE through which the speaker refers to the TARGET, which is a WHOLE person. Although the speaker in (1-a) is primarily interested in the brains of the future employee, they must hire

[1] In this article I do not attempt to distinguish between suffix and desinence (as in Russian *-nica*, which could be segmented as *-nic-a*) because issues of morphophonemics make this segmentation difficult and/or artificial in some cases, and for the purposes of this article, such segmentation is not necessary.

an entire person, and this is a classic example of lexical metonymy. In (1-b) and (1-c) we see the same PART FOR WHOLE metonymy at work, this time signalled by the presence of a suffix. Metonymy is not limited to PART FOR WHOLE relationships; it covers a range of other types of contiguity, including CONTAINED FOR CONTAINER as illustrated in (2). In (2-a) it is not the milk itself that has tipped over, but the glass or carton that the milk is (or was) in. Both Russian and Czech invoke this metonymy when they derive the names for CONTAINERS from nouns that refer to what is CONTAINED in them in (2-b) and (2-c).

Several important observations can be made even on the basis of this modest set of examples. Firstly, there are clear parallels between the use of metonymy in lexicon and in word-formation. Secondly, the presence of a suffix does not in and of itself determine the metonymy designation: we see that Czech *-áč* can be used for both PART FOR WHOLE and CONTAINED FOR CONTAINER. This same suffix can signal a third metonymy designation, namely CHARACTERISTIC FOR ENTITY, as we see in the word *naháč* 'naked person', derived from the adjective *nahý* 'naked'. The Russian suffix in (2-b), *-nica*, can likewise signal a second metonymy: ACTION FOR LOCATION, as in *varnica* 'saltworks' derived from the verb *varit'* 'cook'. Some suffixes are even more versatile, with the maximum number of metonymy designations per suffix at sixteen for Czech, fifteen for Russian, and eleven for Norwegian. A third observation is that the extent of metonymic word-formation is different in different languages: whereas both PART FOR WHOLE and CONTAINED FOR CONTAINER are robustly represented in Russian (nine and three suffixes respectively) and Czech (nine and eleven suffixes respectively), neither of these metonymy designations are attested for Norwegian, which has overall a more restricted system of affixal word-formation. This third observation demonstrates that word-formation systems provide a basis for discovering significant differences among languages. Discovering such differences is the main objective of this article.

[1.1] *Relevant previous scholarship*

The scholarly literature relevant to this article can be divided into roughly two groups: there are works on metonymy that make only scattered mention of word-formation, and there are works on word-formation that rarely make reference to metonymy. This subsection provides a brief overview, focusing on the works that indicate a connection between word-formation and metonymy.

Langacker (2009, 1993) describes metonymy as a pervasive phenomenon, not only in language, but in cognition in general. If this is the case, then we should expect metonymy to play a role across the entire spectrum of the grammar-to-lexicon continuum. However, the majority of linguistic scholarship on metonymy has focused nearly exclusively on the lexicon. Peirsman & Geeraerts (2006) pro-

vide a state-of-the-art inventory of scholarly works on metonymy, but their compilation is nearly exclusively focused on lexical use of metonymy.[2]

Padučeva (2004, 147,163), in a discussion of lexical metonymy, mentions that a metonymy expressed lexically in one language might be expressed via word-formation in another and gives two examples. Though Koch (1999) specifically argues that metonymy plays a major role in word-formation, he offers only a pair of examples from the history of French. Several works analyze metonymy in a single affix or a small group of affixes: Basilio (2006) in connection with three suffixes in Brazilian Portuguese, Panther & Thornburg (2002) in connection with English *-er*, and Radden (2005) in connection with English *-able*. The remaining works on metonymy in word-formation focus on the marginal phenomena of conversion (also known as "zero derivation"; cf. Warren 1999 and Dirven 1999) and compounding (Benczes 2005; Langacker 2009), and all these works examine English.

Traditional works on word-formation are basically lists of affixes (cf. reference grammars such as Švedova 1980; Dokulil 1986; Faarlund et al. 1997). They make no mention of metonymy and the semantic analyses they offer are ad hoc and idiosyncratic, hindering cross-linguistic comparisons. Most other works on word-formation of Russian and Czech follow this pattern, though occasionally with more insight (Šanskij 1968; Townsend 1975; Janda & Townsend 2000; Townsend & Komar 2000; McFadden 1975; Maksimov 1975; Andrews 1996). Theoretical works on word-formation are generally quite abstract (cf. Dokulil 1962; Mel'čuk 1996 and usually offer only a modest set of examples, rather than carrying out the analysis across the system of an entire language. Araeva (2009, 25), in a catalog of hundreds of examples of word-formation from the Kemer dialect of Russian, acknowledges that metonymy plays a role in just three examples. All three of her examples illustrate PART FOR WHOLE/WHOLE FOR PART: *medvežatina* 'bearmeat' derivd from *medved'* 'bear'; *gorošina* 'pea' derived from *gorox* 'peas (collective)'; and *zver'e* 'beasts (collective)' derived from *zver'* 'beast'.

In sum, existing scholarship on metonymy provides little in the way of analysis of word-formation, and the few works that do exist analyze either a single or a handful of affixes, or focus instead on conversion or compounding. These few works are primarily based on English, a language where massive borrowing has compromised the word-formation system. Traditional presentations of word-formation are inventories of affixes, with virtually no mention of metonymy. A systematic study of the role of metonymy in word-formation is lacking.

[2] Peirsman & Geeraerts' (2006) inventory includes one use of metonymy sometimes identified as "grammatical", namely the POTENTIAL FOR ACTUAL metonymy observed in the use of *can* in a phrase like *Can you open the window?*, where the speaker is not interested in the hearer's ability to open windows, but is using this form in place of an imperative. This type of example is perhaps best interpreted as a pragmatic use of a lexical item rather than a grammatical one. At any rate, this use is not as systematic and widespread as the grammatical function of word-formation.

Relating to **Actions**:	ACTION, STATE, CHANGE STATE, EVENT, MANNER, TIME
Relating to **Participants**:	AGENT, PRODUCT, PATIENT, INSTRUMENT
Relating to **Entities**:	ENTITY, ABSTRACTION, CHARACTERISTIC, GROUP, LEADER, MATERIAL, QUANTITY
Relating to **Part for whole**:	PART, WHOLE, CONTAINED, CONTAINER, LOCATED, LOCATION, POSSESSED, POSSESSOR

TABLE 1: Classificatory terms for VEHICLES and TARGETS

[1.2] *Data and methodology*

The goals of the present study have dictated the design of the classification system. There are two goals: 1) to explore the extent of metonymy in word-formation, and 2) to facilitate cross-linguistic comparisons. To this end, it was prudent to devise a classification that would 1) reflect the range of semantic relationships that can be considered metonymic, and 2) be applicable across various languages. Since there is no such previously existing classification for metonymy in word-formation, I chose the best available model: Peirsman & Geeraerts' (2006: henceforth P&G) inventory of metonymy in the lexicon. It was necessary to slightly modify the P&G inventory because it turns out that metonymy is somewhat more diverse in word-formation than in the lexicon.

The P&G inventory of lexical metonymy consists of a series of terms that can serve either as the VEHICLE or as the TARGET in a metonymic relationship. Table 1 presents the terms used in my modified classification for metonymy in word-formation. The terms are organized according to thematic groups. The only term in Table 1 that has no equivalent in the P&G inventory is QUANTITY, which was essential due to the presence of words derived from numerals in all three languages. Otherwise all adjustments that have been made are along the scale of specific to abstract: PARTICIPANT, for example, has been broken down into several more specific terms, and ABSTRACTION has been added as a more abstract version of ENTITY.

A database of TYPES was collected for each of the three languages. A type is a unique combination of three items: 1) a metonymy designation (giving the terms for the VEHICLE and TARGET), 2) a word class designation (giving the word class of both the VEHICLE word and the TARGET word), and 3) an affix. Each type was additionally supplied with an illustrative example. Table 2 on the next page lists some sample entries from the databases.

In order to keep the project manageable and the data comparable across languages, several limitations were imposed. Further restrictions were necessary

Metonymy designation		Word class designation		Suffix	Illustrative example		Language
VEHICLE	TARGET	VEHICLE	TARGET		VEHICLE	TARGET	
PART	WHOLE	noun	noun	*(i)a/jan*	*brjuxo* 'belly'	*brjuxan* 'person with large belly'	Russian
PART	WHOLE	noun	noun	*áč*	*břicho* 'belly'	*břicháč* 'person with large belly'	Czech
CONTAINED	CONTAINER	noun	noun	*nica*	*saxar* 'sugar'	*saxarnica* 'sugar-bowl'	Russian
CONTAINED	CONTAINER	noun	noun	*áč*	*květina* 'flower'	*květináč* 'flower-pot'	Czech
CHARACTERISTIC	ENTITY	qualitative adjective	noun	*áč*	*nahý* 'naked'	*naháč* 'naked person'	Czech
ACTION	LOCATION	verb	noun	∅	*vxodit'* 'enter'	*vxod* 'entrance'	Russian
ACTION	INSTRUMENT	verb	noun	∅ *(fem)*	*péci* 'bake'	*pec* 'oven, stove'	Czech
CHARACTERISTIC	ACTION	qualitative adjective	verb	∅*e*	*søt* 'sweet'	*søte* 'sweeten'	Norwegian
MATERIAL	ENTITY	relational adjective	noun	*ka*	*žestjanoj* 'tin'	*žestjanka* 'tin can'	Russian
ACTION	PRODUCT	verb	noun	*ka*	*sbírat* 'collect'	*sbírka* 'collection'	Czech

TABLE 2: Sample entries in the databases

to remove from consideration data that does not represent metonymic relationships.

This project is limited to suffixal word-formation. This limitation is motivated by two factors: 1) while all three languages also use prefixes to derive words, the majority of word-formation is accomplished via suffixation, 2) the primary purpose of prefixal word-formation in the Slavic languages is for signalling aspect, which is not represented in a commensurate way in a Germanic language such as Norwegian. In regard to the second factor, aspectual suffixes, such as the semelfactive *-nu/-nou* and suffixes used to derive imperfectives were also eliminated from this study. This was done in order to level the playing field so that the databases would represent comparable subsystems across the three languages. This does not, however, mean that metonymy is irrelevant to aspect (cf. Janda 2008; Nesset 2009), just that it was not included in this study.

Although the study reveals that the majority of word-formation is motivated by metonymic relationships, some types of word-formation do not involve metonymy and were thus eliminated from consideration. This includes the formation of hypocoristics (such as Russian *knižka* 'book (dim.)' derived from *kniga* 'book'), formations that merely change the gender (as in Czech *učitelka* 'female teacher' derived from *učitel* 'teacher'), and the formation of deverbal nouns when they lack any specialized meaning (as in Norwegian *maling* in the meaning 'action of painting' from *male* 'paint'; note, however, that *maling* in the meaning 'paint' is metonymic, signalling ACTION FOR MATERIAL).

Any morphological system presents issues of allomorphy, where it is necessary to decide whether a group of items are merely variants of each other or separate units. On issues of allomorphy, I followed the lead of the three reference grammars (Švedova 1980; Dokulil 1986; Faarlund et al. 1997), which are in fairly good agreement. When variants are predictable according to the morphophonemics of the language, they are considered to be allomorphs and thus not separate units. This includes automatic variations due to prosody and phonotactics. Thus, for example, Russian is acknowledged as having only one suffix that could be realized as *-nyj* (as in *mesjačnyj* 'monthly' derived from *mesjac* 'month' via TIME FOR CHARACTERISTIC) or *-noj* (as in *oblastnoj* 'regional' derived from *oblast'* 'region' via LOCATION FOR CHARACTERISTIC) depending on stress. On the other hand, non-automatic variants are treated as separate units, even when an etymological relationship is obvious. Thus Russian *-nica* (as in *saxarnica* 'sugar-bowl') is recognized as distinct from the related *-ica* (as in *teplica* 'hot-house' derived from *teplyj* 'hot' via CHARACTERISTIC FOR LOCATION). Furthermore, all three reference grammars recognize conversion as a type of zero-suffixation (cf. Russian *vxod* 'entrance' derived from *vxodit'* 'enter' via ACTION FOR LOCATION, Czech *pec* 'oven, stove' derived from *péci* 'bake' via ACTION FOR INSTRUMENT, and Norwegian *søte* 'sweeten' derived from *søt* 'sweet' via CHARACTERISTIC FOR ACTION). I do not take a stand on whether

Language	# types	# metonymy designations	# word class designations	# suffixes
Russian	747	110	33	274
Czech	561	105	23	207
Norwegian	177	60	12	57

TABLE 3: Total size of databases in terms of types, metonymy designations, word class designations, and suffixes

zero-morphemes exist, since that issue goes beyond the scope of this study, but I do include examples of conversion in the databases.

Dialectal and colloquial forms are not considered in this study, which also excludes forms that are restricted to a highly marked register (jargon, slang, etc.). Since the aim is to explore systematic types of word-formation, isolated examples are also excluded.

Finally, no kind of frequency information is included in the databases. Each entry consists of a unique type in terms of metonymy designation, word class designation and suffix, and no types are repeated, nor do the databases include any information on type or token frequency. Some types might represent only a couple of derived words, whereas others may represent hundreds of words. And some of the derived words might be relatively rare, whereas others are of high frequency. All of this information is certainly important, but was excluded in this preliminary study, since the goal was to map out the extent of metonymy in word-formation. Frequency can be taken up in future studies.

Table 3 presents the overall dimensions of the databases, which are based primarily on the three reference grammars. These figures cannot be considered absolute metrics given that we are dealing with dynamic open-ended systems and the three grammars may differ in how exhaustive their inventories are. However, the relative sizes are probably indicative of real differences in the three languages. By all measures, the two Slavic languages have much more extensive word-formation systems than Norwegian, but Czech (historically strongly influenced by German) has a somewhat less extensive system than Russian. Over three times as many types are recognized for Czech word-formation than for Norwegian, and Russian yields over four times as many types. This difference is indicative of a different overall strategy between Slavic vs. Germanic languages, where much of the "work" done by word-formation is taken care of by compounding instead (cf. Nesset 2010). In terms of metonymy designations, Russian and Czech are nearly equivalent, and those numbers are nearly double what we find in Norwegian. Overall, the metonymy designations found in word-formation are considerably more diverse than those found in the lexicon. Taken together, there

are 133 different metonymy designations attested across the word-formation systems of the three languages. When we compare these metonymy designations with those found in the lexicon (inventoried in P&G), we find nine designations that are attested only in the lexicon, seventy-nine that are shared by lexicon and word-formation, and fifty-four that are found only in word-formation (for more discussion of this distribution and examples, see Janda forthcoming). Though the phenomenon of metonymy is uniform enough to be classified by the same system across both lexicon and word-formation, it is more diverse in the latter and this increased diversity is largely due to greater combinability of terms.

[2] ANALYSIS

The data in this study can be examined from various different angles, but the purpose of this article is to highlight contrasts. On a number of parameters, the three languages behave very similarly. The purpose of this section is to sort through some of the possible parameters and identify those that yield the most meaningful contrasts.

It is possible, for example, to look at the relationship of metonymy designations to suffixes and the relationship of word class designations to suffixes. It is also possible to probe the metonymy designations in more detail, looking at the relative numbers of VEHICLES vs. TARGETS signalled by a given suffix. Yet another opportunity for comparison is presented by the distribution of bi-directional vs. unidirectional metonymy relationships. In a bi-directional metonymy relationship the same terms can serve as both VEHICLE and TARGET; for example, ACTION FOR AGENT (cf. Russian *tancovščik* 'dancer' derived from *tancevat'* 'dance') and AGENT FOR ACTION (cf. Russian *šoferit'* 'work as a driver' derived from *šofer* 'driver') constitute a bi-directional metonymy relationship. By contrast, PRODUCT FOR AGENT (cf. Russian *lampovščik* 'lamp-maker' derived from *lampa* 'lamp') is a unidirectional metonymy relationship since AGENT FOR PRODUCT is not attested for Russian word-formation. Tautological metonymy relationships such as CHARACTERISTIC FOR CHARACTERISTIC (cf. Russian *veličavyj* 'majestic' derived from *velikij* 'great') constitute a special type of bi-directional relationship.

However, the above-named parameters yield very similar results across the three languages, as summarized in Table 4 on the next page.

The average numbers of metonymy and word class designations per suffix show us that suffixes are not very specific in terms of the metonymies they signal, but they are quite specific as to the word classes they designate. Indeed, the majority of variation associated with word class designations involves the word class of the VEHICLE, not the TARGET. Taken together, these two metrics suggest that we can paraphrase the role of the suffix as follows: take the VEHICLE word and derive a word of word class "X", but the metonymy relation is not usually specified by the suffix. The data in Table 5 on page 252 illustrate this difference in the

Language	Average metonymy designations per suffix	Average word class designations per suffix	Percent of suffixes where # of TARGETS exceeds # of VEHICLES	Percent of bi-directional metonymy relationships
Russian	2.6	1.55	11.0%	43%
Czech	2.6	1.55	12.5%	47%
Norwegian	3.0	1.63	17.5%	45%

TABLE 4: Specificity of suffixes and metonymy designations

specificity of metonymy and word class designations in relation to the Russian suffix *-ina*.

Russian *-ina* can form nouns from adjectives, verbs, and nouns, but the TARGET is always a noun. This suffix is highly non-specific in terms of the metonymy designations it can signal. There are fifteen metonymy designations, with eight different TARGETS. Thus, while Russian *-ina* tells us to form a noun from the VEHICLE word, it does not give us much more information than that. Similar examples are common for both Czech and Norwegian.

The third parameter in Table 4 involves the balance of VEHICLES to TARGETS in metonymy designations. If, as in the case of word class designations, the diversity of metonymy designations were largely a matter of various VEHICLES with a single TARGET, then there would be high determinacy in the system. However, this is not the case. Whereas 60% of suffixes do have a single TARGET, 40% have multiple TARGETS, and from 11% to 17.5% have more TARGETS than VEHICLES. Russian *-ina* shows how non-specific a suffix can be even when the number of TARGETS does not exceed the number of VEHICLES, since *-ina* has eight VEHICLES and eight TARGETS.

The last measure listed in Table 4 involves the balance of bi-directional vs. unidirectional metonymy designations. Once again, the data in Table 5 on the following page can illustrate this phenomenon. We see that the same suffix can even mark both directions of a bi-directional metonymy relation. Russian *-ina* is used to signal both MATERIAL FOR ENTITY in the formation of *l'dina* 'ice-floe' and ENTITY FOR MATERIAL in the formation of *konina* 'horse-meat'. Though it is unusual for a single suffix to signal the opposing directions of a metonymy relationship in this way, it is clearly not impossible, and this is another demonstration of how non-specific suffixes can be in terms of the metonymies they can signal.

However, all of the parameters listed in Table 4 yield very similar dimensions for the three languages. These parameters are valuable for establishing cross-linguistic generalizations, for discovering regularities in how metonymy func-

metonymy designation		word class designation		illustrative example	
VEHICLE	TARGET	VEHICLE	TARGET	VEHICLE	TARGET
CHARACTERISTIC	ABSTRACTION	qual adj	noun	*tixij* 'quiet'	*tišina* 'silence'
CHARACTERISTIC	ENTITY	rel adj	noun	*ženskij* 'female'	*ženščina* 'woman'
ENTITY	ABSTRACTION	noun	noun	*Dostoevskij* 'Dostoevsky'	*dostoevščina* 'Dostoevskian style'
ACTION	PRODUCT	verb	noun	*carapat'* 'scratch'	*carapina* 'scratch'
GROUP	ENTITY	noun	noun	*vinograd* 'grapes'	*vinogradina* 'grape'
GROUP	ABSTRACTION	rel adj	noun	*policejskij* 'police'	*policejščina* 'police repression'
ACTION	EVENT	verb	noun	*krestit'* 'christen'	*krestiny* 'christening'
CHARACTERISTIC	MATERIAL	qual adj	noun	*pušnoj* 'fur-bearing'	*pušnina* 'furs (collect.)'
CHARACTERISTIC	LOCATION	qual adj	noun	*ravnyj* 'equal'	*ravnina* 'plain'
CHARACTERISTIC	ENTITY	qual adj	noun	*rogatyj* 'horned'	*rogatina* 'bear-spear'
CHARACTERISTIC	GROUP	qual adj	noun	*obščij* 'common'	*obščina* 'community'
MATERIAL	ENTITY	noun	noun	*led* 'ice'	*l'dina* 'ice-floe'
ENTITY	MATERIAL	noun	noun	*kon'* 'horse'	*konina* 'horse-meat'
PART	LOCATION	noun	noun	*verx* 'top'	*veršina* 'summit'
PRODUCT	ENTITY	noun	noun	*maslo* 'oil'	*maslina* 'olive-tree'
POSSESSOR	POSSESSED	noun	noun	*gofmejster* 'steward'	*gofmejsterina* 'steward's wife'

TABLE 5: Russian *-ina* as an example of diversity of metonymy and word class designations

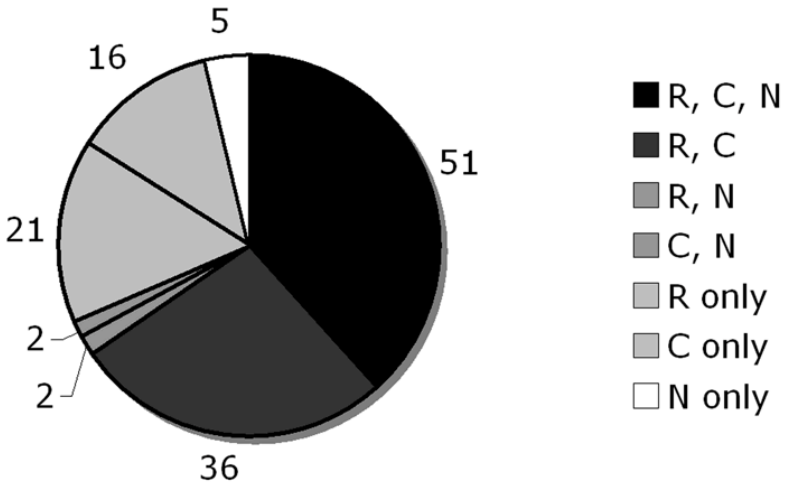


FIGURE 1: Distribution of metonymy designations across the three languages

tions in word-formation. But these parameters do not reveal differences between languages. In order to probe such differences it is necessary to find differences in which metonymies occur in which languages. In the remainder of this section we focus on the following questions: Are there metonymies that occur only in one or two of the three languages? Are there metonymies that are more strongly represented in one or two of the three languages?

Summing across the three languages, a total of 133 metonymy designations are attested in the databases. Figure 1 illustrates the distribution of these metonymies across Russian, Czech, and Norwegian. While fifty-one designations are shared by all three languages, eighty-two are found in only one or two languages. Nearly half of the latter are shared by two languages, with the bulk of these, thirty-six metonymies, attested in Russian and Czech, but absent in Norwegian. The remaining forty-two metonymies are unique to a single language in this group. However, the metonymies that are strongly represented in the languages — represented by several suffixes — still tend to be the same ones. If we look at the metonymies signaled by the most suffixes in each language, the same ten items are among the top fourteen for all three languages. These top ten metonymy designations are listed in Table 6 on the next page.

We can look beyond these similarities to find the metonymies that are proportionally more prominent in one or two languages than in the remainder. Table 7 on page 255 presents such prominent suffixes identified for the two Slavic languages as opposed to Norwegian and for each of the individual languages.

Table 7 begins with a group of metonymy designations that are common in both Russian and Czech, but rare or absent in Norwegian. LOCATION FOR CHAR-

Metonymy designation	Illustrative example		Language
	Vehicle	Target	
ABSTRACTION FOR CHARACTERISTIC	*mysl'* 'thought'	*myslennyj* 'mental'	Russian
ACTION FOR ABSTRACTION	*myslit* 'think'	*myšlenka* 'idea'	Czech
ACTION FOR AGENT	*bake* 'bake'	*baker* 'baker'	Norwegian
ACTION FOR CHARACTERISTIC	*bereč'* 'guard'	*berežnyj* 'careful'	Russian
ACTION FOR INSTRUMENT	*sušit* 'dry'	*sušička* 'dryer'	Czech
ACTION FOR PRODUCT	*stifte* 'establish'	*stiftelse* 'establishment'	Norwegian
CHARACTERISTIC FOR ABSTRACTION	*tixij* 'quiet'	*tišina* 'silence'	Russian
ENTITY FOR CHARACTERISTIC	*Kafka*	*kafkovský* 'Kafkaesque'	Czech
CHARACTERISTIC FOR ENTITY	*tøff* 'tough'	*tøffing* 'tough guy'	Norwegian
ACTION FOR EVENT	*zabastovat'* 'go on strike'	*zabastovka* 'strike'	Russian

TABLE 6: Top ten metonymy designations shared by all three languages

ACTERISTIC is signaled by twenty-two suffixes in Russian and by fourteen suffixes in Czech, but only two suffixes are associated with that metonymy designation in Norwegian. POSSESSOR FOR POSSESSED, signaled by eighteen Russian suffixes and eleven Czech suffixes, is signaled by only one suffix in Norwegian. The remaining metonymy designations in that group are unattested in Norwegian.

The Russian section of Table 7 lists three metonymy designations that are particularly strong in that language. CHARACTERISTIC FOR MATERIAL is associated with nine Russian suffixes, but with only three Czech suffixes and no Norwegian suffixes. The other two designations in this section of Table 7 are exclusive to Russian. These designations suggest that Russian is particularly strong in metonymies that involve CHARACTERISTICS.

Czech excels in deriving nouns via three metonymy relationships that are either unattested or rare in the other two languages. PRODUCT FOR LOCATION is not found in Russian or Norwegian, and CONTAINED FOR CONTAINER is not found in Norwegian; otherwise these three relationships are represented by three or fewer suffixes in the other languages.

		Russian and Czech Illustrative example	
Metonymy designations	# of suffixes	Vehicle	Target
LOCATION FOR CHARACTERISTIC	22 (R), 14 (Cz)	*centr* 'center'	*central'nyj* 'central'
POSSESSOR FOR POSSESSED	18 (R), 11 (Cz)	*kráva* 'cow'	*kraví* 'cow's'
STATE FOR CHARACTERISTIC	12 (R), 10 (Cz)	*želat'* 'want'	*želatel'nyj* 'desirable'
CHARACTERISTIC FOR LOCATION	11 (R), 6 (Cz)	*suxoj* 'dry'	*suša* 'dry land'
PART FOR WHOLE	9 (R), 9 (Cz)	*uši* 'ears'	*ušák* 'bunny'
		Russian Illustrative example	
Metonymy designations	# of suffixes	Vehicle	Target
CHARACTERISTIC FOR MATERIAL	9	*gustoj* 'thick'	*gušča* 'dregs'
INSTRUMENT FOR CHARACTERISTIC	4	*ščipcy* 'tongs'	*ščipcovyj* 'relating to tongs'
CHARACTERISTIC FOR CHARACTERISTIC	4	*velikij* 'great'	*veličavyj* 'stately, majestic'
		Czech Illustrative example	
Metonymy designations	# of suffixes	Vehicle	Target
CONTAINED FOR CONTAINER	11	*písek* 'sand'	*pískoviště* 'sandbox'
PRODUCT FOR LOCATION	6	*mléko* 'milk'	*mlékárna* 'dairy'
QUANTITY FOR ENTITY	6	*sedm* 'seven'	*sedmička* 'number 7 bus, highway, etc.'
		Norwegian Illustrative example	
Metonymy designations	# of suffixes	Vehicle	Target
LOCATION FOR LOCATED	8	*Strømmen*	*strømling* 'person from Strømmen'
PRODUCT FOR AGENT	5	*musikk* 'music'	*musikant* 'musician'

TABLE 7: Language-specific preferences for metonymy designations

The two metonymy designations that are flagged for Norwegian are attested robustly in both Russian and Czech, but are ranked relatively higher (eighth and eleventh most common) in Norwegian. LOCATION FOR LOCATED, though it can identify objects in addition to people in both Russian and Czech, is specialized only to human targets in Norwegian.

It is tempting to speculate on possible cultural parallels to language-specific patterns. In addition to the bias toward CHARACTERISTICS noted above for Russian, Czech appears to be very focused on quantification and commercial transactions. Norwegian preference for LOCATION FOR LOCATED seems to correspond to a strong sense of the connection between location and personal identity in Norway. Future studies could show whether there are indeed cultural parallels to metonymy preferences among languages.

[3] CONCLUSIONS

This article opens up a new parameter for comparing languages, by means of comparing their word-formation systems. Word-formation systems tend to be unwieldy and to appear intractable for cross-linguistic comparison. However, when word-formation is understood in terms of metonymy designations the comparisons can become both feasible and meaningful. This article reports on a pilot study of three languages, Russian, Czech, and Norwegian, using a classification system based on the inventory of lexical metonymic relations known to linguistic scholarship. Surprisingly, we discover that the diversity of metonymy that underlies word-formation is even more extensive than the semantic shifts that motivate lexical metonymy. Some tendencies turn out to be fairly uniform across the three languages, such as the ratio of word classes to suffixes, the degree of metonymic target specificity for suffixes, and the balance of bi-directional vs. unidirectional metonymy relations. There is a top ten list of metonymies that are equally robust in all three languages, but further comparison reveals that some metonymies are proportionally stronger in some languages than in others. Thus despite the many similarities, there are also differences in what metonymic relations different languages invest in. More research needs to be done on the word-formational systems of a greater variety of languages in order to verify and fine-tune the classification system. This line of inquiry has the potential to reveal patterns of semantic association that may have important cultural parallels as well.

ACKNOWLEDGEMENTS

This research has been supported by grants from Norges forskningsråd and Sparebank1.

REFERENCES

Andrews, Edna. 1996. *The semantics of suffixation: agentive substantival suffixes in Modern Standard Russian* (LINCOM Studies in Slavic Linguistis 5). Munich: LINCOM Europa.

Araeva, L. A. 2009. *Slovoobrazovatel'nyj tip*. Moscow.

Basilio, Margarida. 2006. Metaphor and metonymy in word-formation. *DELTA: Revista de Documentação de Estudos em Lingüïstica Teórica e Aplicada* 22. 67–80.

Benczes, Réka. 2005. Creative noun–noun compounds. *Annual Review of Cognitive Linguistics* 3. 250–268.

Dirven, René. 1999. Conversion as a conceptual metonymt of event schemata. In Klaus-Uwe Panther & Günter Radden (eds.), *Metonymy in language and thought* (Human Cognitive Processing 4), 275–287. Amsterdam: John Benjamins.

Dokulil, Miloš. 1962. *Tvoření slov v češtině. 1: Teorie odvozování slov*. Praha: ČSAV.

Dokulil, Miloš. 1986. *Mluvnice češtiny (1)*. Praha: Academia.

Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.

Janda, Laura A. 2008. Metonymy via perfectivization of Russian verbs. In *S ljubov'ju k slovu. Festschrift in honour of Professor Arto Mustajoki on the occasion of his 60th birthday* (Slavica Helsingiensia 35), 77–85.

Janda, Laura A. forthcoming. Metonymy in word-formation. Forthcoming MS.

Janda, Laura A. & Charles E. Townsend. 2000. *Czech* (Languages of the World/Materials 125). Munich, Newcastle: LINCOM Europa.

Koch, Peter. 1999. On the cognitive bases of metonymy and certain types of word formation. In Klaus-Uwe Panther & Günter Radden (eds.), *Metonymy in language and thought* (Human Cognitive Processing 4), 139–167. Amsterdam: John Benjamins.

Kövecses, Zoltán & Günter Radden. 1998. Metonymy: developing a cognitive linguistic view. *Cognitive Linguistics* 9. 37–77.

Langacker, Ronald W. 1993. Reference-point constructions. *Cognitive Linguistics* 4. 1–38.

Langacker, Ronald W. 2009. Metonymic grammar. In Klaus-Uwe Panther, Linda L. Thornburg & Antonio Barcelona (eds.), *Metonymy and metaphor in grammar* (Human Cognitive Processing 25), 45–71. Amsterdam: John Benjamins.

Maksimov, V. I. 1975. *Suffiksal'noe slovoobrazovanie imen suščestvitel'nyx v russkom jazyke*. Leningrad: Izdatel'stvo LGU.

McFadden, Kenneth Dallas. 1975. *The morphophonemics of adjectival derivation in Russian*: University of Michigan dissertation.

Mel'čuk, I. A. 1996. Lexical functions: a tool for the description of lexical relations in a lexicon. In Leo Wanner (ed.), *Lexical functions in lexicography and natural language processing* (Studies in Language Companion Series 31), 37–102. Amsterdam: John Benjamins.

Nesset, Tore. 2009. Metonymy of aspect/aspects of metonymy. *Scando-Slavica* 55. 65–77.

Nesset, Tore. 2010. The art of being negative: metonymical morphological constructions in contrast. This volume.

Padučeva, E. V. 2004. *Dinamičeskie modeli v semantike leksiki*. Moscow: Jazyki slavjanskoj kul'tury.

Panther, Klaus-Uwe & Linda L. Thornburg. 1999. The potentiality for actuality metonymy in English and Hungarian. In Klaus-Uwe Panther & Günter Radden (eds.), *Metonymy in language and thought* (Human Cognitive Processing 4), 337–357. Amsterdam: John Benjamins.

Panther, Klaus-Uwe & Linda L. Thornburg. 2002. The roles of metaphor and metonymy in English *-er* nominals. In René Dirven & Ralf Pörings (eds.), *Metaphor and metonymy in comparison and contrast* (Cognitive Linguistics Research 20), 279–319. Berlin: Mouton de Gruyter.

Panther, Klaus-Uwe & Linda L. Thornburg. 2007. Metonymy. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 236–263. Oxford University Press: Oxford.

Peirsman, Yves & Dirk Geeraerts. 2006. Metonymy as a prototypical category. *Cognitive Linguistics* 17. 269–316.

Radden, Günter. 2005. The ubiquity of metonymy. In José Luis Otal Campo, Ignasi Navarro i Ferrando & Bellés Fortuño Begoña (eds.), *Cognitive and discourse approaches to metaphor and metonymy* (Estudis filològics 19), 11–28. Castelló de la Plana: Publicacions de la Universitat Jaume I.

Radden, Günter & Zoltán Kövecses. 1999. Towards a theory of metonymy. In Klaus-Uwe Panther & Günter Radden (eds.), *Metonymy in language and thought* (Human Cognitive Processing 4), 17–59. Amsterdam: John Benjamins.

Šanskij, N. M. 1968. *Očerki po russkomu slovoobrazovaniju*. Moscow: Izdatel'stvo MGU.

Švedova, N. Ju. (ed.). 1980. *Russkaja grammatika*. Moscow: Nauka.

Townsend, Charles E. 1975. *Russian word-formation*. Columbus, OH: Slavica Publisher.

Townsend, Charles E. & Eric S. Komar. 2000. *Czech through Russian*. Bloomington, IN: Slavica Publishers.

Warren, Beatrice. 1999. Aspects of refernetial metonymy. In Klaus-Uwe Panther & Günter Radden (eds.), *Metonymy in language and thought* (Human Cognitive Processing 4), 121–138. Amsterdam: John Benjamins.

AUTHOR CONTACT INFORMATION

Laura A. Janda
University of Tromsø
Department of Language and Linguistics
NO-9037 Tromsø
Norway
laura.janda@uit.no

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 261–279. (ISSN 1890-9639)
http://www.journals.uio.no/osla

THE ART OF BEING NEGATIVE: METONYMICAL MORPHOLOGICAL CONSTRUCTIONS IN CONTRAST

TORE NESSET
University of Tromsø

ABSTRACT

This paper explores the metonymic nature of grammar by contrasting a Russian and a Norwegian morphological construction that give (mostly) negative characterizations of people. The meanings of the constructions are strikingly similar, and it is argued that they cannot be properly understood without recourse to metonymy. However, while Norwegian employs compounding, the morphological strategy used in the Russian construction is derivation. It is demonstrated that these differences are not idiosyncratic facts about Norwegian and Russian, since there is systemic motivation for the differences between the two languages.

[1] INTRODUCTION

Consider the following example from Ljudmila Ulickaja's short novel (повесть) *Sonečka* which contains the common gender noun *зануда:*[1]

(1) Таня назвала его занудой.
'Tanja called him a bore.'

In the Norwegian translation by Marit Bjerkeng, the word for 'bore' is rendered as the compound *tørrpinn* (lit. 'dry stick'):

(2) Tanja kalte ham en tørrpinn
'Tanja called him a bore.'

The observation that forms the starting point for the present study is that the translation of a Russian common gender noun in *-а* by means of a compound in Norwegian is not a coincidence. In fact, there are numerous word pairs like *зануда — tørrpinn*, some of which are given in Table 1 on the following page. All the words in the table are negative characterizations of people, but they are morphologically different. The Russian words are examples of derivational morphology,

[1] Examples (1) and (2) are from the RuN corpus available at http://www.hf.uio.no/tekstlab/.

Russian	Norwegian	English gloss
жадоба, жад(н)юга	gjerrigknark	skinflint
плакса	tutemikkel	crybaby
сластёна	slikk(e)munn	person with sweet tooth
зубрила	pugghest	rote-learner
трусишка	reddhare	scaredy-cat
зануда	tørrpinn	bore, tedious person
злюка	sinnatagg	cross-patch

TABLE 1: Examples of the Russian Characteristic Derivation Construction and the Norwegian Characteristic Compound Construction

while the Norwegians are compounds. Accordingly, I will talk about two morphological constructions: the "Russian Characteristic Derivation Construction" and the "Norwegian Characteristic Compound Construction". I use the term "construction" about conventionalized pairings of form and function (Goldberg 2006, 3; see also Goldberg 1995 as well as Booij 2005 and 2009, who relates the term to morphology).

In my contrastive analysis of the two constructions, I will focus on the similarity of meaning, as well as the difference in form. In sections [2] through [4], it will be argued that metonymy is pervasive in the meaning of the constructions. Section [5] investigates the negative aspect of the meaning of the constructions, arguing that it arises as the result of the interplay of quantity and quality. In section [6], we turn to form. I propose that the fact that Russian uses a derivational pattern receives systemic motivation from the presence of hypocoristic word-formation and the role of the *a*-declension as a repository for marked persons in Russian. The contribution of the article is summed up in section [7].

[2] METONYMY

In the present study, I will use the term "metonymy" as it is used in contemporary cognitive linguistics:

(3) Metonymy is a cognitive process in which one conceptual entity, the vehicle, provides mental access to another conceptual entity, the target, within the same cognitive model (Radden & Kövecses 1999, 21, see also Padučeva 2004, 157f., and Peirsman & Geeraerts 2006).

Two aspects of this definition deserve mention. First, metonymy is not defined as a trope that is used to embellish poetry, nor is it described as a meaning shift occurring when one word is substituted for another. Rather, metonymy is first

and foremost understood as a cognitive phenomenon that informs the way we think, act and speak. In Lakoff & Johnson's (1980, 35) example *the ham sandwich is waiting for his check*, *the ham sandwich* is used to refer to a person. This person (*the target*) is accessed through the sandwich he has ordered (*the vehicle*), which is a salient property in the restaurant setting. The pervasiveness of examples like this in everyday speech suggests that the ability to perform metonymies is a fundamental property of human cognition.

The second part of the definition that calls for comment is the notion *cognitive model*. Alternative notions used in the definition of metonymy include *domain* (Kövecses 2002) and *domain matrix* Croft (2002). Peirsman & Geeraerts (2006) criticize these notions as vague, and suggest instead a more traditional definition in terms of contiguity. Interesting as this discussion is, it does not have consequences for the present study, and it is therefore not necessary to go into further detail. Under all these approaches, the vehicle and target are closely related in space and/or time. For instance, in the ham sandwich example the sandwich and the person who ordered it are in the same restaurant at the same time. Without this close relationship, the metonymy could not arise.[2]

In recent years, metonymy has attracted considerable attention in cognitive linguistics. Benczes (2006) has emphasized the relevance of metonymy for compounding, and Janda (this volume) explores metonymy in derivational morphology. Langacker (2009, 46) goes as far as to say that "[g]rammar is basically metonymic". This claim will receive further support in the present paper, insofar as the morphological constructions under scrutiny cannot be understood without recourse to metonymy.

[3] A GENERAL SCHEMA FOR THE MORPHOLOGICAL CONSTRUCTIONS

In order to investigate the Russian Characteristic Derivation Construction and the Norwegian Characteristic Compound Construction I set up a database for each of them. The Russian database contains 369 nouns, which are all the nouns of common gender attested in Zaliznjak (1977). Common gender indicates that the nouns in question combine with masculine or feminine agreement targets depending on the biological sex of the referent. The Norwegian database, which comprises 523 compounds, was compiled from the electronic versions of *Norsk riksmålsordbok* (www.ordnett.no), and *Bokmålsordboka* and *Nynorskordboka* (www.dokpro.uio.no). Since dictionaries are often somewhat conservative when it comes to e. g. taboo

[2] Notice that once a metonymy has been created it can be used outside the setting where it arose. For instance, if a regular customer always order a ham sandwich at a certain restaurant, a waiter might well refer to the customer metonymically as the ham sandwich, even if s/he runs into the customer at a different location.

words, a few words attested in Google searches are included in the database.[3]

The analysis of the two databases shows that the meaning of both constructions can be represented as follows:

(4) The nouns under scrutiny denote a PERSON who has a PROPERTY that is CHARACTERISTIC of the person and (often) NEGATIVE.

Let us return to the word pair *зануда — tørrpinn* 'bore' mentioned in the beginning of this study. Both words denote a person who is the carrier of the property of being boring, so in this sense they are arguments to the predicate 'be boring'. Being boring, furthermore, is characteristic of the person in question, and being boring is generally considered a bad thing. It should be pointed out that the description in (4) does not fit all the words in the database equally well. Although the vast majority of Russian common gender nouns denote persons, there are a few animals in the class (e.g. *сивка* 'dark gray horse', *серка* 'gray horse (or cat or dog)'). Among the nouns denoting persons there is a small group of neutral terms (both Church Slavic words like *предтеча* 'forerunner' and more recent borrowings like *коллега* 'colleague'). Moreover, some words involve positive evaluation (e.g. *умница* 'smart person'). However, the majority of the personal nouns involve negative characteristics. Since — as we will see in section [5] — negative evaluation is a matter of degree, it is not possible draw a clear-cut demarcation line between negative and non-negative nouns. A case in point is Norwegian *driblefant* '(excessive) dribbler'. There is nothing wrong with dribbling per se, but in excessive quantities it becomes a problem. The word is attested in clearly derogatory contexts, such as in a song by the Norwegian rock singer Åge Aleksandersen, where *driblefant* is used together with the clearly derogatory *rotsekk* 'slob' to characterize a bad soccer player. However, in recent newspaper prose the same word is sometimes used in a more neutral way to characterize technically skilled soccer players. In the following sentence, for instance, the emphasis is on the high market value of the player in question:

(5) Lyns driblefant har tiltrukket seg massiv oppmerksomhet fra interesserte klubber, men hittil har kun Brann avslørt sin interesse.
(*Dagbladet*, August 19 2009)
'Lyn's [name of Norwegian football club] dribbler has attracted massive attention from interested clubs, but so far only Brann [name of Norwegian football club] has made its interest public.'

[3] Since there is no grammatical marker that unites the class of compounds of interest in the present study, the Norwegian database could not be compiled fully automatically. A number of relevant second components (nouns denoting people, animals or body parts) were identified manually. Then electronic searches were performed in the dictionaries and all compounds ending in these nouns were identified. The compounds with relevant meaning were included in the database. I would like to express my gratitude to Maria Nordrum and Anna Baydimirova for their assistance with the databases.

For the purposes of the present study, words like *driblefant* were included in the database, although they are not used exclusively as derogatory words.

In the beginning of this article, I emphasized the relevance of metonymy for the "Russian Characteristic Derivation Construction" and the "Norwegian Characteristic Compound Construction". We are now in a position to see why. In both constructions we access a person through a salient characteristic, such as tediousness or his/her inclination to perform excessive dribbling. This is a classic example of metonymy, which can be represented as CHARACTERISTIC FOR PERSON. In the terminology of Radden & Kövecses' definition in (3), the characteristic property is the vehicle and the person the target. Both *зануда — tørrpinn* are related to adjectives (*занудный* 'tiresome' and *tørr* 'dry'), but there are also words based on verbs and nouns in the databases. Examples of verb-based words are *driblefant* 'excessive dribbler' (cf. *drible* '(to) dribble') and *торопыга* 'person always in a hurry' (cf. *торопиться* '(to) hurry'). Noun-based words include *løgnhals* 'liar' (lit. 'lie-throat') and *сластёна* 'person with sweet tooth'. In the case of noun-based words, we are arguably dealing with double metonymy. In *løgnhals*, for instance, the noun *løgn* 'lie' metonymically stands for the predicate of telling lies (RESULT FOR PROCESS), which in turn stands for the agent of this predicate, i. e. the teller of lies (AGENT FOR PROCESS). Similarly, in *сластёна* there is arguably a chain of metonymies from the object sweets to eating sweets, and then to the eater of sweets.[4] What these examples show is that metonymy facilitates a precise description of the meaning of the constructions in Russian and Norwegian. In the following section, we will consider additional evidence for the pervasiveness of metonymy.

[4] MORE METONYMY: NORWEGIAN COMPOUNDS

In the Russian Characteristic Derivation Construction the person in question is represented in a highly schematic way. In words like *торопыга* 'person always in a hurry', the derivational suffix *-iga* indicates that we are dealing with a person who is negatively evaluated, but beyond that the suffix does not give us any information about the person in question. Cases like *зануда* 'boring person' are even less informative, since they do not involve an overt suffix at all. Here, language users have to have knowledge about the construction in order to be able to use and interpret the word correctly; if you know that Russian has a construction of deadjectival nouns in the *-а* declension denoting persons, you may be able to tackle words like *зануда* even if the information you have is not sufficient to predict the exact meaning and use of the words in question.

The Norwegian Characteristic Compound Construction is in a sense more informative, insofar as compounds contribute two lexical roots. However, although

[4] From a synchronic point of view, I assume that *сластёна* is more closely related to the noun *сласти* 'sweets' than to the adjective *сладкий* 'sweet'.

the second component of the compounds involves relevant information for the interpretation of the words, the interpretation is not straightforward, insofar as it requires cognitive processes such as metonymy. The compounds in my database can be divided into five types according to the meaning of the second component:

(6) Types of second components in Norwegian compounds:
   a. Person: *klossmajor* 'clumsy person' (lit. 'clumsy major'), *driblefant* 'excessive dribbler' (lit. 'dribble-hobo')
   b. Body part: *kjøtthue* 'meathead', *løgnhals* 'liar' (lit. 'lie-throat')
   c. Animal: *pugghest* 'rote-learner' (lit. 'rote-learning horse'), *stabukk* 'stubborn old mule' (lit. 'stubborn billy-goat')
   d. Object: *tørrpinn* 'bore' (lit. 'dry stick'), *skravlebøtte* 'chatterbox' (lit. 'chatterbucket')
   e. Opaque: *stabeis* 'stubborn person' (*beis* is attested in the meaning 'paint', but that seems irrelevant for the synchronic analysis of *stabeis*).

In the following, we shall take a closer look at types (6a) and (6b), which shed light on one of the main topics of this study, namely metonymy. In compounds like *klossmajor* 'clumsy person' and *driblefant* 'excessive dribbler', the second component denotes a person with specific characteristics. The noun *major* 'major' describes a person with a military rank between captain and lieutenant-colonel. If the compound *klossmajor* were fully compositional, one would therefore expect the meaning to be 'clumsy major'. However, it is not. You do not have to graduate from a military academy in order to be a *klossmajor* — the compound denotes any clumsy person. The word *fant* 'hobo' denotes a person who belongs to a particular social group. However, you don't have to be a hobo in order to classify as a *driblefant.* Any person, regardless of his/her social background can be a *driblefant* as long as s/he dribbles a lot. In fact, since hobos generally are not involved in soccer and similar sports, it would be extremely unlikely that one would be able to find a single instance of a hobo who could be characterized as a *driblefant*. What we see in cases like *klossmajor* and *driblefant*, is that a specific person (an army officer or a hobo) stands for a person in general. Since *major* and *hobo* are hyponyms of *person*, this can be analyzed as the metonymy HYPONYM FOR HYPERNYM (Peirsman & Geeraerts 2006, 277 and 306–408, see also Dirven 1998, 284 for discussion of similar metonymies in noun-verb conversion). In order to interpret Norwegian compounds where the second component denotes a person, language users have to perform a metonymy. If we imagine a language user who for some reason cannot handle metonymies, s/he would be at loss when facing compounds like *klossmajor* and *driblefant*.

Compounds where the second component is a body part also presuppose metonymy. If the meaning of *kjøtthue* 'meathead' and *løgnhals* 'liar' were fully com-

positional, these words would denote a head and a throat, respectively, since *hue* means 'head' and *hals* 'throat'. However, the words denote people, so in order to interpret such compounds language users have to perform a metonymy from a body part to the person with this body parts. In other words, a body part stands for a person. This is an example of the PART FOR WHOLE metonymy — one of the most studied examples of metonymy, known as *pars pro toto* in classical rhetoric. Peirsman & Geeraerts (2006, 309) regard part-whole relations as the prototypical type of metonymy.

At this point the question arises as to how speakers are able to identify the meaning of compounds like *kjøtthue* and *løgnhals*. How do language users know that they have to perform a metonymy? Why don't they misunderstand? Why don't people think that these words denote body parts? I have no pretensions of giving a complete answer to these questions, but the facts about the Norwegian Characteristic Compound Construction shed light on three of the factors involved: context, entrenchment and blocking. Words occur in a context, which often gives the language user hints as to whether a metonymy should be performed or not. If I say *Jon er et kjøtthue* 'John is a meathead', the context prompts a metonymical reading since *Jon* is most likely the name of a person, and not of a head. However, context is not always enough. If you hear me say *For et kjøtthue!* 'What a meat-head!' the (immediate) context does not necessarily exclude reference to a head with a surprisingly large amount of meat on it.

Another observation is that not any body part will do in the construction. The distribution of body parts is clearly skewed in my data. As shown in Table 1 on page 262 and Figure 1 on page 269, there is a strong tendency for the nouns in question to involve the head or some part of it, such as the mouth. Taken together, the head and the throat (which is arguably part of the head) account for 77% of the body part compounds in my database. I suggest that language users can utilize this fact. If s/he encounters a compound involving the head, s/he can with some degree of confidence, at least, infer that the compound is used metonymically for the person as a whole. If, on the other hand, a language user comes across an unfamiliar compound involving, say, the heart, s/he would have a much weaker basis for performing the metonymy from body part to person. The Norwegian word for 'heart', *hjerte*, forms numerous compounds, but these compounds do not stand metonymically for persons. A case in point is *vennehjerte* 'friend's heart', which according to *Norsk Riksmålsordbok* can be used metaphorically to refer to a friend's warm feelings, but cannot be used metonymically to refer to the person who has these feelings. Another example is *sovehjerte* (lit. 'sleep-heart'), which denotes the ability to fall asleep easily, not the person who has this ability.

Body part	Raw frequency	Per cent	Examples
Head/part of head	37	60.7	*kjøtthue* 'meathead', *slikk(e)munn* 'person with sweet tooth' (lit. 'lick-mouth')
Throat	10	16.4	*løgnhals* 'liar' (lit. 'lie-throat'), *skrikhals* 'cry-baby' (lit. 'scream-throat')
Genitals	9	14.8	*hestkuk* 'asshole' (lit. 'horse-penis'), *dovenpeis* 'lazybones' (lit. 'lazy penis')[5]
Torso	2	3.3	*dovenkropp* 'lazybones' (lit. 'lazy body'), *urokropp* 'live wire' (lit. 'unrest-body')
Fur	2	3.3	*svinepels* 'bastard' (lit. 'swine-fur'), *revepels* 'smarty pants' (lit. 'fox-fur')
Hand	1	1.6	*treneve* 'clumsy person' (lit. 'wood-fist')
Total	61	100.1	

TABLE 2: Distribution of body parts in the Norwegian Characteristic Compound Construction

I suggest that the effect described in the previous paragraph is an "entrenchment effect". The head is more frequently used in the construction, and therefore represents a more entrenched pattern in the terminology of Langacker (1987, 59 and 2008, 16–17). The idea that language users utilize entrenchment for the interpretation of compounds is simple, but nevertheless has far-reaching implications for linguistic theory. The entrenchment effect suggests that the language faculty is not, or at least not only, an algorithm for manipulation of abstract symbols, but is sensitive to (type) frequency (for more about the relationship between frequency and linguistic structure, see Bybee 2001, 2007).

In addition to the context and entrenchment effects, I suggest that there is a blocking effect that helps the language users perform metonymies. Some of

[5] Compounds with *peis* as the second component are included in the category "genitals" since *Norsk Riksmålsordbok* defines *peis* as the "penis of an animal, especially a bull" (My translation, TN). However, apart from compounds, this word is not much used in contemporary Norwegian, and many speakers are probably not aware of its meaning. For such speakers, compounds in *peis* are of the opaque type in (6e).

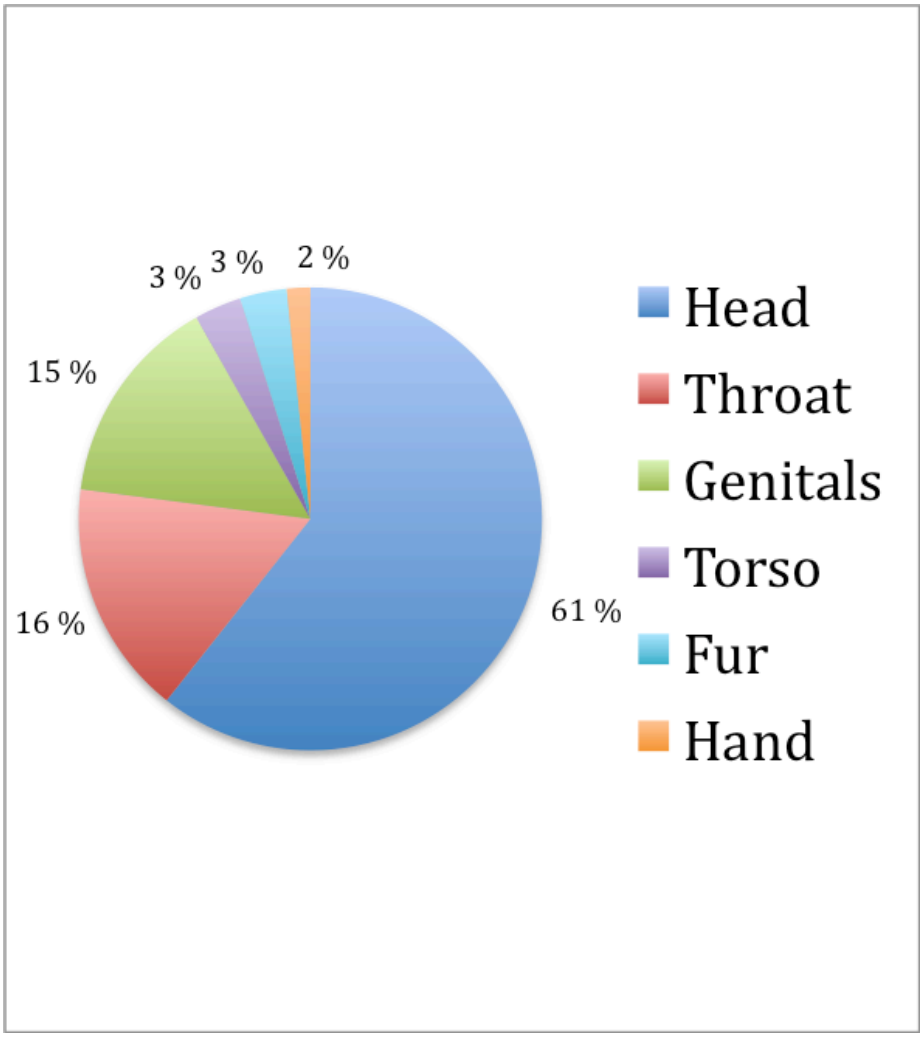


FIGURE 1: Distribution of body parts in the Norwegian Characteristic Compound Construction

the compounds in the database can be interpreted literally. As mentioned, *kjøt-thue* 'meathead' might potentially be used to designate a head with a surprisingly large amount of meat on it. However, world knowledge tells us that this is not a very likely interpretation, since most of us do not encounter situations where the amount of meat on heads is so salient that we need a separate lexical item to refer to it.[6] A word where a literal interpretation is more likely is *revepels*. This word can be used in the meaning 'smarty pants' (lit. 'fox-fur'), but it is equally natural to use the word to refer to the animal's fur or a coat made from the fur. In cases like *slikk(e)munn* 'person with a sweet tooth' (lit. 'lick-mouth'), however, a literal

[6] Notice that the (ir)relevance of the amount of meat on heads is culture specific. In cultures where animal heads are prepared for food, the amount of meat on a head may be crucial. An example that comes to mind is western Norway where the dish *smalahove* is prepared from the head of a sheep.

interpretation is much less likely. World knowledge indicates that mouths cannot lick, at least not independently of the person who "owns" the mouth. Admittedly, one could perhaps conceive of using *slikk(e)munn* about a contraption with some similarity to a mouth, but this would be quite far-fetched. What I suggest is that when a literal interpretation is blocked, the language user is encouraged to perform a metonymy. A thorough discussion of blocking in morphology is beyond the scope of the present article, so the interested reader is referred to Carstairs-McCarthy (1992, 33f.), Haspelmath (2002, 108f. and 249f.) and Aronoff & Fudeman (2005, 218f.) and references therein. Notice, however, that the blocking effect is not an either-or matter, since a literal interpretation can be more or less likely. Once again, we are not dealing with an algorithmic system of strict rules. Rather, the human mind seems to be sensitive to various degrees of likelihood of a literal interpretation in the same way as it is sensitive to a pattern's degree of entrenchment.

In addition to illustrating the context, entrenchment and blocking effects, the data in Table 2 on page 268 and Figure 1 on the preceding page raise an important question: Why is the head so prevalent in the Norwegian Characteristic Compound Construction? Several factors deserve mention. First, in folk models the head is considered the home of the rational self, and it is therefore important in defining the personality, including negative characteristics. Second, the head is crucial for biological functions such as eating and social functions such as speech. Finally, as opposed to e.g. the heart, which is important in folk models as the home of emotions, the head is visible, and it is possible that visibility enhances the use of the head in metonymy. The common denominator for the three factors is salience — the head is salient in relevant folk models, biological and social functions and with regard to visibility. Salience is often regarded as a relevant factor in metonymy (see e.g. the discussion of reference point constructions in Langacker 2008, 83). This study of the Norwegian body part compounds lends support to this idea.

As shown in Table 2 and Figure 1, genitals (male and female) are relatively frequent in the Norwegian Characteristic Compound Construction. In my database, they represent 15% of the body part compounds. Once again, salience seems to be a key word, insofar as the genitals distinguish between males and females. Needless to say, natural gender is a salient property biologically and socially. In addition to this, however, many words for genitals are taboo words, which makes them particularly well suited in derogatory use. This takes us to an important question: what is negative? This is the topic of the next section.

[5] WHAT IS NEGATIVE?

As a first approximation to this question, we may look at (7) and (8), where I list a number of Norwegian and Russian words classified into various categories.

(7) What is "bad" in Norwegian?

a. Immoral behavior: drinking (*fyllefant* 'drunkard'), lying (*ljugarfant* 'liar'), stealing (*tjuvradd* 'petty thief'), laziness (*lathans* 'lazybones'), adultery (*horebukk* 'lecher')
b. Annoying behavior: talking (*skravlebøtte* 'chatterbox'), teasing (*ertekrok* 'teaser')
c. Pointless, idle activity: *pugghest* 'rote-learner', *driblefant* 'excessive dribbler'
d. Physical characteristics: *brillejesus* 'four-eyes', *halteper* 'person who limps'
e. Marginal social groups: *ferdafant* 'tramp', *byfant* 'derogatory word for city dweller', *bondeknøl* 'country bumpkin', *fattiglus* 'pauper', *svenskeradd* 'derogatory word for Swede'
f. Stupidity/clumsiness: *dustemikkel* 'fool', *dummepetter* 'fool', *klossmajor* 'clumsy person'
g. Mentality: anger (*sinnatagg* 'cross-patch'), melancholy (*tåreperse* 'cry-baby'), being difficult (*vriompeis* 'pighead'), dullness (*tørrpinn* 'bore')
h. Cowardice: *reddhare* 'scaredy-cat', *knehøne* 'coward'
i. Showing off: *spradebasse* 'peacock, dandy', *flottenfeier* 'poser'
j. Luck, smartness: *heldiggris* 'lucky dog'

(8) What is "bad" in Russian?

a. Immoral behavior: drinking (*пьяница* 'drunkard', *гуляка* 'boozer, idler'), lying (*враша* 'liar', *врунишка* 'liar'), stealing (*ворюга* 'petty thief'), laziness (*лежебока* 'lazybones'), murder (*убийца* 'murderer')
b. Annoying behavior: talking (*балаболка* 'chatterbox', *болтунишка* 'chatterbox'), teasing (*задира* 'teaser')
c. Pointless, idle activity: *зубрила* 'rote-learner', *мазила* 'dauber', *писака* 'scribbler'
d. Physical characteristics: *калека* 'cripple', *коротконожка* 'short-legged person'
e. Marginal social groups: *бродяга* 'tramp', *бедняга* 'poor fellow', *немчура* 'derogatory word for German'
f. Stupidity/clumsiness: *дурында* 'fool', *недотёпа* 'duffer', *мямля* 'mumbler'
g. Mentality: anger (*злюка* 'cross-patch'), melancholy (*плакса* 'crybaby', *нюня* 'crybaby'), being difficult (*капризнюля* 'capricious child'), dullness (*зануда* 'bore')
h. Cowardice: *трусишка* 'coward'
i. Showing off: *задавака* 'show off', *ломака* 'poser'
j. Luck, smartness: *хитрюга* 'sly, cunning person', *пройда* 'creeper'

This classification is not exhaustive, but (7) and (8) suffice to illustrate two important points. First of all, the same categories recur in both Norwegian and Russian. The Norwegian Characteristic Compound Construction and the Russian Characteristic Derivation Construction therefore do not seem to provide fertile soil for linguists looking for national stereotypes. The second point is that there is no small and clearly delineated set of negative properties that occur in the constructions under scrutiny in this study. On the contrary, almost anything can be considered negative in one way or other. A word like *хитрюга* 'sly, cunning person' is a good example. While being smart or resourceful is a good thing *per se*, as suggested by the gloss, the word *хитрюга* is often used with negative connotations about people who are willing to manipulate others in order to achieve their goals:

(9) Увидев, что Лева запросто перетаскивает пудовые узлы, хитрюга бортмеханик сначала восхитился его силой, потом попросил передвинуть полтонны груза и спокойно смотрел, как Лева в одиночку делает его, бортмеханика, работу (V. Sanin 1987)[7]
'Having seen that Leva was simply dragging packs of one pood [=16.3 kilos] in weight, the sly and cunning flight engineer admired his strength, and then asked him to move half a ton of cargo while calmly looking at Leva doing the flight engineer's job alone.'

For the purposes of the constructions we are interested in, things are not bad in themselves. Whether something is negative or not, depends on how a situation is construed. Construal, which has to do with the human capacity for adopting different perspectives, may be defined as the relationship between a conceptualizer and the conceptualization s/he entertains (Langacker 1987, 128, 1991, 546, 2008, 55ff.). According to Verhagen (2007, 48), construal involves "facets of meaning and grammatical organization which crucially make use of notions such as 'perspective', 'subjectivity', or 'point of view'. What these notions have in common is that they capture aspects of conceptualization that cannot be sufficiently analyzed in terms of properties of the object of conceptualization, but [...] necessarily involve a subject of conceptualization." Clearly, we cannot understand the words in (7) and (8) without taking construal into consideration.

What are the factors that underlie the construal of negative characteristics in the Norwegian and Russian morphological constructions? I suggest that both quality and quantity are important. As for quality, the overviews in (7) and (8) indicate that the adopted perspective is always that of a majority against outliers. If a person is characterized as *fyllefant* 'drunkard' or the corresponding Russian word *пьяница*, s/he shows behavior that is unacceptable from the perspective of the majority's norms. Exactly what is construed as sins and vices presumably

[7] Example from The Russian National Corpus available at http://www.ruscorpora.ru.

varies among cultures and languages. As mentioned, the lists in (7) and (8) are remarkably similar, but since other cultures have other stereotypes, it is likely that morphological constructions in other languages reflect different construals. In any case, it is clear that the meaning of morphological constructions is deeply intertwined with cultural stereotypes. As Enfield (2002, 3) remarks, "[g]rammar is thick with cultural meaning".

Quantity is also a relevant factor. In section [3], I mentioned that dribbling in *driblefant* is not negative *per se*, but becomes problematic in excessive quantities. Similarly, words like *skravlebøtte* 'chatterbox' and its Russian equivalents *балаболка* and *болтунишка* involve excessive talking. Talking is not a bad thing in itself, but in large quantities it can become annoying, especially if there is not much content. These examples suggest that excessive quantity compensates for the lack of graveness of the relevant sin or vice. But there are also examples where it is sufficient to commit the sin only once and still be a full-fledged member of the category. Nouns like *убийца* 'murderer' and *horebukk* 'lecher' illustrate this.[8] Here we are dealing with grave sins that the people in question may only have committed once. However, quality compensates for low quantity.

I depict the relationship between construal, quality and quantity informally in Figure 2 on the next page. Let the origin represent the observer, i. e. the perspective of the majority from which a situation is construed. The vertical axis captures the qualitative dimension, where a high value indicates that something is construed as highly negative. The horizontal axis represents the quantitative dimension, where a high value corresponds to excessive quantity. What can be called "acceptable behavior" is represented by a sector, while everything outside this sector is "bad". In order to be outside the sector, an action must be high with regard to quality or quantity (or both). *Horebukk* 'lecher' and *убийца* 'murderer' receive high values for bad quality, but not for excessive quantity, whereas *skravlebøtte* 'chatterbox' and *балаболка* 'chatterbox' score high for quantity, but not for bad quality. In other words, both quality and quantity contribute to the construal as negative, and a high value for one factor compensates for a low value for the other.

[6] WHY DERIVATION IN RUSSIAN?

In the beginning of this article, we saw that the Norwegian and Russian constructions are similar in meaning, but different in form. So far we have been concerned with meaning, but now we turn to form. Is it possible to explain the differences be-

[8] Not all native speakers of Norwegian share my intuition that *horebukk* can be used felicitously about a person who has committed the relevant sin only once. An (even) clearer example is *hanrei* 'cuckold, deceived husband' — it is enough to be deceived on one occasion to qualify as a full-fledged member of the category. Since the difference between *horebukk* and *hanrei* may reflect a higher tolerance for promiscuity for men in traditional culture, we are arguably dealing with an example of the lexicalization of sexist ideologies in language.

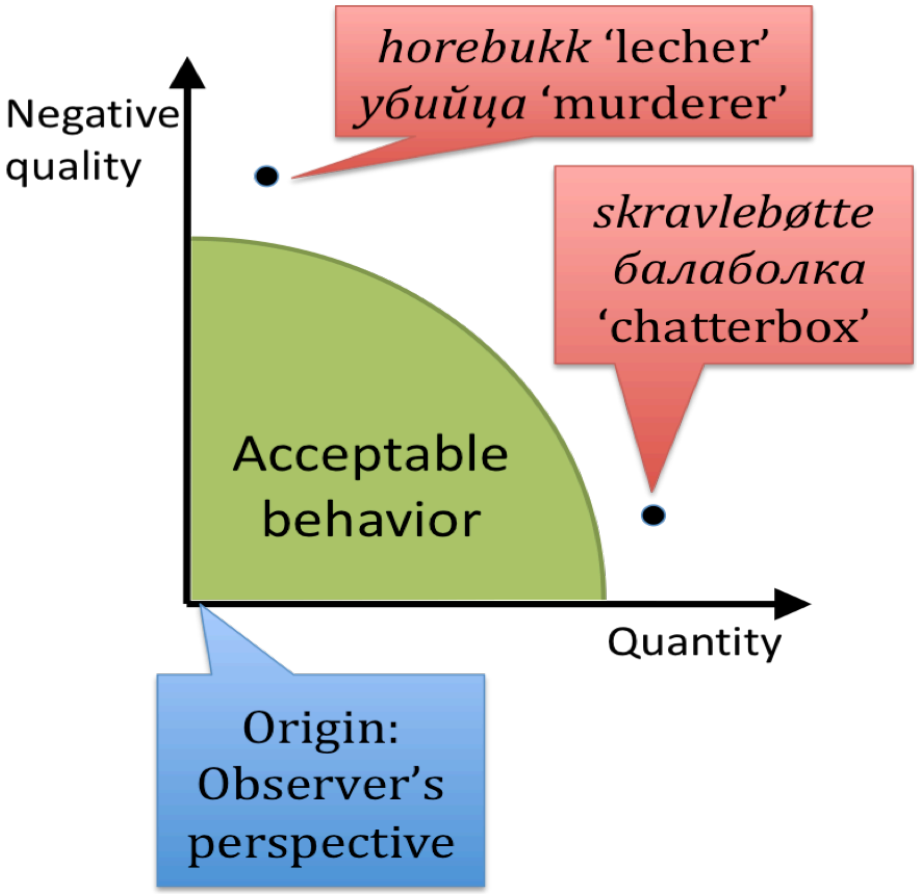


FIGURE 2: Construal, quality and quantity

tween the Norwegian and Russian constructions? And what exactly is it we need to explain? Is it the fact that Norwegian employs compounding that demands an explanation, or should we rather try to motivate the use of derivational morphology in Russian? Adopting a grammaticalization perspective, I suggest focusing on Russian. By grammaticalization, I mean the historical process whereby lexical morphemes develop into grammatical morphemes, i. e. affixes (Bybee et al. 1994, 4; Hopper & Traugott 1993, 2). Derivation is further down the grammaticalization cline, since compounding is closer to the mere juxtaposition of two separate lexical items. The question is why Russian has proceeded further along the grammaticalization cline. In other words, we must ask whether there is systemic motivation in the Russian grammar for expressing negative characteristics by means of derivation.

A first observation is that Russian is heavily invested in derivation. As shown by Janda (this volume), Russian displays almost five times as many derivational suffixes as Norwegian (274 in Russian compared to 57 in Norwegian). Admittedly, Janda's analysis is concerned only with metonymical derivational patterns, but since she shows that metonymy is pervasive in derivation, it stands to reason that

Janda's results are indicative of the word formation system as a whole.

A second observation is that Russian has a rich system of diminutive derivation, a feature that Norwegian word formation lacks. As is commonly observed, the most important function of Russian diminutive suffixes is not to indicate the small size of the referent, but rather to characterize the speaker's emotional relationship to the referent (see e. g. Townsend 1975; Wierzbicka 1992). If a speaker chooses the diminutive form *злючка* 'cross-patch' instead of the neutral *злюка*, this indicates that the speaker has a more emotional relationship to the person in question. In other words, diminutives incorporate a system of evaluation in Russian word formation. This provides systemic motivation for the Russian Characteristic Derivation Construction, which also expresses evaluation through derivational morphology.

A third point concerns the role of the *a*-declension in Russian.[9] This inflection class comprises both inanimate and animate nouns. Among animates, those denoting persons are particularly relevant for our purposes. As pointed out in Nesset (2001, 201), the personal nouns in the *a*-declension largely fall into three categories:[10]

(10) Personal nouns in the *a*-declension:
  a. Females (e. g. *соседка* 'female neighbor')
  b. Hypocoristic forms of names (e. g. *Ваня* from *Иван*)
  c. Characteristic nouns of common gender (e.g. *плакса* 'crybaby')

I suggest that the *a*-declension serves as a repository for marked persons in Russian morphology. Nouns denoting females are often morphologically marked in the sense that they are derived from nouns denoting males. For instance, *соседка* 'female neighbor' is derived from *сосед*, which is used about male neighbors or neighbors in general. From the point of view of feminist theory, one can even argue that words for females are semantically marked, since traditional stereotypes tend to construe women as the "second sex" (de Beauvoir 1993, see also Nesset 2001 for extensive discussion summarized in Dirven et al. 2007, 1231–1232). The hypocoristic nouns in (10b) are semantically marked as well; such nouns are used when the speaker has a special, intimate relationship to the addressee. Finally, the common gender nouns investigated in the present study are marked in the sense that they tend to involve negative evaluation of the referent. In other words, all the groups in (10) involve marked persons. This shows that the fact

[9] I use the term "*a*-declension" to refer to the inflection class of nouns ending in /a/ in the Nominative Singular. This class is called the "first declension" in the Russian tradition, but labeled the "second declension" in Western works on Russian grammar.

[10] In addition, there are kinship terms like *папа* 'daddy' that are related to the hypocoristic forms in (10b) and some masculine characteristic nouns (e.g. *волокита* 'skirtchaser') that resemble the common gender nouns in (10c).

that the Russian Characteristic Derivation Construction involves nouns in the *a*-declension is not an arbitrary idiosyncrasy. On the contrary, the *a*-declension's function as a repository for marked persons in Russian provides systemic motivation for expressing negative evaluation through the Russian Characteristic Derivation Construction.

[7] CONCLUSION

The juxtaposition of the Norwegian Characteristic Compound Construction and the Russian Characteristic Derivation Construction has shown that they are different in form, but strikingly similar in meaning. Both morphological constructions give (mostly) negative characterizations of people, covering large, but very similar sets of sins and vices in both languages. We have seen that the two constructions cannot be properly understood without recourse to metonymy, and the proposed analysis therefore lends support to the idea that metonymy is pervasive in word formation and grammar in general. As for form, the Russian expression of negative characteristics through derivation receives systemic motivation from the language's rich system of diminutive derivation, as well as from the function of the *a*-declension as a repository of marked persons.

This study indicates that detailed contrastive analysis can shed light on the similarities and differences between Norwegian and Russian and identify properties that might have been overlooked in analyses of either language in isolation. But first and foremost this study illustrates how contrastive analysis informs linguistic theory — in our case by demonstrating the pervasiveness of metonymy in grammar.

ACKNOWLEDGMENTS

I would like to thank Hans-Olav Enger, Atle Grønn, Laura A. Janda and audiences in Oslo and Prague for comments on earlier versions of this article. Thanks also to my employer, the University of Tromsø, and Sparebankens gavefond for financial support.

REFERENCES

Aronoff, Mark & Kirsten Fudeman. 2005. *What is morphology?* Oxford: Blackwell.

de Beauvoir, Simone. 1993. *The second sex.* London: David Campbell.

Benczes, Réka. 2006. *Creative compounding in English.* Amsterdam, Philadelphia: John Benjamins.

Booij, Geert. 2005. Compounding and derivation: evidence for Construction Morphology. In Wolfgang U. Dressler, Francesco Gardani & Markus A. Pöchtrager

(eds.), *Morphology and its demarcations: selected papers from the 11th morphology meeting, Vienna, February 2004*, 109–132. Amsterdam, Philadelphia: John Benjamins.

Booij, Geert. 2009. Compounding and Construction Morphology. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 201–216. Oxford: Oxford University Press.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan. 2007. *Frequency and the organization of language*. Oxford: Oxford University Press.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar*. Chicago, London: The University of Chicago Press.

Carstairs-McCarthy, Andrew. 1992. *Current morphology*. London, New York: Routledge.

Croft, William. 2002. The role of domains in the interpretation of metaphors and metonymies. In René Dirven & Ralf Pörings (eds.), *Metaphor and metonymy in comparison and contrast*, 161–205. Amsterdam, Philadelphia: John Benjamins.

Dirven, René. 1998. Conversion as a conceptual metonymy of event schemata. In Klaus-Uwe Panther & Günter Radden (eds.), *Metonymy in language and thought*, 275–288. Amsterdam, Philadelphia: John Benjamins.

Dirven, René, Frank Polzenhagen & Hans-Georg Wolf. 2007. Cognitive linguistics, ideology, and critical discourse analysis. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 1222–1240. Oxford: Oxford University Press.

Enfield, Nick J. 2002. Ethnosyntax: Introduction. In Nick J. Enfield (ed.), *Ethnosyntax: explorations in culture and grammar*, 1–30. Oxford: Oxford University Press.

Goldberg, Adele E. 1995. *Constructions*. Chicago: The Chicago University Press.

Goldberg, Adele E. 2006. *Constructions at work*. Oxford: Oxford University Press.

Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.

Hopper, Paul & Elizabeth Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Hovdenak, Marit, Laurits Killingbergtrø, Arne Lauvhjell, Sigurd Nordlie, Magne Rommetveit & Dagfinn Worren (eds.). 2001. *Nynorskordboka*. Oslo: Det Norske Samlaget. http://www.dokpro.uio.no.

Janda, Laura A. this volume. Russian word-formation in contrast with Czech and Norwegian.

Knudsen, Trygve, Alf Sommerfelt, K. Emil Bødtker & Harald Noreng (eds.). 1937–1957. *Norsk riksmålsordbok*. Oslo: Aschehoug. http://www.ordnett.no.

Kövecses, Zóltan. 2002. *Metaphor: a practical introduction*. Oxford: Oxford University Press.

Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago: The University of Chicago Press.

Landrø, Marit Ingebjørg & Boye Wangensteen. 1986. *Bokmålsordboka*. Bergen, Oslo, Stavanger, Tromsø: Universitetsforlaget. http://www.dokpro.uio.no.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar*, vol. 1. Stanford: Stanford University Press.

Langacker, Ronald W. 1991. *Foundations of cognitive grammar*, vol. 2. Stanford: Stanford University Press.

Langacker, Ronald W. 2008. *Cognitive grammar: a basic introduction*. Oxford: Oxford University Press.

Langacker, Ronald W. 2009. Metonymic grammar. In Karl-Uwe Panther, Linda L. Thornburg & Antonio Barcelona (eds.), *Metonymy and metaphor in grammar*, 45–71. Amsterdam, Philadelphia: John Benjamins.

Nesset, Tore. 2001. How pervasive are sexist ideologies in grammar? In René Dirven, Bruce Hawkins & Esra Sandikcioglu (eds.), *Language and ideology. Volume I: Theoretical cognitive approaches*, 197–226. Amsterdam, Philadelphia: John Benjamins.

Padučeva, Elena Viktorovna. 2004. *Dinamičeskie modeli v semantike leksiki*. Moscow: Jazyki slavjanskoj kul'tury.

Peirsman, Yves & Dirk Geeraerts. 2006. Metonymy as a prototypical category. *Cognitive Linguistics* 17. 269–316.

Radden, Günter & Zóltan Kövecses. 1999. Towards a theory of metonymy. In Karl-Uwe Panther (ed.), *Metonymy in language and thought*, 17–60. Amsterdam, Philadelphia: John Benjamins.

Townsend, Charles E. 1975. *Russian word-formation*. Columbus, OH: Slavica Publishes.

Verhagen, Arie. 2007. Construal and perspectivization. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 48–81. Oxford: Oxford University Press.

Wierzbicka, Anna. 1992. *Semantics, culture and cognition*. Oxford, New York: Oxford University Press.

Zaliznjak, Andrej Anatol'jevič. 1977. *Grammatičeskij slovar' russkogo jazyka*. Moscow: Russkij jazyk.

AUTHOR CONTACT INFORMATION

Tore Nesset
University of Tromsø
Department of Language and Linguistics
NO-9037 Tromsø
Norway
tore.nesset@uit.no

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 281–318. (ISSN 1890-9639)
http://www.journals.uio.no/osla

INPUT FREQUENCIES IN PROCESSING OF VERBAL MORPHOLOGY IN L1 AND L2: EVIDENCE FROM RUSSIAN

ELENA TKACHENKO AND TATIANA CHERNIGOVSKAYA
University of Oslo and St. Petersburg State University

ABSTRACT

In this study we take a usage-based perspective on the analysis of data from the acquisition of verbal morphology by Norwegian adult learners of L2 Russian, as compared to children acquiring Russian as an L1. According to the usage-based theories, language learning is input-driven and frequency of occurrence of grammatical structures and lexical items in the input plays a key role in this process. We have analysed to what extent the acquisition and processing of Russian verbal morphology by children and adult L2 learners is dependent on the input factors, in particular on type and token frequencies. Our analysis of the L2 input based on the written material used in the instruction shows a different distribution of frequencies as compared to the target language at large. The results of the tests that elicited present tense forms of verbs belonging to four different inflectional classes (-AJ-,-A-, -I-, and -OVA-) have demonstrated that for both Russian children and L2 learners type frequency appears to be an important factor, influencing both correct stem recognition and generalisations. The results have also demonstrated token frequency effects. For L2 learners we observed also effects of formal instruction and greater reliance on morphological cues. In spite of the fact that L2 learners did not match completely any of the child groups, there are many similarities between L1 and L2 morphological processing, the main one being the role of frequency.

[1] INTRODUCTION

Usage-based models view language learning as a process first and foremost relying on the general cognitive learning principles (Langacker 1987; Kemmer & Barlow 2000; Langacker 2000; Bybee & Hopper 2001; Tomasello 2003; Goldberg 2006). According to this view the linguistic knowledge is acquired in a bottom-up direction, so that actual language use shapes the nature of generalisations that emerge from the analysis of patterns occurring in the input: "The children are picking up frequent patterns from what they hear around them, and only slowly making more abstract generalisations as the database of related utterances grows" (Ellis 2003, 70).

In recent years a handful of studies have employed usage-based theories to account for patterns of language processing in second language and bilingual acquisition (e. g. Ellis 1998; Gathercole 2007) (see also Ellis 2002; Robinson & Ellis 2008 for a review). In this study we explore the data from the acquisition of verbal morphology by adult L2 learners of Russian, as compared to children acquiring it as an L1. Second language processing presents an interesting field of research because the input of L2 learners may potentially differ in many respects from the input that children are exposed to. In this paper we aim at investigating to what extent the acquisition and processing of Russian verbal morphology by children and adult L2 learners is dependent on the input.

[1.1] *Differences and similarities in the input to L1 and L2 learners*

One of the continuing issues in second language acquisition (SLA) research has for many years been whether L2 learners acquire/learn the language in the same way as children acquiring their first language (L1). The differences between L1 and L2 acquisition are evident in many domains: ultimate attainment, achieved fluency, accuracy and so forth (see e. g. Hyltenstam 1992; Hyltenstam & Abrahamsson 2000; Singleton 2003). The sources of explanations for these differences are, however, somewhat differently shaped in different theoretical approaches: within the generative framework the discussion has focused mainly on the availability of the UG in L2 acquisition (White 1989; Eubank 1991; Schwartz 1992; Flynn 1996; White 2003); cognitive approaches to language acquisition view the process of learning a language as an instantiation of all other types of learning and for this reason try to explain these differences by a range of factors that might influence language processing, including variation in learning environments, adults' better developed analytical thinking, influence of the native language, as well as amount of input received and its properties. Another source of variation in L2 processing may lie in learner-internal factors such as motivation, aptitude, language analytic abilities, attitudes and so forth (Dörnyei & Skehan 2002; Robinson 2002; Masgoret & Gardner 2003; Dörnyei 2005).

Comparing the contexts of L1 and L2 acquisition, there are several points where these differ. The initial point when the process of acquisition starts differs enormously for the children acquiring their first language and adults beginning to learn an L2. When adults start learning a second language, they already possess first, a substantial knowledge of the world, and second, the knowledge of their L1 (Gass 1996); for children the language acquisition and conceptual and cognitive development proceed simultaneously. When children acquire the language, they gradually extract the grammatical forms from the input and build their grammatical system in a piecemeal fashion. Adult L2 learners, especially in a classroom context, can treat language as an object of learning in itself and as a consequence, can deal with the linguistic system structurally. Furthermore,

L2 learners can be exposed to formal instruction and for them explicit rules can be formulated. Thus, while children acquire language implicitly, adult L2 learners have an opportunity for explicit learning (see a discussion on the interface between explicit and implicit learning in SLA in e. g. Ellis 1994; DeKeyser 1997; MacWhinney 1997; Ellis 2005). Views on the role of explicit grammar instruction in second language classrooms have been changing throughout many years of SLA research, and various teaching methods allot it different places and varying degrees of importance (Ellis 1990; Celce-Murcia 1991). Focus on form in the formal instruction to L2 learners is yet another point of difference between L1 and L2 acquisition: while children acquire morphology through communication, and their first focus of attention is meaning, L2 learners are to a lesser or greater extent (depending on the teaching method) focused on grammatical forms, even in communicative methods (Long 1991; Spada 1997; Doughty & Williams 1998; Long & Robinson 1998; Norris & Ortega 2000, 2001). The last point we will mention here is the significant difference in the input for L1 and L2 learners. Children receive a lot of input, and this input is natural; while L2 learners' input can be very limited, and it is characterised by some specific features (Ellis 2003, 72), which we discuss later.

In this study we consider in detail two of the above mentioned differences: formal instruction and focus on form, and the nature of input in the L2 learning context. Later in this paper we will analyse in detail how the target morphological structures are presented to the L2 learners who participated in this study, what kind of instruction the learners received in the target forms, and what is their distribution in the L2 input, as compared to the target language (TL) input at large.

[1.2] *Past tense debate and L2 morphological processing*

Acquisition of verbal morphology has for many years been an issue of much debate within linguistic theory of acquisition, mental grammar representation and cognitive processing. Within the "nature vs. nurture" discussion, an issue concerning the acquisition of the English past tense has attracted much attention and this "past tense debate" continues (Marslen-Wilson & Tyler 1998; Pinker 2001; McClelland & Patterson 2002a,b; Pinker & Ullman 2002; Marslen-Wilson & Tyler 2005). The discussion centres on the question of whether the processing of regular and irregular morphological forms is driven by two distinct mechanisms (the Dual Mechanism Account), or by one single mechanism (the Single Mechanism Account). The proponents of the Dual Mechanism Account claim that regular forms are computed by a rule-processing system, while irregular morphological forms are processed in associative memory (Pinker & Prince 1988; Marcus et al. 1992; Prasada & Pinker 1993; Marcus 1995; Marcus et al. 1995; Clahsen 1997; Ullman 1997; Pinker 1999; Ullman 1999; Pinker 2001; Ullman 2001b,a; Pinker & Ullman 2002; Clahsen 2006; Clahsen & Felser 2006b). The opposite view claims that both

regular and irregular forms are processed by one single mechanism in associative memory (Bybee & Slobin 1982; Rumelhart & McClelland 1986; Langacker 1987; MacWhinney & Leinbach 1991; Plunkett & Marchman 1991, 1993; Bybee 1995; Plunkett & Marchman 1996; Langacker 2000; Bybee & Hopper 2001; Eddington 2002). The major point of disagreement between these competing accounts lies in their predictions regarding the role of input factors in processing of inflectional morphology: whereas the Dual Mechanism Account predicts that frequency factors influence the acquisition and processing of irregular forms, but not regular ones; according to the Single Mechanism Account, processing of all types of inflectional forms, both regular and irregular, depends heavily on input frequencies.

The proponents of the Dual Mechanism Account have lately extended their theories and formulated two different accounts of morphological processing in L2 acquisition. One such model was worked out by Clahsen and his colleagues. According to the Shallow Structure Hypothesis (SSH) formulated recently (Clahsen & Felser 2006b,a), the L2 grammatical processing is different from that of L1 particularly in the area of syntax, while morphological processing in L1 and L2 does not differ and shows a dissociation of rule-based and associative patterning. They advocate this view based on their findings in a study on morphological processing of German participle inflection by native speakers and L2 learners (Hahne et al. 2006). The results of this study have demonstrated that the L2 learners showed different ERP responses to violations of regular and irregular inflection, and their responses were similar to that of the native speakers. The conclusion made was that in L2 and in L1 different processes are involved in regular and irregular inflection. However, this study also included a second ERP experiment on the processing of German noun plurals, and for nouns the findings were different: while the L2 learners performed similarly to native speakers for participles, they showed different patterns for noun plurals. The explanation provided was that the L2 learners' processing of participles was more automatised than their processing of noun plurals due to the fact that German noun plural system is rather unusual with a low frequency regular default and more diverse in inflections and, thus, more difficult for L2 learners to acquire. Consequently, the overall results of this study cannot in full support the hypothesis by Clahsen and his colleagues that L2 morphological processing does not differ from L1 morphological processing. This idea was also criticised in a commentary by Ullman (2006), who pointed out several important differences between the SSH and his own model and argued that most of the data can be explained by the declarative/procedural model (Ullman 1997, 2001a,b,c, 2004).

The declarative/procedural model (DPM) proposed by Ullman was initially worked out for L1 morphological processing. The core assumption of this model is that different aspects of linguistic processing are linked to different brain memory systems. The declarative memory system subserves the associative, lexical-

based processing, whereas the procedural memory subserves the grammatical, rule-based processing. In the same way as other dual-mechanism models, the declarative/procedural model makes specific assumptions with regard to the dissociation of morphological processing of regular and irregular forms: regular forms are processed in the procedural memory, while irregular forms — in the declarative memory. However, for the L2 morphological processing, Ullman hypothesises that "processing of linguistic forms that are computed grammatically by procedural memory in L1 is expected to be dependent to a greater extent upon declarative memory in L2" (Ullman 2001c, 109). Thus, morphological processing in L2 is seen as different from that in L1. Since the lexicon/grammar dissociations posited for L1 can be weaker or even absent in L2, it is predicted that both regular and irregular forms can be computed and processed by the associative memory in L2, and frequency effects can be expected for both regular and irregular forms. This particular point is similar to the predictions made by the usage-based models that also expect that all morphological forms will be influenced by frequency of occurrence.

The usage-based theories view frequency as an important factor in language learning. As Ellis put it, "language processing is intimately tuned to input frequency" (Ellis 2002, 143). This principle holds for L1 and L2 processing, and Ellis claims that "the L1 acquisition sequence — from formulas, through low-scope patterns, to constructions — could serve well as a reasonable default in guiding the investigation of the ways in which exemplars and their type and token frequencies determine the second language acquisition of structure" (Ellis 2002, 170).

TYPE FREQUENCY refers "to the frequency of occurrence of a linguistic pattern, or in other words, to the size of a certain class of words using this pattern" (Gor 2007, 371). TOKEN FREQUENCY, on the other hand, shows "how often a language user encounters a certain word, either by producing it or hearing it produced by other speakers" (Gor 2007, 371). The situation of classroom L2 acquisition usually restricts the L2 learner's vocabulary to the most frequent items. L2 learners may have a very limited access to natural target language input, especially when a second language is learned as a foreign language.[1] (see e. g. Gilmore 2007). For this reason, target language input frequencies can be distorted to some extent in the L2 input. Although L2 learners can be exposed to different verb types in their input, the relative proportions of these types may not be the same as in the TL at large (Gor 2004; Gor & Chernigovskaya 2005; Gor 2007). Formal instruction and

[1] The abbreviated term "L2 learner" can potentially refer to both second language learners (i. e. those who learn the language either naturalistically or in a classroom in the country where this language is spoken as an L1), as well as to foreign language learners (i. e. those who learn the language in instructional settings in a country where this language is not spoken as an L1, usually in the country of their origin) (Nizegorodcew 2007). It should be noted, however, that the distinction between different types of learners is not categorical. For this study this opposition is not relevant (see section [3] for a further description of the L2 subjects).

focus on form can also contribute to such a distortion of input frequencies: when a particular pattern is being learned in a classroom context, it can be temporarily enhanced in the L2 input (Sharwood Smith 1993; Alanen 1995; White 1998), so that this structure can seem very frequent to L2 learners at a certain stage, while it is not necessarily as frequent in the TL at large. Since the vocabulary that is introduced in a classroom context is restricted, token frequencies of individual verbs may not be the same as the ones found in the TL at large. Most obviously, the L2 learners are exposed to the most frequent items, so that they become even more frequent in the L2 input, while the items having lower token frequencies can be even less frequent or be completely lacking in the L2 input. Thus, these differences in the L2 input can lead to the differences in generalisations that L2 learners and children can extract from their linguistic experiences.

Summarising this theoretical section, we can say that the three models that have been suggested for L2 morphological processing differ from each other on three counts: first, whether they theoretically advocate the Dual or the Single mechanism account view; second, whether they consider the principles of morphological processing in L1 and L2 as similar or different; and third, whether they assume that processing of regular and irregular morphological forms in L2 learning should be distinct or not.

	SSH (Clahsen & Felser 2006a,b)	DPM (Ullman 2001c, 2004)	UBA (Ellis 2002, 2003)
Initial theoretical position on morphological processing	Dual mechanism	Dual mechanism	Single mechanism
L2 vs. L1 morphological processing	Similar	Different	Similar
Differences between regular and irregular processing in L2	Different	Similar	Similar

TABLE 1: Theoretical positions of different theories of L2 morphological processing

The usage-based approach (UBA) that we take as a point of departure in this study advocates the Single Mechanism Account position. Consequently, we shall test whether the acquisition of verbal morphology is influenced by frequency factors. Our predictions are as follows:

(i) The verbal classes that are most frequent in the input should be acquired earlier than the verbal classes that occur rarely (type frequency effect)

(ii) The verbal tokens that occur frequently should be acquired better than the verbs for which frequency of occurrence in the input is low (token frequency effect)

(iii) Type and token frequency effects should be found in both L1 and L2 data.

In order to address the questions concerning input frequencies, we need further to describe the target structures and analyse their distribution in the input, which we turn to in the next section.

[2] RUSSIAN PRESENT TENSE MORPHOLOGY AND INPUT FREQUENCIES

Russian is a morphologically rich language, with numerous verb classes and complex inflectional paradigms. There are two sets of inflections for conjugating verbs in the present tense form, which distinguish between 1st and 2nd conjugation (e. g. Švedova 1980, 647). In the Cognitive Grammar approach these alternative endings can be described in terms of schemas that allow capture of a generalisation of the two conjugational patterns (Nesset 2008).

In addition to adding the inflection of either 1st or 2nd conjugation, the stems of the verbs are subject to some alternations between the forms of the present tense and imperative subparadigms, in which the stem ends in a consonant, and the past tense and infinitive subparadigms, in which the stem ends in a vowel (e. g. Švedova 1980, 646, Nesset 2004, 66–67). Different types of alternations served as a basis for categorising Russian verbs into several inflectional classes (see e. g. Cornyn 1948; Jakobson 1948; Švedova 1980; Scatton 1984; Nesset 1996).

In this study we chose to focus on the acquisition of four verbal classes: -AJ-, -A-, -I-, and -OVA- (according to Jakobson's 1948 classification), as these differ in type frequency, productivity and morphological complexity and thus seem to be the most suitable for testing our hypotheses. Table 2 on the following page below characterises these four classes in terms of 1) the alternations that occur with the stem in the present tense forms in relation to the stem final segments in the infinitive/past tense forms, 2) the conjugation type, 3) the complexity of paradigm determined by the presence or absence of consonant mutations and stress shifts in the present tense forms, 4) type frequency of the inflectional class and productivity of the patterns.

The verbs belonging to the -AJ- and -A- classes look similar in the infinitive and past tense forms, however, their present tense inflectional forms differ: while the -AJ- class is morphologically simple and besides the suffix alternations (a $\sim$ aj) no other changes occur in the stem, the -A- class is morphologically complex, having both consonant mutations and stress shifts in present tense forms. These two classes also differ in type frequency: the -AJ- class has very high type frequency and is very productive, whereas the -A- class has low type frequency (Townsend 1975; Zaliznjak 1980; Slioussar 2003). The -I- class is similar to the -A- class in its

Class	Alternation (inf/past ∼ present)	Conjug. type	Morpho-logical complexity	Type frequency/ Productivity	Examples (Inf/3pl/1sg)
-AJ-	a ∼ aj	1	—	Very high/ Productive	igr*á*-t igr*áj*-ut igr*áj*-u
-A-	a ∼ ∅	1	Consonant mutations, stress shifts	Low/ Unproductive	pis*á*-t píš-ut piš-ú
-I-	i ∼ ∅	2	Consonant mutations, stress shifts	High/ Productive	nos*í*-t nós'-at noš-ú
-OVA-	ova ∼ uj	1	Suffix alternation	Medium/ Productive	ris*ová*-t ris*új*-ut ris*új*-u

TABLE 2: Description of the four inflectional classes in Russian

morphological complexity, as many verbs belonging to this inflectional pattern have consonant mutations and stress shifts in several present tense forms, but in contrast to the -A- class, the -I- class has high type frequency (Townsend 1975; Zaliznjak 1980; Slioussar 2003). The -OVA- class can be said to have both medium morphological complexity (as this pattern involves suffix alternation -ova ∼ -uj-), and medium type frequency. A particular characteristic of this class is that the suffix -OVA- can be perceived as a clear morphological cue, which can unambiguously point at belonging to this inflectional pattern. Thus, we may suggest that the morphological complexity of the –OVA-class might be compensated by the strong morphological cue, making this pattern easy to identify for the learners. However, the -OVA- verbs tend to have quite low token frequency (Gagarina 2002), which can also slow down the acquisition of this pattern (see e. g. the critical mass hypothesis in Marchman & Bates 1994).

Since the type and token frequencies in the L2 input may differ from the distributions found in the TL at large, there is a need to analyse it in more detail. The L2 input has been analysed in three respects: 1) the explicit explanations on different inflectional patterns provided in the grammar books used by our L2 subjects; 2) the presentation of the present tense formation in the textbooks used in the instruction; 3) the distribution of the four verb classes in the L2 input, as well as correspondences in token frequency rates for the items that were included in the test.

Two grammar books on the syllabus for Norwegian L2 learners of Russian (Mathiassen 1996; Bach Nielsen 2003) present explicitly various patterns of verbal inflections separately, and these are exemplified by several verbs for each pattern. Only one of the grammars (Mathiassen 1996) touches upon the issue of productivity of the inflectional patterns and lists five productive classes, three of them are included in this study -AJ-, -I-, -OVA-), as well as several unproductive classes, among them the -A- class. Consonant mutations and stress shifts are explained thoroughly in the grammar books. Thus, the grammar books for L2 learners provide a systematic description of the verbal system and the present tense formation of different verb types. It is, however, questionable to what extent the learners can apply this explicit information and the rules when they face a real task of conjugating a verb in the present tense (Robinson 1996).

The analysis of the textbooks used in the instruction (Bjerkeng et al. 2000; Bjerkeng & Bräger 2002) shows that to a large extent the learners are encouraged to learn the conjugation patterns on an item-by-item basis. No explicit explanations regarding various patterns of inflections are provided in the textbooks, but different patterns are exemplified by present tense conjugational paradigms of several frequent verbs. With the exception of the -OVA- class, none of the verb classes is presented separately and prominently as a salient class. Consequently, the learners may be inclined to rely on rote learning, rather than on generation of the target forms by application of any formal rule. However, such rote learning can lead to generalisations of some patterns that occur in the input frequently enough to be represented abstractly as a schema. An important role in this process of forming generalisations is played by type frequency of the patterns (Bybee 1995).

For the purpose of this study we tried to estimate the input frequencies in the L2 input and analysed the distribution of verb types in the L2 input. This analysis is restricted to the written material used in the instruction in the beginners' course that the majority of the L2 subjects in this study took at the University of Oslo. The instructional material analysed included the following main sources: Texts in the instructional set *Sosedi* (Bjerkeng et al. 2000; Bjerkeng & Bräger 2002); exercises focused on verbal conjugation and tense formation from *Russian in Exercises* (Chavronina & Širočenskaja 1999, 14–33);[2] the compendium used in the practical Russian and text reading class, which included authentic texts from Russian literature, fairy-tales, business documentation, dialogue examples etc.

All the verbs in the following sources were registered (approximately 5,700 forms in total), then lemmatised and tagged with grammatical information (tense,

[2] Since the learners are expected to produce the inflectional forms in the exercises, it is debatable to what extent we can call this input in the strict sense. In spite of the fact that such forms can rather be regarded as output, we consider it legitimate for our purposes to regard these uses as input, since the exercises are usually checked or completed in the classroom.

person and number of the form used, conjugation type and inflectional class). The frequencies of the verbs belonging to the four classes focused on in this study in the TL and in L2 input are summarised in Table 3.[3]

	-AJ-	-A-	-I-	-OVA-
TL TYPE FREQUENCIES[4]	11,814 (43%)	940 (3%)	7,019 (25%)	2,816 (10%)
L2 type frequencies: Number of uses (total 5,700)	1480 (26%)	453 (8%)	1221 (21%)	102 (2%)
L2 TYPE FREQUENCIES Number of different verbal lexemes (total 960)	296 (31%)	40 (4%)	225 23%)	48 (5%)

TABLE 3: Distribution of Russian verbs in L2 input across the four classes in focus

The distribution of the four verbal classes under consideration is different in the L2 input compared to the TL input. Although the -AJ- and -I- classes are the two largest classes both in the TL and L2 input, the difference in their type frequencies in the L2 input is less prominent. On the other hand, the type frequency of the -OVA- class is much lower in the L2 input than in the TL input, and this class is very similar to the -A- class in the type frequency rate in the L2 input. This may be due to the fact that -OVA- verbs tend to have low token frequency, and for this reason many of them do not occur in the input for L2 learners, as it is typically restricted to the most frequent lexical items. Summing up these findings, we can say that in the L2 input, the -AJ- and -I- classes have high type frequency, whereas the -OVA- and -A- classes have low type frequency.

Taking into account type frequency and morphological complexity separately, the four classes under consideration can be placed in a matrix, as shown in Table 4 on the facing page.

Several predictions can be made about the acquisition of these classes by Russian children and L2 language learners. Isolating the frequency factor from morphological complexity, the predictions formulated below regard rates of correct

[3] The counts represented here include all verb forms, rather than only present tense forms. The argument for doing this is that the glossaries to the textbooks provide the information on other forms, which are sufficient for figuring out the whole present tense paradigm of the verb. The observations made in the classroom have also shown that the teacher tended to explicitly provide the present tense forms for new verbs that occurred in the texts.

[4] The numbers here are based on the database counts performed on Zaliznjak's Grammatical Dictionary, which includes 27 408 verbs. Due to some inconsistencies found in the referred sources (Slioussar 2003; Gor 2004) the numbers here are approximations.

	High type frequency	Low type frequency
Low morphological complexity	-AJ-	-OVA-[5]
High morphological complexity	-I-	-A-

TABLE 4: Type frequency and morphological complexity of the four Russian classes

stem recognition, rather than correctly produced forms.

(i) The two classes with high type frequency, the -AJ- and -I- classes, should be acquired better and earlier by the learners. As these classes have relatively high type frequency in both TL and L2 input, we do not expect any differences between the children acquiring Russian as their L1 and adult L2 learners of Russian.

(ii) The two most frequent patterns are also expected to be applied more frequently to nonce verbs, as well as serve as a basis for overgeneralisation errors.

(iii) The -A- class, having the lowest type frequency and being morphologically complex, should be acquired later by the children, and L2 learners may have problems with mastering this pattern.

(iv) Since the type frequency of the -OVA- class turned out to differ in the TL and L2 input, we can expect to find some differences in the acquisition of this class by children and by adult L2 learners.

[3] METHODOLOGY

[3.1] *Subjects*

Four groups of informants participated in the study: three groups of Russian children (aged 4, 6 and 8) and one group of Norwegian learners of L2 Russian (Table 5 on the next page). The children were recruited for the experiment from several kindergartens and schools in St. Petersburg. The Norwegian learners of L2 Russian were all recruited from the Russian language programme at the University of Oslo.

The L2 learners in the study, although they were all recruited from the same educational programme, varied to some extent in several respects. With one exception, the age of the L2 learners was between 20 and 30 years old. The vast majority of the subjects studied Russian intensively (classes 4–5 days a week). The

[5] Although the –AJ- and –OVA-classes are not exactly on the same level of morphological complexity (since the –OVA-class involves the suffix alternation), as we claimed above, this can be compensated by the strong morphological cue, helping the learners to identify the –OVA-class and inflect its members correctly.

Group	Number of subjects	Males	Females
Russian children 4 y. o.	30	15	15
Russian children 6 y. o.	30	15	15
Russian children 8 y. o.	21	11	10
Adult L2 learners	25	10	15

TABLE 5: Overview of subjects in the study

majority of them had visited Russia for at least several weeks, whereas some lived in Russia for longer periods of time (1–3 months on average). With the exception of four subjects who studied Russian at high school, all other subjects started learning Russian at the University. The overall period the subjects were studying Russian varied from 0.5–6 years.

[3.2] *Testing material and procedure*

All subjects were tested orally and individually with the same test battery, which consisted of two tests. The content in the tests was the same, but they differed in the form the stimuli were presented: in one of the tests the stimuli were presented in the infinitive, in the other — in the past tense plural form. The order of the tests and the order of the stimuli in each of them were controlled for: approximately half of the subjects performed the infinitive test first, and then the past tense test, and the other half was tested in the reverse order; the order of the verbs within the tests was random, furthermore, the order of the stimuli was counterbalanced, so that within each group half of the participants received the items in one order, and the other half in the other.

The testing material included 80 verbal stimuli in each test, evenly representing four Russian inflectional classes: -AJ-, -A-, -I-, -OVA- (described above). Half of the stimuli were real Russian verbs, and half were nonce verbs. Nonce verbs were created by changing the initial consonant segments of the real verbs, and they were meant to test the subjects' ability to generalise the inflectional patterns and use them productively. Within each class the real verbs were balanced for token frequency, at least for the TL at large. Token frequency counts were based on the Russian Frequency Dictionary (Zasorina & Agraev 1977).

Separate counts were performed to estimate the token frequency of the stems in the test for L2 learners, based on the materials included in the analysis in the previous section. The verbs in the test of which the stems had at least four occurrences in the sample were considered as having a high token frequency, while the verbal stimuli of which the stems occurred in the corpus less than four times or did not occur at all, were considered as low token frequency. With the exception of 8 items, L2 token frequency rates turned out to coincide with the token

frequency rates for the TL at large. We didn't consider this difference to be significant, and thus use the same token frequency rates in our analysis of results from both children and L2 learners, which also allows for a more accurate comparison between the groups.

The testing procedure was based on a design first used in Berko (1958): the subjects were shown a picture to each stimuli, and the experimenter told them either what the figures in the pictures like to do (in the infinitive condition), or what they were doing yesterday (in the past tense condition). The target forms for responses were present tense forms 3rd person plural and 1st person singular, these forms were elicited by asking the subjects to say what the figures in the pictures ARE doing (3pl), and then say the same thing about themselves (1sg). The testing was audio-taped, and later the responses were transcribed and systematically coded.

[3.3] *Measurements of subjects' performance*

We present the results on the subjects' performance in the tests measured on several main variables: total correct performance; correct stem recognition for each inflectional class; correctly produced forms for high and low token frequency verbs in each class; and generalisation patterns in responses to nonce stimuli.

The difference between measuring correct performance in terms of stem recognition vs. correctly produced forms needs a more detailed explanation. Measuring the performance in terms of correct stem recognition is done in order to eliminate the influence of the morphological complexity and thus assess the performance on different verb classes on a fair basis. The -A- and -I- classes are morphologically complex and involve consonant mutations and stress shifts in the present tense paradigm, which can result in lower rates of correctly produced forms. For example, for the verb *pisát* (write) the correct target forms are *píšut* (3pl) and *pišú* (1sg). Only these forms are considered correct in terms of correctly produced forms. On the other hand, when either consonant mutations are missing (*písut**, *pisú**), or stress shifts are incorrectly applied (*pišút**, *píšu**), or both (*pisút**, *písu**) the responses are incorrect in form, but they still can be considered correct in terms of stem recognition. When a wrong inflectional pattern is applied to an item, the responses are considered incorrect in terms of stem recognition: e. g. *pisájut**, *pisáju** (-AJ- pattern is applied to a verb from the -A- class).

[4] RESULTS

[4.1] *Total correct performance*

Total correct performance in two Russian tests for four subject groups is illustrated in Figure 1 on the following page.

The figure shows that all subjects' performance was better in the test with the stimuli in the infinitive, than with the stimuli in the past tense form. For

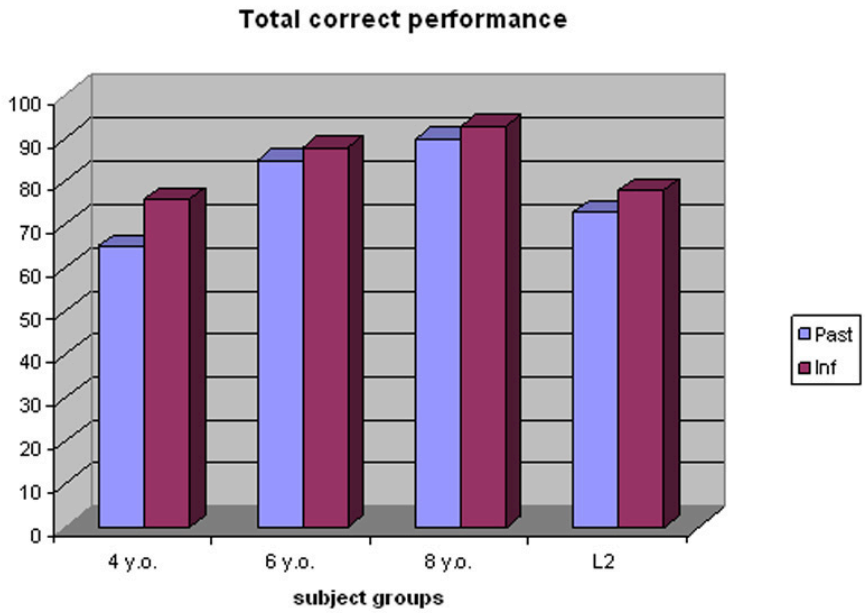


FIGURE 1: Total correct performance in the infinitive and past tense Russian test

L2 learners this difference might probably be explained by the influence of their learning setting: first, the infinitive is a base form cited in the dictionaries and for this reason this form may seem more prominent for L2 learners; second, in the tasks offered in the instruction, L2 learners typically generate inflectional forms from the infinitive. Thus, they are more used to this task than to a task of generating present tense forms from the past tense of the verbs. For the children, however, we should look for a different source of this unbalanced performance in the infinitive and past tense conditions.

We suggest several potential reasons for the children's better performance in the infinitive compared to the past tense condition. One possible explanation could be that the past tense forms occur less frequently in their input than the infinitives. However, according to an assumption expressed in Nesset (2004, 70), "while the infinitive is the citation form used in grammars and dictionaries, ...the more frequent past tense forms enjoy a more central status in speakers' and hearers' mental grammars". However, due to lack of any corpora of child-directed speech which could be used to test this assumption empirically, this explanation cannot be taken as plausible. Another potential explanation for this fact might lie in a higher cognitive complexity of a reference to past events as compared to "here and now". Our data suggest that some children had difficulties in distinguishing past and present forms in the test, which mirrored in more frequent repetitions of the stimuli in the past tense condition, especially for the youngest age group. This can be a result of an incomplete acquisition of the category of aspect: all the stimuli in the test were imperfectives, and previous studies on acquisition

of Russian have provided evidence that young children may associate all imperfectives (either in the present or past tense) with incomplete events (Kiebzak-Mandera et al. 1997; Protassova 1997; Ceytlin 2000; Kiebzak-Mandera 2000; Gagarina 2003). It should be noted, that in principle the design of the task allowed for the use of the past tense forms in responses (e. g. *X and Y painted yesterday. Can you tell me what they do/are doing (in the picture) today? – Today they also painted*). In such cases the experimenter usually insisted on a different response by stressing the present tense context with an adverbial '*now*' instead of '*today*', but in Russian it is not ungrammatical to use the past tense form with the adverbial '*now*' (in the meaning that they have just been doing something), and some of the children failed to recover from the repetition of the stimuli even in such more accurate and enhanced present tense contexts.

In order to find out whether the differences in total correct performance were significant, and whether they hold across the groups, a mixed between-within subjects design ANOVA was performed with TEST as a within-subject variable with two levels (past tense condition vs. infinitive condition), and AGE and SEX as between-subject variables. We found a significant effect for TEST (Wilks' lambda = .692, $F(1, 98) = 43.6$, $p < .0005$) and for interaction TEST×AGE (Wilks' Lambda = .854, $F(3, 98) = 5.575$, $p = .001$). Among the between-subjects variables, the results revealed a statistically significant effect for AGE ($F(3, 98) = 26.359$, $p \leq .0005$). Post hoc comparisons using the Tukey HSD test have shown that 4-year-old Russian children differed significantly from 6- and 8-year-old children on total correct performance ($p \leq .0005$), while the difference between Russian 6-year-olds and 8-year-olds turned out to be not statistically significant ($p = .254$). The L2 learners' total correct performance was not significantly different from Russian 4-year-olds ($p = .191$), but was significantly different from 6- and 8-year-old Russian children ($p \leq .0005$).

[4.2] *Correct stem recognition rates for different verb classes*

Further, we compared correct recognition by verb type across the subject groups. Means for correct stem recognition for each class in two tests and by subject groups are illustrated in Figure 2 on the next page. A mixed between-within subjects design ANOVA was performed with two within-subject variables — CLASS (-AJ-, -I-, -OVA-, -A-) and TEST (past condition vs. infinitive condition), and two between-subjects variables — AGE (4 y.o. children, 6 y.o. children, 8 y.o. children, L2 learners), and SEX (male vs. female). This statistical analysis revealed the main effects for CLASS (Wilks' Lambda = .204, $F(3, 96) = 124.736$, $p \leq .0005$) and TEST (Wilks' Lambda = .712, $F(1, 98) = 39.631$, $p \leq .0005$), and several interaction effects: CLASS×AGE (Wilks' Lambda = .376, $F(9, 234) = 12.865$, $p \leq .0005$), TEST×AGE (Wilks' Lambda = .844, $F(3, 98) = 6.046$, $p = .001$), CLASS×TEST (Wilks' Lambda = .816, $F(3, 96) = 7.202$, $p \leq .0005$). These results suggest that

the subject groups recognised the four verb classes differently, and there were also significant differences in correct stem recognitions depending on the test condition (past vs. infinitive). A general trend we observe is that the verbal classes were recognised better in the test with stimuli in the infinitive than in the past tense form, with the exception of correct stem recognitions for the -AJ- class by L2 learners.

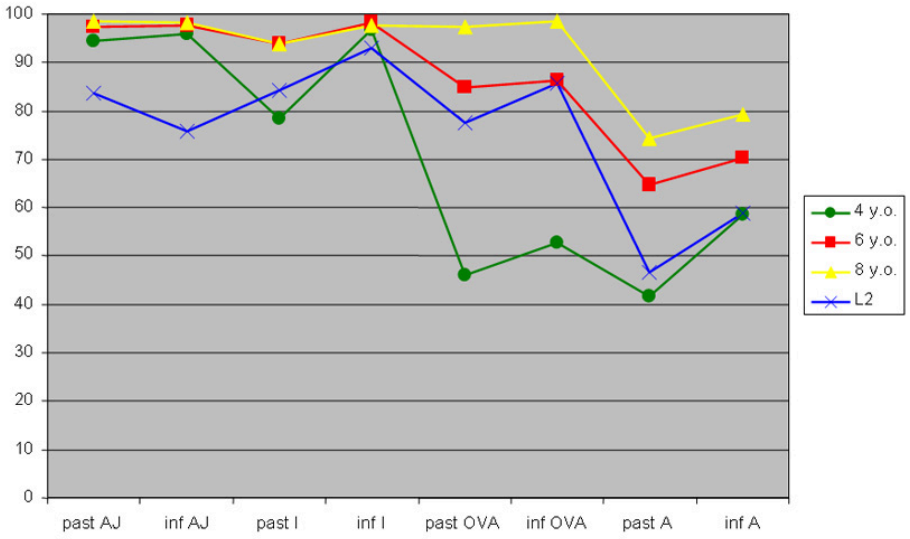


FIGURE 2: Correct stem recognition of different classes real verbs in two Russian tests by subject groups

As Figure 2 shows, all groups of Russian children in both tests recognised the -AJ- class close to 100% correct, and their means are much higher than that for L2 learners. These results suggest that by the age of four the Russian children have mastered the -AJ- class, while L2 learners lag significantly behind in their correct recognition of this class. Post hoc comparisons reveal that L2 learners' means of correct recognition for this class is significantly lower than the means for Russian children ($p \leq .0005$ for all pairs).

For the -I- class the means for correct stem recognition are higher in the infinitive condition than in the past tense condition, i.e. there is a significant effect for TEST (Wilks' Lambda = .724, $F(1, 102) = 38.945, p \leq .0005$) as well as interaction effect TEST×AGE (Wilks' Lambda = .847, $F(3, 102) = 6.16, p = .001$). 6- and 8-year-old Russian children recognised the -I- verbs better than 4-year-old children and L2 learners, especially in the past tense condition. Post hoc comparisons confirm this finding: the difference between 4-year-old children and Norwegian L2 learners of Russian is not statistically significant ($p = .973$), whereas these groups' performance on -I-class is significantly lower than that of 6- and 8-year-olds. Nevertheless, L2 learners and 4-year-olds' rates of stem recognitions are also relatively high on I-class.

Thus far, the results on correct stem recognition of the two most frequent classes, the -AJ- and -I- classes, show that all subject groups have high rates of stem recognition of these stems. The -AJ- class is recognised slightly better than the -I- class, especially in the past tense test condition and by 4-year-old children. On the other hand, L2 learners have a higher percentage of correct recognitions for the -I- class in the infinitive test condition than for the -AJ- class, which can probably be explained by their stronger reliance on the potential morphological cue — the verbs of the -I- class have a different thematic vowel (*-i-*), than all other stimuli in the test, which all end in *-at/-ali* and thus can represent a competition of the -AJ-, -A- and -OVA- inflectional patterns.

We can see more variation between the groups in the results on the less frequent -OVA- and -A-classes. As Figure 2 on the preceding page illustrates, correct stem recognition of the -OVA- class is distinctly lower for 4-year-old Russian children than for all other subject groups, and this difference is statistically significant ($p \leq .0005$ for all groups), while the difference between 6- and 8-year-old children does not reach statistic significance (6 y. o.–8 y. o. $p = .127$). L2 learners performed on -OVA- verbs close to 6-year-old Russian children, the difference between their means was not statistically significant ($p = .886$), but L2 learners' performance was significantly lower on the -OVA- class than that by Russian 8-year-olds ($p = .031$).

These findings suggest that by the age of four, Russian children have not yet acquired the -OVA- pattern. Presumably, this inflectional pattern is acquired later, between the ages of four and six, because already at the age of six we observe a drastic increase in correct performance on this class. These findings are consistent with the results of the earlier studies of acquisition of Russian verbal morphology by children, which report that "there is an abrupt jump in the rates of stem recognition for the -OVA- class between the ages of 4 and 5" (Gor & Chernigovskaya 2004, 6). Concerning L2 learners, as expected due to low type frequency of this pattern in the L2 input, they were worse at recognising the -OVA- pattern than the most frequent -AJ- and -I- patterns in the past test condition; however in the infinitive test condition the rates of correct stem recognition of the -OVA- pattern (86%) are higher than correct stem recognition of the frequent -AJ- class (76%). This fact can probably be explained by the influence of a clear morphological cue (suffix *-ova-* in the infinitive and past tense forms) that characterises this inflectional class, which was also made prominent in the L2 instruction, as discussed earlier.

Figure 2 illustrates clearly that the rates of correct stem recognition of the least frequent -A- class are noticeably lower for all subject groups compared to their correct stem recognition rates of the other three classes we considered before. Even 8-year-old Russian children, who seem to be the most proficient group, performed below 80% correct on this class. These results suggest that the -A-

class, having the lowest type frequency and the high degree of morphological complexity, has not yet been completely acquired by this age by monolingual Russian children. As seen in the figure, Russian 4-year-old children and L2 learners have a remarkably low rate of correct stem recognition of this class, while the recognition rate of 6-year-old and 8-year-old Russian children gradually becomes better. Post hoc comparisons show that the difference between 4-year-olds and L2 learners is not statistically significant ($p = .944$), and that these groups differ significantly from 6- and 8-year-olds (4 y. o.– 6 y.o $p = .001$; 4 y. o.–8 y. o. $p \leq .0005$; L2–6 y. o. $p = .010$, L2–8 y. o. $p \leq .0005$). These results demonstrate, that the class with the lowest type frequency and high morphological complexity is acquired late and is recognised correctly less frequently by all groups of subjects.

The results on correct stem recognition suggest that there is a similarity between L1 and L2 acquisition and processing of verbal morphology: there is a correlation between the type frequency of the pattern and the rates of correct stem recognition for all groups. However, L2 learners do not match any of the child age groups completely, which can be explained by the influence of formal instruction, as well as the differences found in the distribution of type frequencies in the L2 input.

[4.3] *Generalisations in response to nonce verbs*

In the previous section we have demonstrated that the rates of correct recognition of a particular class correlate strongly with the type frequency of this particular inflectional pattern. Although this can be an indication of the importance of type frequency in the order of acquisition of different inflectional patterns, to make this point stronger we need to analyse which patterns the subjects preferred in their responses to nonce verbs. In order to conjugate a nonce verb, the subjects are supposed to apply a particular abstract schema which has emerged on the base of the real verbal items from the subjects' previous linguistic experience. The actual usage of different verbs leads to formation of several schemas. Thus, the subjects are faced with a problem of choice between several schemas which may be available. According to Bybee (1995, 430), "the likelihood of the schema being extended to new items is directly dependent upon two factors: (i) the defining properties of the schema and (ii) its strength, the latter property being derivable from the number of items that reinforce the schema", "the higher the type frequency of the pattern described in the schema, the greater are its chances of applying to new items".

Potentially, any of the existing schemas could be applied to nonce items provided that the nonce item meets the properties of the schema. However, the schemas which were applied most frequently by the majority of the subjects were the schemas representing the four inflectional classes under consideration. For

this reason we decided to focus on these four generalisation patterns and exclude from our analysis other types of responses.[6] The four generalisation patterns we consider here are described in Table 6 on the following page.

As described in section [3], half of the Russian tests consisted of nonce verbs (i. e. 40 nonce verbs in each of the tests). Nonce verbs were created from real verbs, and each of them was classified as belonging to one of the four classes (-AJ-, -A-, -I-, -OVA-): e.g. a nonce stimuli *kisat* matched the real verb *pisat* 'write' and thus was classified as a nonce verb in the -A- class). However, as long as the final segment of verbs belonging to the -AJ-, -A-, and -OVA- classes is the same (*-at/-ali*) in the stimuli, potentially any of these three inflectional patterns could be applied to nonce verbs. It is less likely that any of these three patterns with the thematic vowel *-a-* would be applied to the nonce verbs which have a thematic vowel *-i-* in the stem (i. e. nonce verbs from the -I- class). The schema representing the -I- class has different properties, and it is thus not expected that it will be applied to more than 25% of the nonce items.

If type frequency influences the choice of one of the schemas for application to nonce stimuli, we can expect that the frequent patterns are used for generalisations more often than the patterns that have low type frequency. Thus, we can expect a general trend that the -AJ- and -I- patterns are generalised frequently by all subject groups; whereas the least frequent -A- pattern will not be frequently applied to nonce verbs.

The percentages of generalisations of the -AJ-, -A-, -I- and -OVA- inflectional patterns are illustrated in two figures below. Figure 3 on page 301 illustrates the proportions of generalisations used for the nonce verbs in the test with stimuli in the past tense, whereas Figure 4 on page 301 illustrates the proportions of generalisations in the test with stimuli in the infinitive.

The figures show a general tendency of preference for GEN>-AJ- in response to nonce verbs, however the relative proportions of -AJ- generalisations differed across groups and depending on the test. To estimate effects of within- and between-subjects variables a mixed between-within subjects ANOVA was performed with TEST (past vs. infinitive) and GENERALISATION TYPE (AJ, A, I, OVA) as within-subject variables, and AGE (4 y. o., 6 y. o., 8 y. o. and L2 learners) and SEX (male vs. female) as between-subjects variables. The statistical analysis revealed statistically significant main effects for TEST (Wilks' Lambda = .699, $F(1, 98) = 42.112$, $p \leq .0005$), GENERALISATION TYPE (Wilks' Lambda = .196, $F(3, 96) = 131.158$, $p \leq .0005$) and AGE ($F(3, 98)$ = 8.812, $p \leq .0005$), as well as interaction effects for TEST×AGE (Wilks' Lambda = .789, $F(3, 98) = 8.715$, $p \leq .0005$), GENERALISATION TYPE×AGE (Wilks' Lambda = .624, $F(9, 234) = 5.556$, $p \leq .0005$),

[6] These other response types to nonce verbs included not only the applications of other inflectional patterns than those we consider here, but also responses when the subjects replied in the past form or infinitive, which we considered as "repetition of the stimuli".

Generalisation type	Description	Examples (inf/past plural – present 1sg/3pl)
GEN>-AJ-	The form is generated by adding *-j-* to a stem with a thematic vowel *-a-*, before adding the ending.	*kis**a**-t/kis**a**-li* – *kis**aj**-u/kis**aj**-ut* *mylov**a**-t/mylov**a**-li* – *mylov**aj**-u/mylov**aj**-ut*
GEN>-I-	The form is generated by adding the ending directly to the stem (either with or without consonant mutations) with a thematic vowel *-i-* (the vowel is eliminated before the endings).	*nad**i**-t/nad**i**-li* – *na**ž**-u/na**d**-at*
GEN>-A-	The form is generated by adding the ending directly to the stem (either with or without consonant mutations) with a thematic vowel *-a-* (the vowel is eliminated before the endings).	*kis**a**-t/kis**a**-li* – *ki**s/š**-u/ki**s/š**-ut* *okož**a**-t/okož**a**-li* – *okož-u/okož-ut* *tintov**a**-t/tintov**a**-li* – *tinto**vl**-u/tinto**v**-ut*
GEN>-OVA-	The form is generated using the suffix *-uj-* before the ending. The stimuli do not necessarily have the *-ova-* suffix in the infinitive/past tense stem.	*kis**a**-t/kis**a**-li* – *kis**uj**-u/kis**uj**-ut* *tint**ova**-t/tint**ova**-li* – *tintov**uj**-u/tintov**uj**-ut* *tint**ova**-t/tint**ova**-li* – *tint**uj**-u/tint**uj**-ut* *nad**i**-t/nad**i**-li* – *nad**uj**-u/nad**uj**-ut*

TABLE 6: Generalisation types

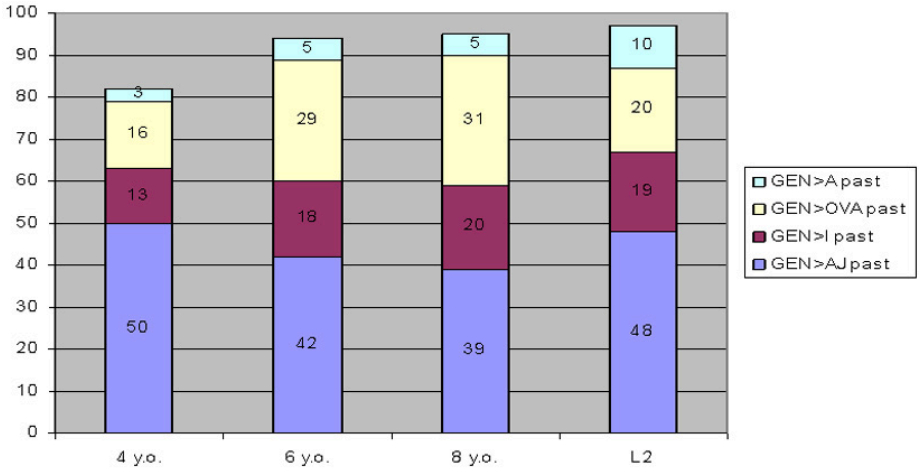


FIGURE 3: Generalisations in responses to nonce verbs (past test condition)

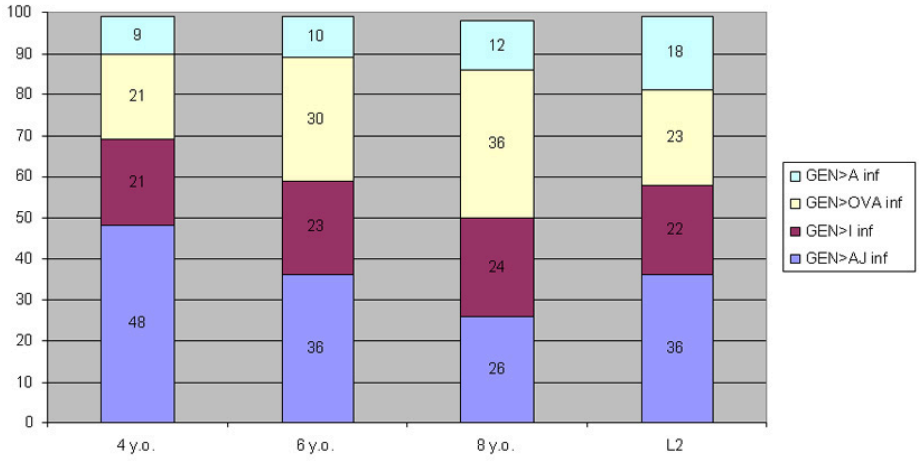


FIGURE 4: Generalisations in responses to nonce verbs (infinitive test condition)

and TEST×GENERALISATION TYPE (Wilks' Lambda = .590, $F(3, 96) = 22.197$, $p \leq .0005$). These results suggest that the differences in generalisation types found between the subject groups are significant, and that the generalisation types used differed depending on the test condition, whereas the subjects' responses did not depend on their sex.

Although the -AJ- generalisations are the most frequent for all subject groups, we can observe a tendency that the use of this pattern is higher in the past tense than in the infinitive condition, and with age the percentage of -AJ- generalisations decreases. Post hoc comparisons using the Tukey HSD test show significant difference between 4-year-old children and 6- and 8-year old children (4 y. o.–6 y. o. $p = .013$, 4 y. o.–8 y. o. $p \leq .0005$). L2 learners generalised the -AJ- pattern in their responses to nonce verbs to the same extent as Russian 4-year-old and 6-year-old children (L2 vs. 4.y. o. $p = .173$; L2 vs. 6 y. o. $p = .804$), but differently

from Russian 8-year-olds ($p = .042$).

As seen in Figure 3 on the previous page and Figure 4 on the preceding page, the use of the -AJ- pattern exceeds 25%, which means that this productive pattern was applied not only to nonce stimuli classified as belonging to the -AJ- class, but also to other nonce stimuli. Most frequently, it was applied to a large part of the -A- stimuli, as the rate for -A- generalisations is extremely low, but it could also be applied to -OVA- nonce stimuli, especially by 4-year-old children who have not yet acquired the -OVA- inflectional pattern. This explains the higher use of the -AJ- generalisation by 4-year-old children.

The -I-class pattern is also applied to nonce verbs quite frequently by all subjects. Although the proportions of -I- generalisations are not very high in the figures, we should bear in mind that the schema representing this inflectional pattern has restrictions in terms of the thematic vowel, so potentially it could only be applied to 1/4 of all nonce verbs included in the test. The figure illustrates that the application of the -I- pattern is close to 25% for all subject groups. Post hoc comparisons reveal that the differences in the use of -I- generalisations were significant for 4-year-old children as compared to all other groups (4 y. o.–6 y. o. $p = .043$, 4 y. o.–8 y. o. $p = .002$, 4 y. o.–L2 $p = .025$).

This result may seem surprising in the light of the high type frequency of this pattern and also quite high rates of stem recognition of the -I- class which we reported on in the previous section. This can probably be explained by two facts:

(i) 4-year-old children tended to reply with repetition of the stimuli, either in the infinitive or the past tense form, more often than other subject groups (repetitions of the stimuli occurred at the rate of up to 10% for real and nonce -I- verbs together)

(ii) 4-year-old children tended also to overuse a general V+j strategy in their replies, so that when this strategy was applied to the verbs of the -I- class, the results were coded as generalisations of the -IJ- pattern, which could also have resulted in the lower rates of -I- generalisations.

L2 learners replied to most nonce -I- verbs with an appropriate inflectional pattern: the -I- generalisation rate is 19% in the past test, and 22% in the infinitive test. Post hoc comparisons have shown that these rates are significantly different from the rates of 4-year-old children ($p = .025$), but they do not reach significance as compared to -I- generalisations used by 6-year-old children (18%/23%) ($p = .987$) and by 8-year-old children (20%/24%) ($p = .792$).

The analysis of generalisations of the -OVA- inflectional pattern should also be related to our previous findings on correct stem recognitions. The results on generalisations used in responses to nonce verbs show in fact that 4-year-old children did apply this inflectional pattern in 16% and 21% in the past and infinitive

test conditions, respectively. However, they differ significantly in their use of this generalisation pattern from 6- and 8-year-olds (4 y. o.–6 y. o. $p = .001$, 4 y. o.–8 y. o. $p \leq .0005$), who tend to apply this pattern not only to -OVA- nonce verbs, but also to nonce verbs from other classes (because their rates exceed 25%).

L2 learners' productive use of the -OVA- pattern is not significantly different from that of 4-year-old Russian children ($p = .767$), but is significantly different from 6- and 8-year-old children (L2–6 y. o. $p = .032$, L2–8 y. o. $p = .001$). As we already discussed above, the -OVA- class turned out to match the -A- class in type frequency, and in contrast to the TL input frequencies, the -OVA- class turned out to have low type frequency in the L2 input. On the other hand, the L2 learners could have acquired this low frequent pattern due to prominence and enhancement of this pattern in the L2 instruction — this was the only class which was explicitly presented as a salient pattern. This might have helped the L2 learners to form proper generalisations and a schema, which they seem to apply to nonce items in the test.

The figures demonstrate clearly that the -A- pattern was the least preferred generalisation pattern in all subject groups, which can be explained by combination of low type frequency of the -A- class and high degree of morphological complexity. We can also observe that the percentage of -A- generalisations increases gradually with age for Russian children. L2 learners have the highest percentage of generalisations of the -A- pattern in both tests (10% in the past test, 18% in the infinitive test), and their results are significantly different from 4- and 6-year-old children (L2–4 y. o. $p = .003$, L2–6 y. o. $p = .021$), but similar to 8-year-old children ($p = .107$). This may be surprising taking into consideration that the -A- class has low type frequency in the L2 input. One possible explanation for this can be ascribed to the influence of formal instruction and focus on form. As mentioned earlier, a special focus in the L2 instruction is given to verbal forms that have "irregular" conjugational patterns, i. e. the patterns which have consonant mutations and stress shifts. This focus on irregularities could have made this low frequency pattern with consonant mutations and stress shifts more enhanced in the L2 input, so that L2 learners were inclined to pay more attention to this pattern and be aware of the morphological complexities it implies. Maybe therefore they were deliberately trying to apply this pattern to nonce verbs more frequently than other subject groups.

Our analysis of generalisation patterns suggests that type frequency is important in the subjects' choice of the schema to be applied to nonce verbs: the most frequent -AJ- pattern is preferred for generalisations. However, the correlation is less evident due to the possible influence of other factors. L2 learners showed higher percentages of use of low type frequency -A- and -OVA- patterns to nonce verbs than was expected. We have shown that these findings can be explained by the influence of formal instruction and focus on form, which is the case in second

language acquisition in a classroom setting.

[4.4] *Influence of token frequency*

Thus far, we have discussed the effects of type frequency on correct stem recognition and use of different patterns for generalisations in responses to nonce verbs. As pointed out by the theoretical framework we take as a departure point in this study, token frequency is also an important factor which can play a role in acquisition. Within the usage-based account we expect that token frequency effects can be found in all verb types in both L1 and L2 acquisition. We consider separately token frequency effects in rates of correct stem recognition and in rates of correctly produced forms.

As was described in the methodology section, each inflectional class in the test was represented by 5 verbs having high token frequency and 5 verbs having low token frequency. Thus, we will look here at whether the subjects recognised high token frequency verbs better than low token frequency verbs, and whether this tendency was similar for all subject groups and for all types of verbs. Correct stem recognition rates for high token frequency and low token frequency verbs in the two test conditions are illustrated in the following figures: Figure 5 for 4-year-old Russian children, Figure 6 on the facing page for 6-year-old Russian children, Figure 7 on page 306 for 8-year-old Russian children, and Figure 8 on page 306 for L2 learners. With several exceptions, where the recognition rates for low and high token frequency stimuli are close to equal, we can observe a tendency that the high token frequency verbs were recognised better than the low token frequency verbs. This trend holds across subject groups, in both tests and mostly for all verb classes.

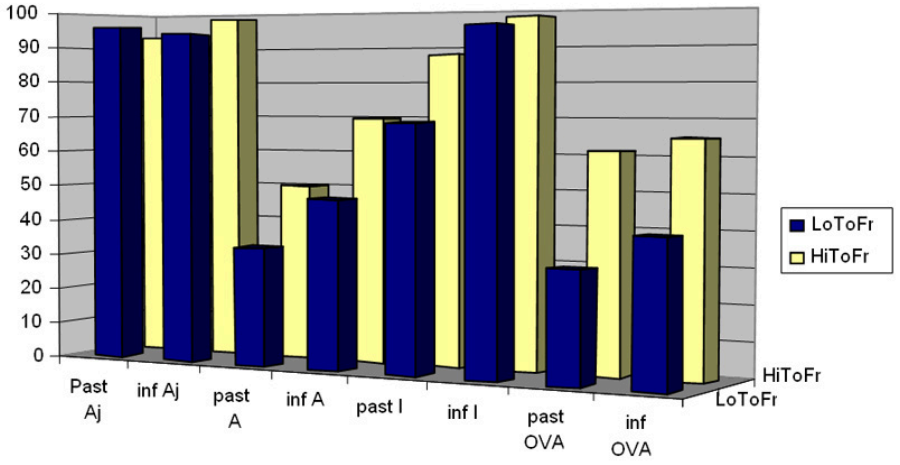


FIGURE 5: Token frequency effects in correct stem recognition rates for 4-year-olds

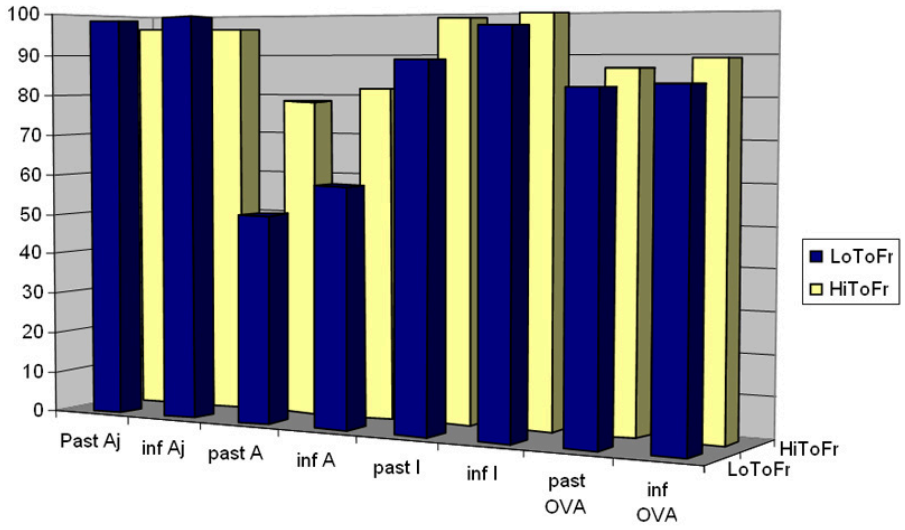


FIGURE 6: Token frequency effects in correct stem recognition rates for 6-year-olds

To check whether the differences visible in the figures are statistically significant, a series of paired-samples t-tests were performed on the data, in which we compared 8 pairs for every subject group (4 classes $\times$ 2 tests, high token frequency vs. low token frequency items). This statistical analysis has shown that not all the differences between high token frequency items and low token frequency items reach the level of statistical significance. Token frequency effects did not reach statistical significance for the -AJ- class in either test condition for all Russian children nor for L2 learners in the past test condition. Token frequency effects again did not reach statistical significance for the -I- class in the infinitive test condition for all groups Russian children, nor in either test condition for L2 learners; nor for the -OVA- class for children at the age 6 and 8 in either test conditions.

At first sight these findings can provide evidence against token frequency effects in some classes, but a closer look reveals an interesting tendency: token frequency effects appear to be non-significant for the inflectional classes on which the subjects perform close to asymptote, i. e. the better in general the subjects perform on a particular class, the less evident are token frequency effects and the subjects tend to perform equally well on both high and low token frequency verbs. A similar finding has been reported in the previous studies which investigated token frequency by regularity interaction, in particular, it was claimed that token frequency effects are evident for all types of inflections (regular and irregular), but frequency effects can disappear and become less visible as the performance is close to asymptote (Ellis 1998). This is the case for our data: relating our findings on token frequency effects to the results on correct stem recognition rates for different classes across subject groups, we can notice that token frequency

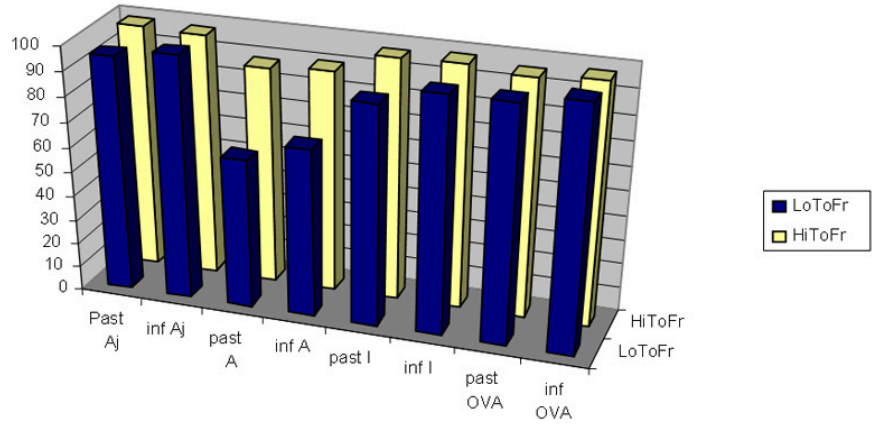


FIGURE 7: Token frequency effects in correct stem recognition rates for 8-year-olds

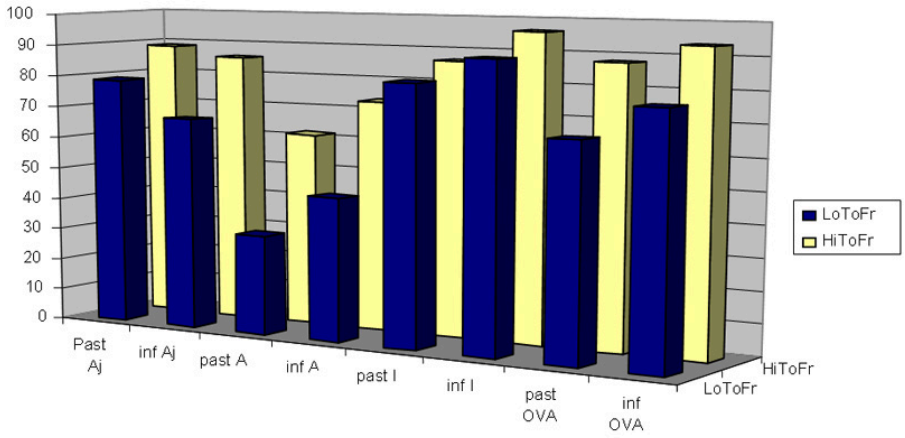


FIGURE 8: Token frequency effects in correct stem recognition rates for L2 learners

effects turned out to be statistically insignificant when the rates of correct stem recognition were higher than 80%.

Thus, our findings do not contradict the predictions of the single-mechanism account that token frequency influences the acquisition of all verbal classes.

We predicted as well that token frequency will influence percentages of correctly produced forms in both L1 and L2 acquisition. The figures below illustrate correctly produced forms for each inflectional class in two test conditions for high and low token frequency verbs: Figure 9 on the next page represents the data for 4-year-old Russian children, Figure 10 on the facing page for 6-year-old Russian children, Figure 11 on page 308 for Russian 8-year-olds, and Figure 12 on page 308 for L2 learners. In the same vein as for the rates of correct stem recognition, the figures show that there is a tendency that correct forms were produced more often for high token frequency than for low token frequency items.

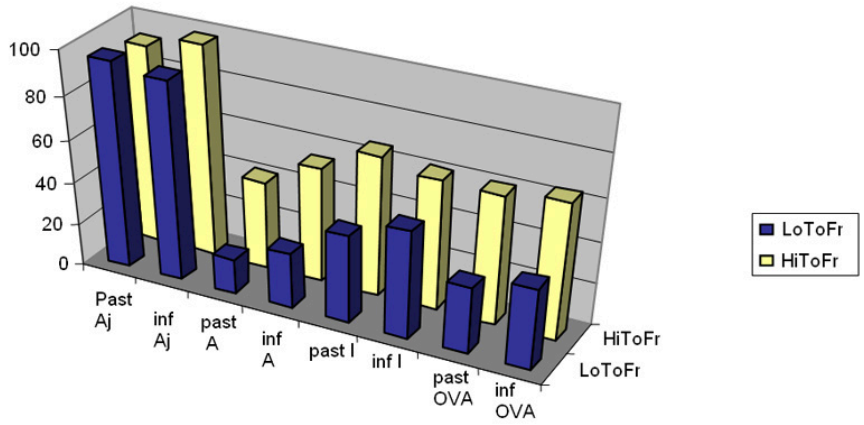


FIGURE 9: Token frequency effects in correctly produced forms for 4-year-olds

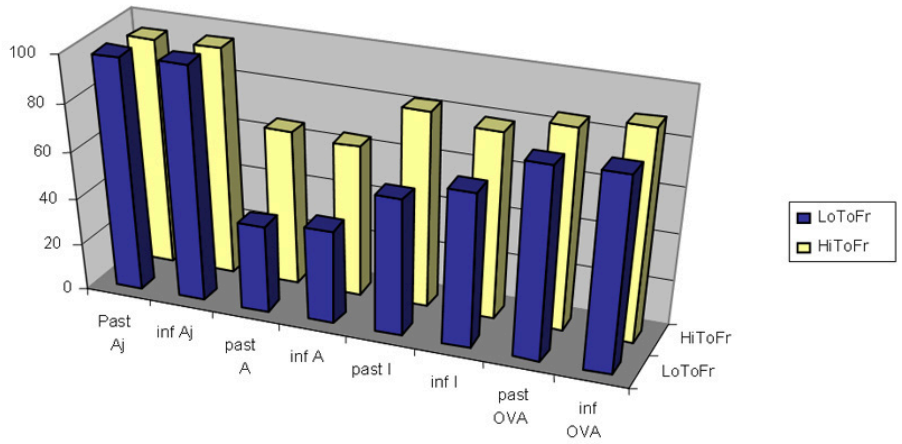


FIGURE 10: Token frequency effects in correctly produced forms for 6-year-olds

When these effects were tested for statistical significance, the statistical analysis has shown that token frequency effects did not reach the level of statistical significance for the -AJ- class in either testing condition for any child groups nor for L2 learners in the past test condition. Token frequency effects were also not statistically significant for the -OVA- class in either test conditions for 6- and 8-year-old Russian children. For all other classes the differences in correctly produced forms from high token frequency and low token frequency stimuli were statistically significant.

Here again we observe that token frequency effects are not evident for some classes. Similarly to our observations made for token frequency effects in correct stem recognitions, the effects are not significant for the -AJ- and -OVA- classes for those subject groups, who perform on these classes close to asymptote. However, these results are different for the -I- class: whereas token frequency effects for stem recognitions were not significant for the -I- verbs in groups who per-

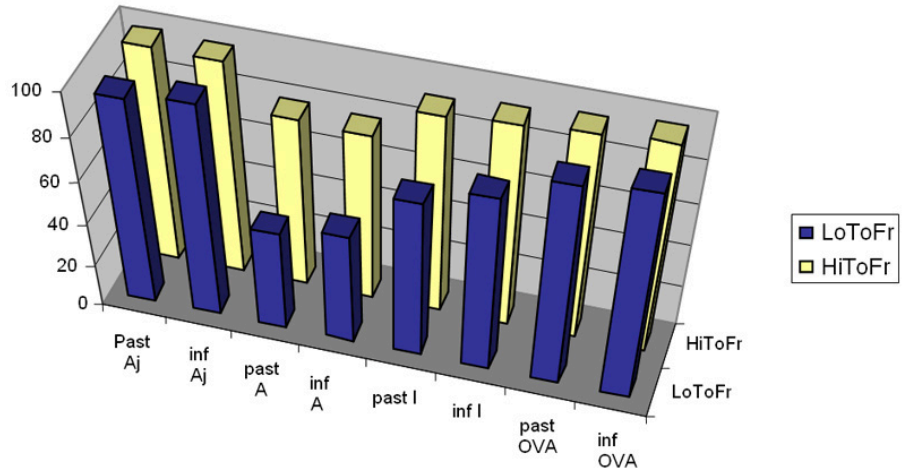


FIGURE 11: Token frequency effects in correctly produced forms for 8-year-olds

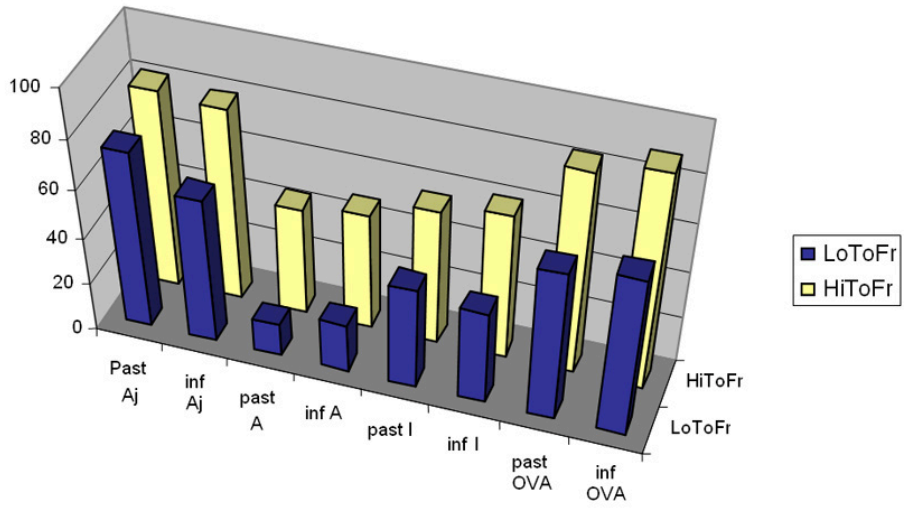


FIGURE 12: Token frequency effects in correctly produced forms for L2 learners

formed over 80% correct on this class, token frequency effects in correctly produced forms are statistically significant for the -I- class in all subject groups.

This phenomenon can be explained by the influence of morphological complexity. When the performance is estimated in terms of correctly produced forms, morphological complexity comes into play. The -I- and -A- classes are characterised by high morphological complexity (consonant mutations and stress shifts). For this reason, the performance measured in terms of correctly produced forms for these morphologically complex classes is lower than the performance measured in terms of correct stem recognitions, which only registers whether the inflectional class was recognised correctly, but disregards whether the subjects applied all consonant mutations and stress shifts appropriately.

Thus, these results accord with the assumption expressed above that token frequency influences the acquisition of all inflectional classes, but these effects may be less evident and decrease with a higher proficiency.

[5] DISCUSSION AND CONCLUSION

In the present study our primary goal was to compare the data on L1 and L2 acquisition of Russian verbal morphology. Studies on L2 morphological processing present an interesting field of research in attempts to shed more light on the question of what role input factors play in morphological processing because L2 learners are potentially exposed to less input than children acquiring this language as an L1. In addition, the L2 classroom learning context differs in several respects from the language environment in which children acquire a language.

Our study has demonstrated that there are some similarities in the acquisition and processing of verbal morphology by Russian children and adult L2 learners. Referring back to different theoretical positions with regard to L2 morphological processing, the fact that L1 and L2 processing of morphology is similar, might support the Shallow Structure Hypothesis (Clahsen & Felser 2006b,a) and the usage-based approach (Ellis 2002, 2003, 2008).

The similarities that we found regard the role of type and token frequencies in the acquisition of different inflectional patterns and verbs. Throughout our analysis, we have shown that the classes that are characterised by high type frequency in the input are acquired earlier by children and are performed better by L2 learners. The effect of type frequency has also been demonstrated in generalisations the subjects made in response to nonce stimuli, and we observed that morphologically complex patterns which can be called "irregular" were not an exception. This contradicts the main assumption of the Dual Mechanism Account, which assumes that only regular patterns can be generalised. Frequent patterns were easily applied to nonce verbs whenever the stimuli met the requirements of the properties of a particular schema, supporting thus the assumptions made by the usage-based account. The pattern with the highest type frequency was preferred for generalisations by children and L2 learners, however, certain developmental tendencies appeared with age: as children acquired more patterns, they started applying them more frequently to nonce stimuli as well.

Although the L2 learners showed a similar trend to Russian children in their preference of the most frequent inflectional pattern in generalisations, they applied less frequent patterns more often than could have been expected based on our type frequency analysis in the L2 input. This might be a result of formal instruction which manifested itself in a stronger reliance on a clear morphological marker in choosing the -OVA- pattern for generalisation, and also in L2 learners' higher awareness of conjugational patterns with irregularities, and thus higher percentage of use of a rather infrequent -A- pattern for nonce stimuli.

Our results demonstrate that token frequency influences both rates of correct stem recognitions and correctly produced forms. As was reported by the previous studies (Ellis 1998), the token frequency effects are most influential at the beginning stages of acquisition and diminish with more learning. In accordance with this, we observed that token frequency effects were not significant for those classes and test conditions, in which L2 learners performed close to 100%. Our data have shown that frequency effects exist in both "regular" and "irregular" classes. This particular finding contradicts the Dual Mechanism account of morphological processing predicting frequency effects for irregular patterns but not for regular ones, and supports the predictions made by the Single Mechanism account. This phenomenon was observed in both children and adult L2 learners.

Consequently, a following conclusion may be drawn from our data. In spite of the fact that L2 learners did not match completely any of the child groups in our data, we suggest that at least one of the key underlying principles of morphology acquisition is the same for L1 and L2 context: the process depends on the type frequency. However, for L2 learners other factors may also come into play and we suggest that this might be a result of different language learning environments and different input characteristics. High type frequency of a particular pattern in the L2 input plays an important role in successful acquisition of this pattern, both in terms of correct stem recognition and in its applicability to nonce items. However, low type frequency does not necessarily mean that L2 learners will not acquire the pattern: provided that the pattern has some clear morphological cues and that attention is paid to it in the instruction, it is likely that L2 learners will be able to recognise the pattern, and also apply it to nonce items which meet the requirements on its properties.

The present study demonstrates that the usage-based account of language acquisition is a promising theoretical framework for the field of SLA research. It looks at L1 and L2 acquisition from the same theoretical position and draws lines explaining similarities and differences between them without assuming that different learning principles underlie the L1 and L2 acquisition.

REFERENCES

Alanen, Riikka. 1995. Input enhancement and rule presentation in second language acquisition. In Richard Schmidt (ed.), *Attention and awareness in foreign language learning and teaching*, 259–302. Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.

Bach Nielsen, Erik. 2003. *Grammatika: russisk grammatikk* Gyldendals gymnasiale grammatikker. København: Gyldendal.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14. 150–177.

Bjerkeng, Marit & Tatiana Bräger. 2002. *Sosedi 2: begynnerkurs i russisk*. Bergen: Fagbokforlaget.

Bjerkeng, Marit, Tatiana Bräger & Trond Gunnar Nordenstam. 2000. *Sosedi 1: begynnerkurs i russisk*. Bergen: Fagbokforlaget.

Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5). 425–55.

Bybee, Joan & Paul Hopper. 2001. *Frequency and the emergence of linguistic structure* (Typological studies in language 45). Amsterdam: John Benjamins.

Bybee, Joan & Dan Slobin. 1982. Rules and schemas in the development and use of the English past tense. *Language* 58(2). 265–289.

Celce-Murcia, Marianne. 1991. Grammar pedagogy in second and foreign language teaching. *TESOL Quarterly* 25(3). 459–480.

Ceytlin, S. N. 2000. *Jazyk i rebenok: lingvistika detskoj rechi. (Language and child: linguistics of the child speech)*. Moskva: Vlados.

Chavronina, S. A. & A. I. Širočenskaja. 1999. *Russian in exercises*. Moscow: Russky Yazyk 9th edn.

Clahsen, Harald. 1997. The representations of participles in the German mental lexicon: Evidence for the dual-mechanism model. *Yearbook of morphology* 1996. 73–95.

Clahsen, Harald. 2006. Dual-mechanism morphology. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, vol. 4, 1–5. Elsevier.

Clahsen, Harald & Claudia Felser. 2006a. Continuity and shallow structures in language processing. *Applied Psycholinguistics* 27(1). 107–126.

Clahsen, Harald & Claudia Felser. 2006b. Grammatical processing in language learners. *Applied Psycholinguistics* 27(1). 3–42.

Cornyn, William S. 1948. On the classification of Russian verbs. *Language* 24(1). 64–75.

DeKeyser, Robert M. 1997. Beyond explicit rule learning. *Studies in Second Language Acquisition* 19(2). 195–221.

Dörnyei, Zoltán. 2005. *The psychology of the language learner: individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates. Zoltan Dörnyeifig.

Dörnyei, Zoltán & Peter Skehan. 2002. *Individual differences in second language acquisition*. Amsterdam: John Benjamins.

Doughty, Catherine & Jessica Williams. 1998. Pedagogical choices in focus on form. In Catherine Doughty & Jessica Williams (eds.), *Focus on form in classroom second language acquisition*, 197–261. Cambridge: Cambridge University Press.

Eddington, David. 2002. Dissociation in Italian conjugations: a single-route account. *Brain and Language* 81(1-3). 291–302.

Ellis, Nick C. (ed.). 1994. *Implicit and explicit learning of languages*. San Diego: Academic Press.

Ellis, Nick C. 1998. Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes* 13(2). 307–336.

Ellis, Nick C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24(2). 143–188.

Ellis, Nick C. 2003. Constructions, chunking, and connectionism: the emergence of second language structure. In Catherine Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*, 63–103. Oxford: Blackwell Publishing.

Ellis, Nick C. 2005. At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition* 27(2). 305–352.

Ellis, Nick C. 2008. Usage-based and form-focused language acquisition. In Peter Robinson & Nick C. Ellis (eds.), *Handbook of cognitive linguistics and second language acquisition*, 372–405. New York - London: Routledge.

Ellis, Rod. 1990. *Instructed second language acquisition: learning in the classroom* Applied language studies. Oxford: Blackwell.

Eubank, Lynn (ed.). 1991. *Point counterpoint: universal grammar in the second language* (Language acquisition & language disorders 3). Amsterdam: John Benjamins.

Flynn, Suzanne. 1996. A parameter-setting approach to SLA. In William C. Ritchie & Tej K. Bhatia (eds.), *The handbook of second language acquisition*, 121–158. San Diego: Academic Press.

Gagarina, Natalia. 2002. Productivity of the verbal inflectional classes (the case of Russian). In Katarzyna Dziubalska-Kolaczuk & Jarek Weckwerth (eds.), *Future challenges for natural linguistics* (Lincom studies in theoretical linguistics 30), 153–164. Wien: Lincom Europa.

Gagarina, Natalia. 2003. The early verb development and demarcation of stages in three Russian-speaking children. In Dagmar Bittner, Wolfgang U. Dressler, Marianne Kilani-Schoch & Peter Jordens (eds.), *Development of verb inflection in first language acquisition: a cross-linguistic perspective* (Studies on Language Acquisition 21), 131–169. Berlin, New York: Mouton de Gruyter.

Gass, Susan. 1996. Second language acquisition and linguistic theory: The role of language transfer. In William S. Ritchie & Tej Bhatia (eds.), *Handbook of second language acquisition*, 317–345. San Diego, Calif.: Academic Press.

Gathercole, Virginia C. 2007. Miami and North Wales, so far and yet so near: Constructivist account of morpho-syntactic development in bilingual children. *International Journal of Bilingual Education and Bilingualism* 10(3). 224–247.

Gilmore, Alex. 2007. Authentic materials and authenticity in foreign language learning. *Language Teaching* 40. 97–118.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Gor, Kira. 2004. The rules and probabilities model of native and second language morphological processing. In L. Verbitskaya & Tatiana Chernigovskaya (eds.), *Teoreticheskie problemy jazykoznanija. sbornik statej k 140-letiju kafedry obschego jazykoznanija (Theoretical problems in linguistics. Papers dedicated to 140th anniversary of the Department of General Linguistics, St. Petersburg State University)*, 51–75. St. Petersburg: Philological Faculty of St. Petersburg State University Press.

Gor, Kira. 2007. Experimental study of first and second language morphological processing. In Monica Gonzalez-Marques, Irene Mittelberg, Seana Coulson & Michael J. Spivey (eds.), *Methods in cognitive linguistics*, 367–398. Amsterdam: John Benjamins.

Gor, Kira & Tatiana Chernigovskaya. 2004. Generation of complex verbal morphology in first and second language acquisition: Evidence from Russian. *Nordlyd. Tromsø University Working Papers on Language and Linguistics* 31(6). 819–833. `http://www.ub.uit.no/munin/nordlyd/viewissue.php?id=6`.

Gor, Kira & Tatiana Chernigovskaya. 2005. Formal instruction and the acquisition of verbal morphology. In Alex Housen & Michel Pierrard (eds.), *Investigations in instructed second language acquisition*, 103–139. New York: Mouton De Gruyter.

Hahne, Anja, Jutta L. Mueller & Harald Clahsen. 2006. Morphological processing in a second language: Behavioral and event-related brain potential evidence for storage and decomposition. *Journal of cognitive neuroscience* 18(1). 121–134.

Hyltenstam, Kenneth. 1992. Non-native features of near-native speakers. On the ultimate attainment of childhood L2 learners. In R. Harris (ed.), *Cognitive processing in bilinguals*, 351–368. Amsterdam: Elsevier Science Publishers.

Hyltenstam, Kenneth & Niclas Abrahamsson. 2000. Who can become native-like in a second language? All, some, or none? On the maturational constraints controversy in second language acquisition. *Studia linguistica* 54(2). 150–166.

Jakobson, Roman. 1948. Russian conjugation. *Word* 4(3). 155–167.

Kemmer, Suzanne & Michael Barlow (eds.). 2000. *Usage-based models of language*. Stanford, Calif.: CSLI Publications.

Kiebzak-Mandera, Dorota. 2000. Formation of the verb system in Russian children. *Psychology of Language and Communication* 4. 27–46.

Kiebzak-Mandera, Dorota, Magdalena Smoczynska & Ekaterina Protassova. 1997. Acquisition of Russian verb morphology: The early stages. In Wolfgang U. Dressler (ed.), *Studies in pre- and protomorphology*, 101–114. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar*. Stanford, Calif.: Stanford University Press.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 1–63. Stanford: CSLI Publications.

Long, Michael. 1991. Focus on form: A design feature in language teaching methodology. In Kees De Bot, Claire Kramsch & Ralph B. Ginsberg (eds.), *Foreign language research in cross-cultural perspective* (Studies in bilingualism 2), 39–52. Amsterdam: John Benjamins.

Long, Michael H. & Peter Robinson. 1998. Focus on form. theory, research and practice. In Catherine Doughty & Jessica Williams (eds.), *Focus on form in classroom second language acquisition*, 15–41. Cambridge: Cambridge University Press.

MacWhinney, Brian. 1997. Implicit and explicit processes. *Studies in Second Language Acquisition* 19(2). 277–281.

MacWhinney, Brian & Jared Leinbach. 1991. Implementations are not conceptualizations: revising the verb learning model. *Cognition* 40(1–2). 121–157.

Marchman, Virginia & Elizabeth Bates. 1994. Continuity in lexical and morphological development: a test of the critical mass hypothesis. *Journal of Child Language* 21. 399–366.

Marcus, Gary. 1995. The acquisition of the English past tense in children and multilayered connectionist networks. *Cognition* 56(3). 271–279.

Marcus, Gary, Ursula Brinkmann, Harald Clahsen, Richard Wiese & Steven Pinker. 1995. German inflection: the exception that proves the rule. *Cognitive Psychology* 29(3). 189–256.

Marcus, Gary, Steven Pinker, Michael Ullman, Michelle Hollander, T. J. Rosen & Fei Xu. 1992. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 57(4). i+iii+v+vi+1–178.

Marslen-Wilson, William & Lorraine K. Tyler. 1998. Rules, representations, and the English past tense. *Trends in Cognitive Sciences* 2(11). 428–435.

Marslen-Wilson, William & Lorraine K. Tyler. 2005. The lexicon, grammar, and the past tense: Dissociation revisited. In Michael Tomasello & Dan Slobin (eds.), *Beyond nature-nurture: Essays in honor of elizabeth bates*, 263–280. Mahwah, NJ: Lawrence Erlbaum.

Masgoret, Anne-Marie & Robert C. Gardner. 2003. Attitudes, motivation, and second language learning: a meta-analysis of studies conducted by Gardner and associates. *Language Learning* 53(1). 123–63.

Mathiassen, Terje. 1996. *Russisk grammatikk*. Oslo: Universitetsforlaget 2nd edn.

McClelland, James L. & Karalyn Patterson. 2002a. Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences* 6(11). 265–272.

McClelland, James L. & Karalyn Patterson. 2002b. 'Words Or Rules' cannot exploit the regularity in exceptions (reply to Pinker and Ullman). *Trends in Cognitive Sciences* 6(11). 464–465.

Nesset, Tore. 1996. *Russian conjugation revisited: a cognitive approach to aspects of russian verb inflection*. Oslo: University of Oslo, Faculty of Arts.

Nesset, Tore. 2004. Truncation without truncation? *Poljarnyj Vestnik* 7. 56–73.

Nesset, Tore. 2008. *Abstract phonology in a concrete model: Cognitive linguistics and the morphology-phonology interface.* Berlin: Mouton de Gruyter.

Nizegorodcew, Anna. 2007. *Input for instructed L2 learners*. Clevedon: Multilingual Matters.

Norris, John M. & Lourdes Ortega. 2000. Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis. *Language Learning* 50. 417–528.

Norris, John M. & Lourdes Ortega. 2001. Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning* 51(S1). 157–213.

Pinker, Steven. 1999. *Words and rules: the ingredients of language*. London: Weidenfeld & Nicolson.

Pinker, Steven. 2001. Four decades of rules and associations, or whatever happened to the past tense debate? In E. Dupoux (ed.), *Language, the brain, and cognitive development: Papers in honor of Jacques Mehler*, 157–179. Cambridge, MA: MIT Press.

Pinker, Steven & Alan Prince. 1988. On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28. 73–193.

Pinker, Steven & Michael Ullman. 2002. The past and future of the past tense. *Trends in Cognitive Sciences* 6(11). 456–463.

Plunkett, Kim & Virginia Marchman. 1991. U-shaped learning and frequency effects in a multi-layered perceptron: implications for child language acquisition. *Cognition* 38. 43–102.

Plunkett, Kim & Virginia Marchman. 1993. From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition* 48. 21–69.

Plunkett, Kim & Virginia Marchman. 1996. Learning from a connectionist model of the English past tense. *Cognition* 61. 299–308.

Prasada, Sandeep & Steven Pinker. 1993. Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes* 8(1). 1–56.

Protassova, Ekaterina. 1997. Transition from babbling to word structure. *Papers and Studies in Contrastive Linguistics* 33. 153–158.

Robinson, Peter. 1996. *Consciousness, rules, and instructed second language acquisition* (Theoretical Studies in Second Language Acquisition 7). New York: Peter Lang.

Robinson, Peter. 2002. *Individual differences and instructed language learning* (Language learning and language teaching 2). Amsterdam: John Benjamins.

Robinson, Peter & Nick C. Ellis (eds.). 2008. *Handbook of cognitive linguistics and second language acquisition*. New York, London: Routledge.

Rumelhart, David E. & James L. McClelland. 1986. On learning the past tenses of English verbs. In David E. Rumelhart & James L. McClelland (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. II, 216–271.

Scatton, Ernest A. 1984. Two stems or not two stems, is that the question? In Sophia Lubensky & Donald K. Jarvis (eds.), *Teaching, learning, acquiring Russian*, 213–227. Columbus: Slavica.

Schwartz, Bonnie D. 1992. Testing between UG-based and problem-solving models of L2A: Developmental sequence data. *Language Acquisition* 2. 1–19.

Sharwood Smith, M. 1993. Input enhancement in instructed SLA. *Studies in Second Language Acquisition* 15(2). 165–179.

Singleton, David. 2003. Age and second language acquisition. *Annual review of applied linguistics* 21. 77–89.

Slioussar, Natalja A. 2003. *Psicholingvističeskoe issledovanie struktury mental'nogo leksikona na materiale russkih glagolov. (Psycholinguistic study of the mental lexicon structure based on the Russian verb investigation)*. St.Petersburg SPbGU MA thesis.

Spada, Nina. 1997. Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching* 30. 73–87.

Tomasello, Michael. 2003. *Constructing a language: a usage-based theory of language acquisition*. Cambridge, Mass.: Harvard University Press.

Townsend, Charles E. 1975. *Russian word-formation*. Columbus, Ohio: Slavica Publishers corr. repr. edn. By Charles E. Townsend.

Ullman, Michael. 1997. A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience* 9(2). 266–276.

Ullman, Michael. 1999. Acceptability ratings of regular and irregular past-tense forms: Evidence for a dual-system model of language from word frequency and phonological neighbourhood effects. *Language and Cognitive Processes* 14(1). 47–67.

Ullman, Michael. 2001a. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research* 30(1). 37–69.

Ullman, Michael. 2001b. The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition* 4(1). 105–122.

Ullman, Michael. 2001c. A neurocognitive perspective on language: The declarative/procedural model. *Nature reviews. Neuroscience* 2(10). 717–726.

Ullman, Michael. 2004. Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92(1-2). 231–270.

Ullman, Michael. 2006. The declarative/procedural model and the shallow structure hypothesis. (commentary on Clahsen & Felser). *Applied Psycholinguistics* 27(1). 97–105.

Švedova, N. Ju. 1980. *Russkaja grammatika*. Moskva: Nauka.

White, Joanna. 1998. Getting the learners' attention. A typographical input enhancement study. In Catherine Doughty & Jessica Williams (eds.), *Focus on form in classroom second language acquisition*, 85–113. Cambridge: Cambridge University Press.

White, Lydia. 1989. *Universal grammar and second language acquisition* (Language acquisition & language disorders 1). Amsterdam: Benjamins.

White, Lydia. 2003. On the nature of interlanguage representation: Universal Grammar in the second language. In Catherine Doughty & M. Long (eds.), *Handbook of second language acquisition*, 19–42. Oxford: Blackwell.

Zaliznjak, A. A. 1980. *Grammaticeskij slovar' russkogo jazyka : slovoizmenenie*. Moskva: Russkij jazyk. A. A. Zaliznjak.

Zasorina, L. N. & V. A. Agraev. 1977. *Chastotnyj slovar' russkogo jazyka : okolo 40 000 slov (Russian frequency dictionary: about 40 000 words)*. Moskva: Russkij jazyk.

AUTHOR CONTACT INFORMATION

Elena Tkachenko
Department of Linguistics and Scandinavian Studies
University of Oslo
Norway
elena.tkachenko@lui.hio.no

Tatiana Chernigovskaya
Department of General Linguistics, Philological Faculty
St. Petersburg State University
Russia
tatiana.chernigovskaya@gmail.com

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 319–338. (ISSN 1890-9639)
http://www.journals.uio.no/osla

ANIMAL SOUNDS: A HUMAN VANTAGE POINT

EKATERINA RAKHILINA
Russian State University for the Humanities
[Translated from Russian by Irena Marijanović]

ABSTRACT

Why is it that the lexicon is often shunned and ignored by linguists, or in other words, treated with disdain and generally considered to be the *Ugly Duckling* of the linguistic family? This paper is both an attempt to redress the balance and it is designed as a tentative, initial contribution to the study of verbs of sound. Here, the central focus is devoted to a small subsection of verbs of sound, namely to the verbs denoting animal sounds used with metaphorical reference to human beings. The paper also attempts to sketch possible situations and parameters which are relevant for human beings and which appear to be cross-linguistic universals. The discussion is for the most part focused on Russian but examples from other languages, such as English, German, Estonian, and so forth, are also included.

[1] INTRODUCTION

There is a tradition according to which language is believed to be grammar and not the lexicon. This tradition has been upheld in linguistics for centuries, and for understandable reasons. It is curious, however, that this way of thinking about language is present in some fashion in the minds of "ordinary" native speakers. Indeed, when someone who is learning a foreign language says: "I seem to have picked a rather hard (or easy) language to learn", what one has in mind is that the inflectional system or the conjugational patterns of the language in question are complex. And not the fact that one has to learn a rather large, or conversely, a rather small number of words in order to speak freely: this aspect is not understood as an integral part of the language learning process. This situation can even be seen as enigmatic in some sense because one would believe that what the "ordinary" speakers and hearers say and hear are, in fact, words, that is, the very substance of language. But it is precisely the latter which is being ignored.

Strictly speaking, it is because of such a generally dismissive attitude towards the lexicon that we know so little about it. For instance, at the present time, we are unable to answer the question of whether a given language is simple or complex (or, in other words, whether it is poor or rich) when taking the lexical bulk as one's vantage point. Some linguists will argue that, generally, it makes no difference because all languages inherently have an equal wordage, compensating for

any gaps in their lexicon in some other fashion. To be quite honest, before I got involved in comparative research on lexis, I was of the same opinion. However, in order to provide a more accurate assessment of this issue, further, more comprehensive studies are needed. Indeed, it would be foolish to directly compare the lexical items of different languages by consulting the relevant dictionary entries. Every dictionary adheres to its own principles: some pay close attention to polysemy, whereas others ignore it; some include derivational morphology and yet others even prepositional or adverbial collocations (of the type *to go out*) as separate entries. Since this problem cannot be resolved in a straightforward manner, we should refrain not only from hastily appraising the lexical bulk of a given language, but also the volume or weight of the semantic field, and only then can we compare the two.

In their time, the morphologists also encountered a similar problem, that of an incommensurability of grammatical descriptions. Then, at the beginning of the '70s, this became an impetus behind a rapid development of the theoretical and practical aspects in the field of grammatical typology (cf., for example, Kholodovič (1969), Comrie (1976), Dahl (1985) and others). Now it seems the time of lexical typology may have come. The initial steps in this direction have already been made. If we are to speak of the advances made in this field on the international scene, then first of all we need to mention the projects of MPI for Psycholinguistics at Nijmegen (Majid & Bowerman (2007); Majid et al. (2008)), Cliff Goddard and Anna Wierzbicka (see, for instance, Goddard & Wierzbicka (2002)), Åke Viberg (Viberg 2002) and John Newman (Newman 1997, 2002, 2009); for a more detailed overview of the field, see Koptjevskaja-Tamm et al. (2007) and Koptjevskaja-Tamm (2009). In Russia, on the other hand, there is a well established branch of "diachronic" typology (see Dybo (1996), Zaliznjak (2001, 2009)) as well as synchronic research. With regard to the latter, we can mention an already completed research project on the "aquatic" verbs of motion Majsak (2007) and an ongoing project on the verbs of pain (Bonč-Osmolovskaja et al. (2009), Bricyn & Rakhilina (2009)). However, this is but a beginning. There ought to be many more research projects of this kind, and to gain an insight into the general picture such projects should address different problems in the field of lexical linguistics.

The present paper is a small step towards a future project which, for the time being, remains my own private reverie: it will ultimately focus on the verbs of sound. This field is in itself highly interesting because it provides a very subtle denotative basis for comparing different languages. Indeed, there is almost a complete absence of a "video sequence", so to say, which determines the situation structure of the verbs of motion (see Majsak (2007)) or the verbs of deformation (see the project on *cutting & breaking*). Strictly speaking, such verbs are simply untranslatable from one language to another: how should *šumet'* ('to rustle') be distinguished from *gudet'* ('to drone') so that we can find the correct translation?

One could assume that instead of the perceptual there is a common acoustic plane, but from a denotative point of view it is equally impossible to rely on sound as it is on perception: the sound of the forest and the sound of water are two entirely different sounds.

However, the world of sounds is extremely variegated (which implies that my imagined future project will be rather large). Indeed, there are sounds made by human beings (e.g. the Russian verb *šarkat'* ('to shuffle')) and those made by animals (e.g. *ryčat'* ('to bellow')); and then there are sounds produced by natural objects (e.g. *žurčat'* ('to murmur') and those by artefacts (e.g. *ljazgat'* ('to clank')), among which musical instruments represent a special class (e.g. *brenčat'*). It is also possible to establish further subclasses within each of these four classes (see the analysis in Stojnova (2008)).

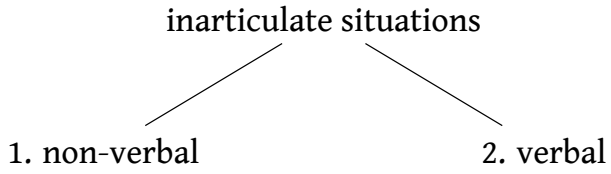
Here we consider the most cheerful subclass of all, namely, animal sounds. It is a known fact that animal sounds are used metaphorically with reference to human beings. A question arises: just which human sounds are rendered as "animal", "avian" or "insect"? The simple answer to this question is: inarticulate sounds. In fact, sounds that animals make may be likened to human sounds only if they do not convey the information customarily associated with human speech, or if humans do not identify such sounds as essentially "animal". However, the classification of these sounds, and of their corresponding situations, is of much interest: the present paper addresses the development of one such classification system and discusses the examples of its "work" based on a small sample of typological material. The lexico-typological component of this work is an attempt to understand which sound situations are so cognitively relevant that special lexical markers assigned to them are found in the languages of the world. We also raise the issue of to what extent these markers are unique with regard to their semantics, that is, whether there are rules and regularities to be found in the process of metaphor selection, or whether every metaphor in every language is simply unique. It is clear that we can speak with greater certainty of the existence of lexical typology, even in this unprototypical lexical domain, the more rules and regularities we uncover.

It should be pointed out in advance that we managed to examine only a few languages and can therefore boast of only rudimentary findings. It seems to us, however, that they should be of a considerable linguistic interest in the view of the novelty of the field of lexical typology. We are dealing, first of all, with Russian material, as well as English, German, Norwegian, Italian, Armenian, Czech, Bulgarian, Hindi and Estonian. I am deeply grateful to all my colleagues who provided me with the necessary data from the languages they are familiar with: L. Janda, T. Nesset, V. Khuršudjan, T. I. Reznikova, L. V. Khokhlova, U. Sturop, A. Van'kaeva, D. Stanulevič, and E. Tančeva.

[2] GENERAL CLASSIFICATION: A FRAGMENT.

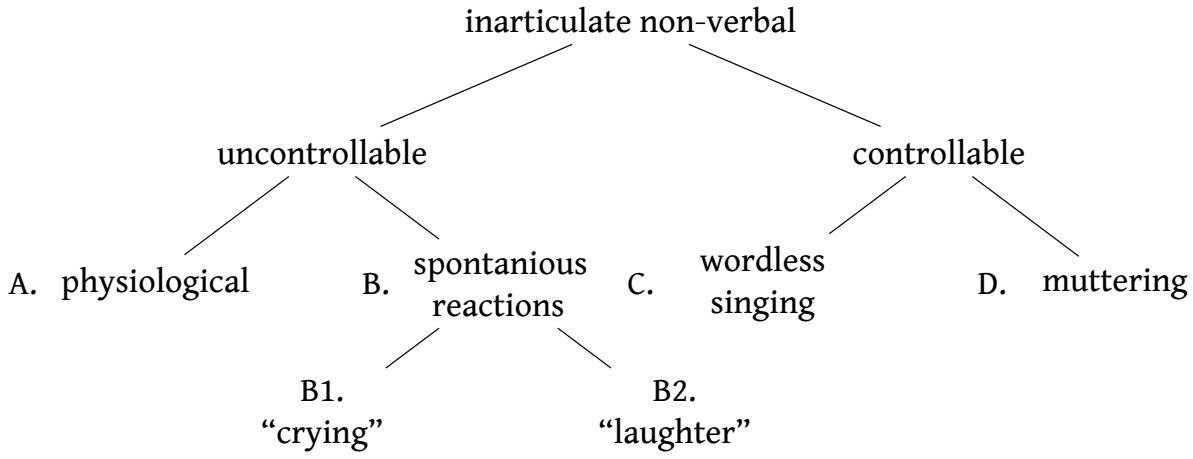
[2.1] *Non-verbal inarticulate sounds.*

And so, we take as our starting point the fact that inarticulate sounds made by humans in many languages of the world may, as metaphors, come to resemble animal sounds. The corresponding situations can be divided, with a certain grain of conventionality, into non-verbal and verbal.



We start with the non-verbal sounds, which can further be uncontrollable and controllable. The non-verbal uncontrollable sounds can, in their turn, be divided into physiological sounds, which are engendered by bodily processes (e.g. Russ. *krjakhtet'* ('to groan'), *čikhat'* ('to sneeze'), *khripet'* ('to wheeze'), *khrapet'* ('to snore'), etc.), and the spontaneous (non-verbal) reactions to an external situation. The latter can either be negative, such as "crying", or positive, such as "laughter". With regard to the non-verbal controllable sounds, all kinds of muttering lacking a specific addressee may be considered to belong here, as well as wordless singing.

Therefore, the first of the two branches of our classification has the following appearance:



We shall systematically examine its terminal subclasses with regard to animal metaphors in the world languages.

A. "Physiological sounds"

As already mentioned, we have in mind the spontaneous bodily reactions which are somehow or other accompanied by sounds. This takes place precisely when

one's stomach starts "making sounds" independently of one's volition, for example, after a meal. Such a situation stands out and is almost always rendered by an animal metaphor. It is as if there were a wild animal in one's stomach, usually a dog or a bear, as in Russ. *určat'* ('to growl') or Eng. *growl* (of a dog and a bear). Next, we turn our attention to this unique contamination of a dog and a "good" bear, which we revisit once more later in the paper. We also observe that no reference is made to a wolf although in principle, as we shall see later, a dog may be easily mixed with a wolf in other cases.

Pigs, as a source of metaphors, play the most significant role in this set of situations. It is a question of a characteristic wheezing (grunting) sound which by virtue of its phonetic quality is easily likened to various physiological processes: in Armenian, the sound will be interpreted as 'to emit a death rattle'; in Estonian as belching; in Kalmuck as snoring; in Bulgarian as groaning (including groaning from pleasure). A regrettable exception is Russian in which all above mentioned situations (snoring, groaning, and so on) are lexicalised, that is, each situation is expressed by its own verb, but not with the help of a zoological metaphor. Furthermore, the verb *khrjukat'* ('to grunt') does not have a stable, conventional context in which it could be applied to humans.

A curious situation obtains in Czech in which the corresponding verb is *hrochtat* and, like in Bulgarian, it is interpreted as groaning, in particular when one is lifting something heavy. However, the same verb can be translated into other languages not only as 'to grunt (of a pig)' but also as 'to hippo' (something in the vein of 'to make hippo-like sounds') because in Czech the word for hippopotamus itself (*hroch*) is directly associated with that sound.[1]

Another important source for the physiological domain are the cries of "clamorous" and "shrill" birds, above all the sounds produced by geese and crows. However, they only imitate the catarrhal voice in our sample (Bulg. *grača* 'hoarse voice' – lit. of geese and crows; Est. *kraaksuma* 'hoarse voice after having recovered from a cold' – lit. of crows).

As we can see, the whole physiological domain represents onomatopoeic sounds in their pure form: the metaphor's donor domain is chosen on the basis of a sound as such without any recourse to animal imagery. Precisely because of that the borrowing domain is very homogeneous and is replicated from language to language. In addition, the range of resulting recipient meanings itself is not very wide, although here one may expect the unexpected, for instance 'death rattle' as in Armenian, or 'belching' as in Estonian.

[1] This latter circumstance compels us to think about the possible lexical amalgamations of animal names and/or their sounds in the world languages. Even a most preliminary investigation reveals a high frequency of amalgamation of croaking (frog) and quacking (duck), especially in Norwegian, Czech, Estonian and other languages. An amalgamation of sounds produced by animal young is also possible. So, according to the data provided by our expert in Kalmuck, snoring is likened to the grunting of a piglet (not of a fully grown adult pig) which, in its turn, is amalgamated with the purring of kittens.

B. Spontaneous reactions

As already mentioned, spontaneous reactions can be both positive (see Section B2) and negative (see Section B1), and certainly with regard to language as a whole, the domain of negative assessment is developed more thoroughly than the positive, although the negative and the positive assessments may be amalgamated from time to time (see Section B3). For the moment, however, we are only talking about the non-verbal human reactions in this segment of our classification system, that is, about those reactions which are not accompanied by words. This means that we are dealing with what could conditionally be named "crying" and "laughter", and the varieties of these. We start the discussion with "crying".

B1. "Crying"

That "crying" is a variegated phenomenon is already adequately demonstrated by Russian – the different kinds of "crying" are so significant for humans that they are lexicalised in different ways. In Russian, there are at least four different zoological "crying" metaphors, each with its own specific meaning: *revet'* ('to roar') (lit. pertaining to a large, predatory animal, prototypically to a bear, but also a lion, tiger, etc. but not to a wolf or, for that matter, a fox); *vyt'* ('to howl') (of a wolf or a dog); *skulit'* ('to whine') (of a wolf and a dog), and *piščhat'* ('to peep') (of chicks and mice – the two are not separated in Russian). It is quite obvious that like the whole domain of the zoological metaphor, and in particular those metaphors included in the negative domain, all these meanings are either coarse or very vulgar. In spite of this, such metaphors have a sufficiently wide usage. So, *revet'* is used with reference to children, denoting a very loud crying accompanied by tears, which does not refer to infants. It is the verb *piščat'* that is used to describe "infant crying". A man cannot *revet'* in that sense – when referring to a male person, this verb denotes a very loud and aggressive voice, that is, a verbal reaction (cf. with a more characteristic verb *vzrevet'* ('to give a roar')). In principle, an adult female cannot *revet'* unless she is consciously likened to a child. *Vyt'* (cf. also the inceptive verb *vzvyt'* ('to give a wail')) means to cry from an intense pain without tears and it resembles a wolf's cry. The other characteristic context for *vyt'* is the female keening over the body of a deceased (also without tears). *Skulit'* differs from *vyt'* with regard to both the intensity of the sound and original meaning: a fully grown wild animal can howl (*vyt'*) out of loneliness, it is said, as for instance a wolf howls at the moon, but a whelp whines (*skulit'*). In addition, whining is associated with making a request. For example, a dog whines at the table to persuade his owners into sharing some of their food with it. Because of that the Russian verb *skulit'* is more "inoffensive" than *vyt'* and is associated with a plaintive as well as suppliant crying.

In principle, a very similar situation obtains in other languages we examined: it focuses above all on wolves and dogs. Armenian is of an indisputable interest

here, in which mooing is included in the same group. As we shall see, bovine mooing is associated with another verbal recipient domain, and the unexpectedness of its occurrence compelled us to scrutinise the Russian material more attentively. And here we stumbled upon a discovery: in Russian, the range of sources for the verb *revet'*, comprising a set of large predatory animals, includes not only bears but also bulls that, unlike cows, do not *myčat'* ('to moo') but they do *revet'*. However, if we were to assume that the sounds made by bulls and cows, that is bellowing and mooing respectively[2], can be amalgamated in a given language, then the case observed in Armenian becomes better motivated.

B2. "Laughter"

Here, the source of the zoological metaphor in Russian appears to be horses (*ržat'* ('to neigh')) and geese (*gogotat'* ('to gaggle')). Both of these verbs denote a very loud and raucous laughter, in the latter case the laughter is more "discreet" with regard to sound, as is the sound of the source. A close juxtaposition between a "homogeneous" and more "staccato" laughter is encountered in Armenian: between 'chirr' (of grasshoppers) and 'bleat' (of sheep) respectively. In English, the source of a metaphor for loud laughter is the owl (cf. the verb *to hoot* and also the expression *that was a real hoot* meaning 'it was very funny'). With regard to geese, their characteristic cry is used differently in English. It forms a zoomorphic metaphor in the artefactual domain: the English verb *to honk* imitates the sound of a car horn and is translated into Russian as 'bibikat''. A separate problem, of course, presents the question of artefacts as the sources of sound. Presently, it is not possible to delve deeper into this topic but the very fact that this field is structured in quite an interesting way, something which we shall encounter later on, and is easily captured by the zoological metaphor is very significant.

B3. "Crying"/"Laughter" (mixed reactions)

We should bear in mind that the positive and negative reactions cannot always be teased apart. It happens that both extremities are expressed by a single lexeme. An example of this kind is the Russian verb *vizžat'* ('to squeal both from pain and happiness'). It is true that this verb cannot be called zoomorphic in the full sense of the word because it does not juxtapose a human to an animal sound (here we have in mind the squealing of a pig), so strictly speaking, this is not the case of a

[2] The names of domestic animals are always fairly well distinguished lexically on the basis of sex, cf. Russ. *kozel – koza* ('billy-goat – nanny-goat'), *baran – ovca* ('ram – ewe'), *kot – koška* ('tomcat – cat'), etc.

semantic shift.[3] However, positive and negative reactions in metaphors are not contrasted in precisely the same manner, e.g. compare the English verb *to howl* (of wolves) meaning 'to cry or laugh very loudly' to the verb *to roar*, which carries approximately the same meaning and is in its literal meaning associated with lions, but not with bears or bulls as its corresponding Russian translation, *revet'*, is. Incidentally, the Russian verb *revet'* also belongs to the class of mixed reactions, but only when it is used with a plural subject[4] (cf. *tolpa revela* ('the crowd roared') may mean that the crowd is saluting the leader, or demanding his immediate resignation). We address the question concerning the role of the plural subject later in the paper.

Presently, we address the issue of non-verbal, uncontrollable sound situations as the recipients of zoological metaphors. In principle, it is not simple to distinguish the latter from the verbal ones precisely because humans have control over these sounds, which means that they utter them consciously. The basic parameter which could be used here to establish a more or less unambiguous border between the classes could be the presence or absence of an addressee. Indeed, prototypical speech is always addressed to someone; it fulfils a particular communicative task. The situations which we are presently examining preclude communication in the full sense of the word: here the acoustic substance, even if it contains words, is not addressed to anyone, with the possible exception of the speaker himself. However, even communication of this kind can be subdivided into distinct subclasses sufficiently connected with the world of animal sounds. These are "wordless singing" (C) and "conversation with oneself" (D); in addition, we have identified one more class (E) which may be tentatively called "singing without music".

C. Wordless singing

The situation in which wordless singing obtains is sufficiently significant in many languages. In Norwegian, for instance, there is a special underived verb, *å nynne*, which is used to describe it. In Russian, it is expressed by means of a metaphor, the "feline" verb *murlykat'* ('to purr'). It does not simply describe a cat's voice (cf. with the Russian verb *mjaukat'* ('to mew') which lacks such a meaning) but a

[3] On the other hand, the semantic shift is clearly discernable in the English equivalent of *vizžat'*, namely *to squeal*. Here we have in mind the expression *to squeal on smb* which in its pure "human" sense means 'to turn informer; betray an accomplice or secret'. We observe that in Russian the same meaning may be rendered by an acoustic metaphor: *stučat' na kogo-l.*(lit., 'to knock on smb.'). However, in Russian, this mapping is realised not through the domain of animal sounds, but through an entirely different causal domain, that of the verbs of sound. For this reason, the metaphorical imagery regarding denunciation differs in Russian and English. In English, it is a vociferous announcement delivered in a particularly unpleasant tone of voice, as a shriek, of the type: *"Mar' Ivanna, Petrov opjat' na uroke pljuetsja!"* ('Mar Ivanna, Petrov is spitting again in the lesson!'). In Russian, on the other hand, it is a secret intimation, like rapping, cf. a different way of designating denunciation in Russian, namely, *naušničestvo*.

[4] On the behaviour of the singular subject see B1

voice of a *satisfied* cat, which is semantically close to the canine/ursine verb *určat'* ('to growl') (see above). In other European languages, there is a more widespread link with the world of insects. In English, for instance, the verb *to hum* is used with that purpose in mind: it denotes a "quiet" buzzing that is characteristic of mosquitoes and flies but not bees (see below about the apian buzzing that is very often juxtaposed to that of mosquitoes). In particular, the verb *to hum* describes the manner, known to all, in which Winnie the Pooh sang his grumblings. This verb may be applicable to the act of singing with one's mouth closed, that is, completely without words. The German verb *summen* behaves in a very similar way; however, it extends to mosquitoes and bees. The French verb *bourdonner*, with the same meaning, describes the buzzing of flies, beetles, and humming birds.

D. Conversation with oneself

It is a known fact that this situation occupies an even more prominent place in the lives of humans than singing does. Humans, in particular although not exclusively old men, talk to themselves, and such a condition ought to have a name. In Russian, there is an underived onomatopoeic verb *bormotat'* ('to mutter'), but in many other languages this onomatopoeia is associated with low animal or avian sounds, like those made by a bear, as in German, or a chicken, as in English. It is interesting that the Norwegian cognate of the English verb *to cluck*, which is used in this situation, namely *å klukke*, has a more specific meaning: it does not simply mean 'to mutter to oneself' but 'to laugh quietly at oneself'.

E. Singing without music

A few more words about singing. Not only is wordless singing distinguished zoomorphically in many languages, but also "singing without music" that is off-key and unpleasant to the ear, that is, what in Russian is expressed by the collocation *tjanut' kota za khvost* ('to pull the cat by its tail'), cf. an example from the RNC (Russian National Corpus):[5]

(1) Snizu igrali muzyku, khotja, kak vsegda, tjanuli kota za khvost (Asar Ėppel', *Pomazannik i Vera* 1990–2000)
'They were playing music downstairs although, as always, it sounded abysmally off-key' .

[5] Another meaning of this phrase is 'to hold back, detain', cf. also *tjanut' rezinu* ('to drag one's feet') or *volynit' volynku* ('to dawdle over smth') . Note also from the RNC:

(i) – Tak bystro? – A čego tyanut' kota za khvost. Raz-raz – i gotovo! – Čto ona delaet? Legla spat'? (V. P. Kataev, *Dorogoj, milyj deduška*)
'So quickly? But why beat about the bush? One-two-three and it's done! What is she doing? Has she gone to bed?'

In Armenian, the "feline" metaphor is indirectly realised by the verb with the literal meaning 'to mew', which means 'to sing badly in a high-pitched voice'. In English, one identifies badly performed violin music and uses the verb *to squeak* in its literal sense to refer to it, but this verb is also zoomorphic since it may refer to the squeaking of mice. It is similar to the Russian verb *skripet'* ('to screech') in the sense that it is used to refer to doors and car brakes. In this way, the instrument in question, *violin* in English, if shoddily made or amateurishly played, emits a sound similar to that which in Russian gets its name from the word *skripka* (i.e. 'that which squeaks').

[2.2] *Inarticulate speech*

Presently, we examine the other branch of the classification system, to which pertains everything which is connected with speech, even inarticulate speech. It appears that here it is possible to distinguish the following four classes, each of which will be examined in turn:

A. Inarticulate speech (of infants or adults)

B. Approving/disapproving reactions

C. Plurality "speakers"

D. Semiotically meaningful speech

A. Inarticulate speech (of infants or adults)

Infants cannot pronounce words: their "conversation" resembles more the singing of birds than human speech; hence, we have the Russian verb *gulit'*, and the English *to coo* with the same meaning, where both verbs describe one of the "pigeon" sounds (cf. see below *vorkovanie* ('cooing')).

But adults may also speak in such a fashion that it is sometimes difficult to discern what they are saying: their speech can be both unintelligible and incoherent. This effect is engendered by either of the following two mutually exclusive causes: when the speed of the discourse is either too slow or too fast. If the speech is too slow, unsure, having a staccato-like quality, and consequently too disjointed, then it describes a stuttering speaker, as it were, who makes endless pauses or self-corrections (like the Russian verbs *bekat'* ('to bleat'),[6] mekat' ('to bleat') , originally of a goat), or un unstructured discourse (like the Russian verb *myčat'* ('to moo'), originally of a cow). Both types of slow speech correspond to a recognisable situation of a D-student at an oral exam: in Italian, this situation is associated with the bray of a donkey (*ragliare*).

[6] It is curious that according to our data the Norwegian verb *å breke*, with the same literal meaning 'to bleat', utilises a completely different acoustic segment: not intermittency but the colour of someone's voice; it denotes an unpleasant male voice.

However, the feeling of incoherence is also created when the speech, on the other hand, is too rapid, so that one's thoughts fail to take actual shape; they interrupt one another (cf. the Russian onomatopoeic verb *taratorit'* ('to jabber')). *Taratorit'* is not zoomorphic but in Russian this verb has a zoomorphic quasi-synonym, *strekotat'* ('to chirr') (of a grasshopper). The latter describes not only human, more precisely female chatter, but also the sounds made by such artefacts as sewing machines or type writers. In this connection, we can mention the verb *treščat'* which describes, apart from the sound of breaking wood, the "conversation" of magpies; hence, the verb's applicability to female speech.

On the whole, a similar situation obtains in Crimean Tartar: as in Russian and many other languages, rapid female speech is also coded in this language by specific lexical means. However, such speech is associated with the more traditional artefacts of that culture: in Crimean Tartar, 'to chatter' is rendered by the same verb which also denotes 'clatter', 'shrill sounds of zurna' and 'clicking of the beads on the abacus'.

In English, rapid incoherent speech can be rendered as canine yapping, which according to our informants, is not gender specific, e.g. *they were just yapping away the whole night* which may be associated with both men and women. Furthermore, in Estonian, we encountered two zoomorphic verbal metaphors of the same kind, which according to our informants are also not gender specific. These are the verbs *kaagutama* (lit. referring to the clucking of chickens) and *prääksuma* (lit. 'to quack and metaphorically 'to talk rubbish (especially pertaining to children)). However, we note that in Norwegian cackling (*å kakle*) is in its metaphorical meaning interpreted as associated only with women: as a 'loud meaningless conversation or laughter'.

Thus, rapid female speech (always pejoratively judged as incoherent, meaningless, etc.) appears to be a significant parameter in the sound domain. Evidently, it is necessary to distinguish female laughter as a separate class in typological questionnaires on this topic. In English, along with the verb *to yap*, which is unmarked with respect to gender, there is also a metaphor especially associated with the female chatter and laughter: these two activities are coded by the verb *to tweet* whose literal meaning denotes avian "conversation", something along the lines of the Russian verbs 'ščebetat' ('to chirp'), 'čirikat' ('to twitter').[7]

B. Verbal reactions

In this section, we examine the examples of verbal inarticulate reactions. Just as non-verbal, verbal reactions are divided into positive and negative, but they do not form a further mixed class since the contrast is sufficiently discrete. As before, the domain of negative, that is, disapproving reactions is developed significantly

[7] It should be pointed out that *ščebetat'*, and to a lesser extent *čirikat'*, may also be applied to female speech but only in a positive way, see below for further discussion.

more thoroughly, and we start the discussion with the latter.

B1. Disapproving reactions

Disapproving verbal reactions represent a verbal confrontation to someone else's speech, action, or a situation generally speaking. One can oppose all that to a varying degree, starting with a simple expression of disapproval, to an intense enmity, which may turn into aggression, so that the verbal reactions are scaled and thus subdivided into smaller classes.

B1.1. Weak confrontation

FYRKAT' ('to snort'). A very weak type of confrontation is described by the Russian verb *fyrkat'*, which in its literal meaning is associated with a brief characteristic sound made by a horse or a dog upon coming out of the water. It is used metaphorically as an expression of refusal or (passive) disapproval, cf. the following example from the RNC:

(2) Podružki peregljanulis', fyrknuli, podkhvatili drug druga pod ruku i pribavili šagu – komu nravitsja byt' ob"ektom rozygryša (Semen Daniljuk, *Rublevaja zona*)
'The two friends exchanged looks, snorted disapprovingly, took each other under the arm and quickened the pace. Who enjoys falling victim to a practical joke?'

It is important to bear in mind that, even in the metaphorical meaning, the verb in question retains the sound that accompanies it despite the fact that the verb becomes semantically loaded. That is as a rule a special kind of sound, although it is usually accompanied by some form of speech, (cf. from D. Dontsova: *Samo ničego ne portitsja – fyrknula Ol'ga* ('I only hope nothing gets spoilt – Olga snorted.). Note, however, the unmarked *fyrknula i ušla* ('she snorted and left') – it is possible that something may have been said disapprovingly, and, on the other hand, it may not have been. In English, there is a direct analogue for this Russian metaphor but its source is different, being namely grunting which, as we have seen, transcends a simple physiological onomatopoeia in Russian.

VEREŠČAT' ('to churr') Furthermore, Russia has its own verb of verbal reaction associated with the sound of a pig, namely, *vereščat'*, although it is possible that this verb, just like *vizžat'* ('to squeal'), is not entirely zoomorphic. When applied to humans, it may denote a sharp, rumbling sound accompanying an act of passive

resistance (particularly that of children), which, as a rule, is easily suppressed.[8] In addition, the confrontation may also be verbal having the same quality of timbre.

VORČAT' ('to grumble'). An ordinary grumbling is a form of passive verbal opposition, and is of a lower register: the speaker has an objection – he does not agree – but his disagreeing is so passive in nature that it is clearly not directed towards the source of displeasure, but towards himself, so to speak. This situation is metaphorically represented in German through the verb *brummen*, which literally describes the sounds made by a bear, bull or a swarm or bees, something which in itself presents a curious case of contamination. We have already noticed the association of a bull with a bear; therefore it occasions very little surprise. But the association with a swarm of bees is rather interesting. We notice that the sound of a swarm is clearly distinguished from the sound of a single bee, and they belong to the different classes of the zoological metaphor. This is especially evident in the Russian material: a bee buzzes (*žužžat'*) like a small machine such as a spindle or an electrical shaver, but a swarm drones (*gudet'*) like bells or heavy airplanes above the airport. Furthermore, the German verb *brummen* may also apply to the drone/hum of the airplanes. It is precisely the whole swarm and not a single bee that is amalgamated with a bear and a bull in German. In Armenian, a tom cat grumbles in approximately the same manner, perhaps softer: it does not mew (mewing is phonetically not similar to grumbling); it does not purr (purring is a positive and not a negative reaction). And that what could be termed "to whir" may be applied to humans (cf. with the possible meaning 'to whir from spite' in Armenian).[9]

ŠIPET' ('to hiss'). This is yet another type of a passive reaction. This verb is not directed towards the source of displeasure but towards those individuals in the surrounding environment and not only towards oneself. It is true that this action is performed clandestinely, silently; hence, we have the following type of a zoological metaphor: the hissing of a snake. One (usually a woman as she is less likely to engage in open conflict) hisses from jealousy and malice in many languages, including in the Slavonic and Germanic.

[8] Compare, however, the following example from the RNC:

(i) Vereščit, ručonkami soprotivljaetsja i ne daetsja ni v kakuju, pop s nej izmučilsja, no krestik na nee nadet' tak i ne smog. (Ėduard Volodarskij, *Dnevnik ubijcy*)
'She is hissing, resisting with her small hands and is not about to give in for anything. The priest was exhausted from fighting her, but he still did not manage to put the little cross on her.'

[9] Compare here with the expression found in the EANC (Eastern Armenian National Corpus) which is literally translated into Russian as ' i čtoby my vmeste perežili [peremurčali] našu grust'' ('so that we may overcome [over-whir] our grief') (`www.eanc.net`).

However, Norwegian uses its own resources to express the "hissing" metaphor. The animal used is a type of a polar mouse – a lemming – a rather small, yellowish, almost tailless creature, like a hamster or a ground squirrel, whom Norwegians perceive as being quite angry. From time to time, the lemmings migrate and when they do, hordes of lemmings tend to occupy large spaces. The hissing of lemmings is denoted by the verb *å frese*, and this verb when applied to humans means that one is feeling angry and this emotion becomes overwhelming, as if one were seething with rage, but one is quite incapable of doing anything about it (cf. a similar metaphorical use of Russian *kipet'* 'boil'). Evidently, such "undirected" hissing is of a somewhat different type than what we have just described. Furthermore, it is not gender specific, that is, it is not seen as a solely female quality.

And we arrive at the last variety of hissing which we have encountered in our material. It also denotes a type of snake hissing but it has a different semantics, aptly expressed by the English verb *to hiss*. This verb refers to a feeling of indignation felt by a large group of people, for instance, in a grandstand during a football match or in the theatre. We observe that in Russian this type of situation has its own non-zoomorphic but onomatopoeic means of lexical marking, namely, the verb *šikat'* ('to boo, hiss'), since hissing is already "reserved" to denote female malice or jealousy.

B1.2. Aggressive confrontation

Aggressive confrontation could be called, with a certain degree of conventionality, a "canine" reaction because dogs, and to a lesser degree wolves,[10] serve as the donors of this type of metaphor. In Russian, these are the verbs *vjakat'* ('to yap'), *ogryzat'sja* ('to snarl'), and *ryčat'* ('to bellow'). They describe a confrontational verbal reaction that escalates in its intensity and is openly directed towards the source of displeasure. In English this is expressed by the verbs *to growl* and *to snarl* (also implicating the escalation in intensity, cf. *I told him we needed to leave and he just* GROWLED *at me, or If they* SNARL *at each other they are really fighting*. It is evident from these examples that the confrontation is so aggressive that the one who engages in it is wholly capable of emerging from it as a victor. A laconic and vivid interpretation of a pair of Norwegian cognates corresponding to these English quasi-synonyms provided by Tore Nesset may serve as a good illustration. The verbs in question are *å knurre* (lit. of a dog or a wolf, but not of a bear) and the more aggressive *å snerre*: "If I suggest that we should do something, and my addressee *knurre*, that means that he doesn't want to do it, but we'll end up doing it anyway. But if he *snerre*, then nothing will come of it."

[10] In Italian, according to some of our informants, it is lupine and not canine reactions that are mapped onto human beings, cf. *ringhiare* – lit. 'to growl (of a wolf)', metaphorically 'to react brusquely (of people)'.

B2. Approving reactions

There are very few reactions of this kind. In Russian, we found two in fact: *krjaknut'* ('to give a quack'), as a reaction of surprise as well as approval in response to an unexpected action from one's partner, and *myčat'* ('to moo'), which refers to a both verbal and non-verbal reaction of pleasure (for instance, when one's back is being scratched or when one is eating something delicious). The latter is attested in Bulgarian, *muča*. Furthermore, the verb *grukham* is used in Bulgarian as a lexical marker of pleasure with the literal meaning 'to grunt'. It is known that this verb has a metaphorical semantics different from the verb *muča*. It seems that the meaning is closer to a satisfied growling. In Russian, growling cannot be applied to humans, something which is possible in Estonian, in which the verb *mõmimisema* literally denotes growling of a satisfied bear.

C. The plural subject

The zoological metaphor helps distinguish a class of plural subjects which are important for humans. The conversation of couples that are in love (cf. Russ. *vorkovat'* ('to coo') – lit. of pigeons) as well as when two people hurl abuse at each other (cf. Russ. *lajat'sja* ('to yap')) are usually marked in different languages. Furthermore, as we have already examined it in some detail, women who chatter and laugh, and possibly children too (see 2.A above), present an important type of plural subject. *Crowd* in connection with non-verbal reactions (1.A) was also mentioned as a relevant plural subject. We have already discussed the fact that in the context of such a plural subject, as *crowd*, the verb revet' ('to roar') (lit. of a bear, lion or bull) is not perceived in the same fashion as in the context of a singular subject. The former denotes an evaluative non-verbal reaction, whereas the latter crying of a child or an aggressive male voice. Moreover, in Russian, the *crowd* may also *galdet'* ('to make a racket') like a number of large birds such as jackdaws or crows (cf. the Bulgarian verb *gracha*, lit. of crows or geese). This is a description of a simultaneous yet incoherent loud speech characteristic of a large number of people. By itself, such speech is neither reactive nor evaluative, in contrast to *revet'* or to the second example of the zoological metaphor with the plural subject, the English verb *hiss* (2.B1), but the speaker himself judges it negatively as an unnecessary clamour. Quite another matter is the Russian verb *gudet'* ('to drone'), literally pertaining to a swarm bees or a beehive (cf. also the English verb *to buzz* (bee, mosquito)), which also relates to "orderly" noise of the same kind, for instance, when a group of people purposefully discusses an idea. In this situation, the evaluation may also be positive.

When speaking of verbal situations which in a given language may be expressed as animal sounds, in particular as verbal reactions, we can clearly distinguish them from the physiological sounds: it is precisely speech that possesses its own semantics. Relaying on a metaphor, we certainly cannot exactly repro-

duce what was in fact said, but we do know HOW it was said, and can therefore infer the contents. Thus it is not only the phonetic form of a sound, intrinsic to a given animal as in the physiological zoological metaphors, that is significant in verbal reactions, but the linguistic form of the animal itself. In other words, the metaphors are not created here simply on the basis of the onomatopoeic effects, but on the basis of more complex semantic considerations.

The following group (2.D) comprises verbs which either use onomatopoeia indirectly or do not have any onomatopoeic associations at all: these are the lexemes which code verbal sign situations present in a given culture by relying on the animal form.

D. Semiotically significant speech

A good example of a semiotically meaningful metaphor is the Russian verb *zudet'* ('to buzz') (lit. of a mosquito) in its metaphorical meaning 'to bother, to nag someone with one and the same advice, request or tale of moral edification'. *Čto ty zudiš'!* ('Why are you on my case!) may be heard as a response to an insistent request *wash the dishes* and even *to don't procrastinate, do your homework*. Such a response will necessarily be crude and impolite but not impossible to make. It is an interesting fact that in Hindi a verb denoting the buzzing of a fly performs the very same function. It is clear that in the given case the metaphor relies on the monotonous sound made by an insect that is intent on capturing its prey and is perceived as an irritating but minor nuisance.

We also note that a similar verb in Italian – *ronzare* (pertaining to mosquitoes, flies, and bees) – evolves in a completely different fashion. Its acoustic component in the recipient domain is made wholly subordinate to the idea of a purposive circular motion, which is associated with all these insects. This verb undergoes not a metaphoric but a metonymic shift and, as a result, it means 'to hang around a girl'.

The meaning of the Russian verb *brekhat'* ('to yap') is also quite removed from the onomatopoeia: it does not bring to mind the canine barking. However, the origin of this metaphor is clear: it comes from the idea of "empty" barking, barking without a reason, which erroneously informs the owner of the danger. A very similar idea is present in the Bulgarian verb *laja* which is used particularly in the situation when the politicians talk a lot and without making much sense.

The semantics of mapping of both the English verb *cackle* to mean 'to care for someone (with a touch of excessiveness) and the Bulgarian *kudkudjakam* to mean 'to find oneself in a panic (of women)' can be traced to the image of a stupid and restive brooding hen. It is also probably possible to explain the origin of the German metaphor associated with the word *kollern* ('cry of a turkey'), namely, 'to speak angrily', in that it refers to the "angry" facial expression of an important person. But why is 'to neigh' (of horses) in Hindi when applied to humans under-

stood as 'to speak with false enthusiasm' (cf. Russ. *veščat'* ('to prophesy, broadcast'))? Or why does the Russian verb *karkat'* ('to croak') (of crows), but Armenian 'to croak (of frogs)', mean 'to predict bad development of events in the "human" domain when in Hindi the verb 'croak' is significantly closer to an onomatopoeia: the monotonous repetition of a crow's cry means 'to repeat the same thing' when associated with human speech.

[3] A LITTLE BIT MORE LINGUISTICS

It seems that the task of building a typology of such "weather-beaten", semiotically saturated recipient meanings will not be a simple matter because here it is difficult to compare the lexemes from different languages, but perhaps the number of such unadulterated cases may not be very high in the end. However, there is a different, no less serious danger with regard to the development of such a typology: the illusion of a facile juxtaposition of lexemes. Let us take, for instance, the English verb *to bark*, which is easily and correctly translated into Russian by the verb *lajat'* because both of these verbs presuppose one and the same type of subject (dog) and, generally speaking, imply one and the same sound. It is clear that not every English expression with bark will be translated as *lajat'* and vice versa[11], but, on the whole, this should not interfere with their comparability.

However, casting a more attentive glance at these verbs should raise some questions. In particular, it seems that the metaphorical verb *to bark* is naturally used with the collocation *to bark commands* ('to issue military orders in a brusque, especially shrill, tone of voice'). But such a mapping is impossible in Russian: in this case, one would rather say *rjavkat'* ('to bellow, bawl') and not **lajat'* owning to wholly linguistic reasons. The point is that the Russian verb *lajat'* can only be interpreted iteratively and it simply cannot describe a single act of barking (a single bark, as it were), whereas the English *bark* is clearly more neutral in this respect.

Most probably, the parameter 'singular : iterative', as well as 'discrete : non-discrete', should be regarded as relevant, albeit to a lesser degree than the type of subject.

And incidentally, there is one more interesting question: how is the quantum of sound specified in different languages? Indeed, it is well known that the quantum of food stuffs is lexically coded, e.g. Russ. *ne s"el ni kroški* ('he didn't eat a crumb'), as well as the quantum of liquids, e.g. Russ. *ne popil ni kapli* ('he didn't drink a drop'). According to all traditional theories of metaphor, emotions may

[11] For instance, in English, there is a phase *you're barking up the wrong tree* which conjures up a hunting scene in which a dog is chasing a cat: the cat is sitting in the tree while the dog is barking at it. Here the dog is barking at the "wrong" tree – the one without a cat. Consequently, the meaning that emerges is 'you're swearing in vain', that is, more literally 'that's not the person to be barking at'. Of course, this meaning is not to be directly found in Russian. See Dobrovol'skij & E. (2005) on the theoretical aspects of the culturally conditioned specificity of phrasal metaphors.

be equated with liquids because the quanta of liquids are suitable to emotions, e.g. compare *ne boyalsja/ne ljubil ni kapli* (lit. he wasn't afraid/didn't love (her) one single drop). And what about the sounds? We usually say: *On ne izdal ni zvuka* (He didn't utter a sound) – but what sound? How is such a sound to be marked lexically? What is it compared to, if this is a metaphor? In such cases, one usually says in Russian, *ne piknul* (he didn't make a peep), if the person in question was ordered to do something, and although he did not agree with it, he did not remonstrate against it either.[12]

This is not an instance of a zoological metaphor. In contemporary Russian, the verb pikat', when in the imperfective aspect and having an iterative semantics, is more strongly connected to the artefactual instruments, such as radios, telephones, etc., and their electronic "peeping" sound, cf. an example from the RNC:

(3) telefon u nego pikaet každye pjat' minut (Andrej Belozerov, *Čajka*)
'his telephone keeps beeping every five minutes'

In English, the quantum of sound is zoomorphic: it corresponds to a single peep of a chick, e.g. she didn't make a peep. In Norwegian, on the other hand, it is a single duck "quack", e.g. *han sa ikke et kvekk* 'he didn't say a word' (lit. 'didn't say a quack').

[4] CONCLUSION

We have already mentioned that the present paper does not lay claim to comprehensiveness and completeness. Its task was merely to draw the attention to lexical typology, in particular to the problem regarding the construction of a typology of sound verbs. Is this task realisable in the domain of the zoological metaphor, for instance?

Our material shows that:

- the same sounds are categorised differently in different languages (in particular with regard to the opposition iterativity vs. singularity)
- the amalgamation of sounds ("human", "animal" and "artefactual") happens differently in different languages

These facts, at first sight, hinder the development of some universal system in the lexicon. At the same time, we have seen on the basis of the examined linguistic material that:

[12] Cf. from the RNC:

(i) A Strekalovykh tak pripugnem, čto piknut' ne posmejut (Leonid Juzefovič, *Kostjum Arlekina*)
'But I'll scare the Strekalovs so much that they won't dare to make a sound.'

- the same significant sound situations and their parameters are lexically distinguished in different languages, and they are expressly coded by lexical means, often zoomorphically, but sometimes also by their own verbs of sound or ordinary verbs
- it is clear that such situations and parameters are relevant for humans and are independent of a particular language and culture. It would be possible to say that they are universal (and that, of course, needs to be shown), or, at the very least, that they lay claim to universality.

In all probability, a typology of a linguistic domain that is developed in this way must be possible – shall we try and develop it?

ACKNOWLEDGMENTS

This work is supported by a grant from the RFFI (09-06-00364-a).

REFERENCES

Bonč-Osmolovskaja, A. A., E. V. Rakhilina & T. I. Reznikova. 2009. Conceptualization of Pain: a Database for Lexical Typology. In P. Bosch, D. Gabelaia & J. Lang (eds.), *Logic, language, and computation*, 110–123. Heidelberg: Springer.

Bricyn, V. M. & E. V. Rakhilina. 2009. *Koncept* bol' *v tipologičeskom osveščenii*. Kiev: Dm.Buraho Publishing House.

Comrie, B. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge U. Press. (rev. ed. 1981).

Dahl, Ö. 1985. *Tense and aspect systems*. Oxford: Blackwell.

Dobrovol'skij, D. & Piirainen E. 2005. *Figurative languages*. Amsterdam: Cross-cultural and cross-linguistic perspectives.

Dybo, A. V. 1996. *Semantičeskaja rekonstrukcija v altajskoj ėtimologii*. Moscow.

Goddard, C. & A. Wierzbicka. 2002. *Meaning and universal grammar: Theory and empirical findings*. Amsterdam: Benjamins.

Kholodovič, A. A. 1969. *Tipologija kauzativnykh konstrukcij*. Leningrad: Nauka.

Koptjevskaja-Tamm, M. 2009. Approaching Lexical typology. In M. Vanhove (ed.), *From polysemy to semantic change: Towards a typology of lexical semantic associations.*, Amsterdam and Philadelphia: Benjamins.

Koptjevskaja-Tamm, M., M. Vanhove & P. Koch. 2007. Typological approaches to lexical semantics. *Linguistic Typology* 11(1). 159–186.

Majid, A., J. S. Boster & M. Bowerman. 2008. The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition* 109(2). 235–250.

Majid, A. & M. Bowerman. 2007. Cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics* 18(2).

Majsak, T.A. 2007. *Aquamotion - glagoly dviženija v vode: leksičeskaja tipologija*. Moscow: IndrikBri.

Newman, J. 1997. *The linguistics of Giving*. Amsterdam, Benjamins.

Newman, J. 2002. *The Linguistics of Sitting, Standing, and Lying*. Amsterdam: Benjamins.

Newman, J. 2009. The linguistics of *eating* and *drinking*. In *Studies in typological linguistics*, vol. 84, Amsterdam and Philadelphia: John Benjamins.

Stojnova, N. A. 2008. *Glagoly zvučanija v russkom jazyke*. Moscow State University MA thesis.

Viberg, Å. 2002. The verbs of perception. In Oesterreicher W. Haspelmath M., König E. & W. Raible (eds.), *Language typology and language universals*, Berlin: De Gruyter. An International Handbook.

Zaliznjak, A.A. 2001. Semantičeskaja derivacija v sinkhronii i diakhronii. *Voprosy jazykoznanija* 2. 13–25.

Zaliznjak, A.A. 2009. Catalogue of Semantic Shifts: towards a Typology of Semantic Derivation. In M. Vanhove (ed.), *Towards a Typology of Lexical Semantic Associations*, Paris: Mouton de Gruyter.

AUTHOR CONTACT INFORMATION

Ekaterina Rakhilina
Russian State University for the Humanities
Moscow, Russia
rakhilina@gmail.com

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 339–346. (ISSN 1890-9639)
http://www.journals.uio.no/osla

CONTRASTIVE LEXICOGRAPHICAL PERSPECTIVES ON THE RUSSIAN LANGUAGE

OLGA IVANISHCHEVA
Murmansk State Pedagogical University

ABSTRACT

The article is devoted to the problems of the representation of culture elements in a bilingual dictionary. It is shown that cultural information concerning realia is necessary in dictionary definitions. Facts about different types of realia may be included as additional information after the actual translation.

[1] INTRODUCTION

The present article is devoted to the problem of the representation of culture elements in a bilingual dictionary.

The research hypothesis is as follows: during lexicographical processing of culture elements in a bilingual dictionary, the selection of information should be defined by cognitive and contrastive principles in order to determine the knowledge of a dictionary user and help the people of two different cultures understand each other. In connection with the hypothesis put forward, we seek to develop a theory of representation of cultural information in a bilingual dictionary.

The article focuses on the following issues:

- The need for the representation of cultural information in bilingual dictionaries in order to provide additional and necessary information to the translation.
- The strategies concering the inclusion of cultural information, as additional commentary relevant to the translation in bilingual dictionaries.

[2] CULTURE ELEMENTS

By culture elements relevant for a bilingual dictionary, we understand the so-called realia, i.e. a culture-specific vocabulary, that is, a vocabulary with a cultural component of meaning, lexical and phraseological units denoting realia specific for a certain culture, such as artifacts, aspects of social life, and historical events. In other words, a kind of vocabulary which demands further commentary because the denotatum (object) is highly specific for a certain culture.

In our research, phrases denoting facts of non-verbal behaviour are also included in the group of such realia, e.g. ни пуха ни пера (literally: neither down, nor a feather; Good luck); не сглазить бы (Knock on wood!); с первым апреля (April Fools). We consider nomenclature names (e.g. валерьянка (valerian drops), индийский чай (Indian tea), ливерная (liverwurst), докторская колбаса (literally: doctor's sausage), cigarettes "Казбек" (Kazbek) and "Беломорканал" (Belomorkanal), shops "Детский мир" (literally: Children's World, Detskij Mir)) and phrases with cultural connotations as a separate group (e.g. десять лет без права переписки (ten years of prison without the right to write letters), московские кухни (Moscow kitchens), пятая графа ("the fifth column")) as a separate group.

[3] DICTIONARY USER

One of the most remarkable tendencies of modern lexicography is the appeal to the dictionary user. As repeatedly mentioned in linguistic research, the issue of creating an active dictionary is becoming more and more important. For example, L.V. Shcherba emphasizes that it is necessary to have two explanatory translation dictionaries for each pair of languages: for Russians with explanations in Russian and for non-native Russian speakers with explanations in corresponding languages. According to L.V. Shcherba, these four dictionaries would allow a non-native speaker to read and understand books in the original language, as well as to grasp the authentic meaning of foreign words (Shcherba & Matusevich 1993, 7).

The main task of a passive dictionary is to provide word meanings. In certain contexts, the reader can find the necessary translation on his/her own using her native language. In an active dictionary, the reader does not search for the explanation of the word to be translated (he/she knows it already) but for instructions helping him/her to find an exact equivalent in the target language (Gak 1995, 53).

The issue concerning the "amount of cultural "connotation" in a bilingual dictionary" has repeatedly been brought up by V.P. Berkov (see e.g., Berkov (1975) and Berkov (2004)). Berkov points out that all serious dictionaries include some linguistic and cultural knowledge. But the way this knowledge is represented is not systematic. Some cultural phenomena should be extensively commented on, including the differences and similarities which exist in different cultures (Berkov 1975, 418).

Bilingual dictionaries usually do not include any cultural information. See, for example, the words relating to the realia of Russian history and culture in Russian-Foreign language dictionaries: Бурлак (Burlak) – *pramdrager* (Berkov 1994); *barge hauler* (Smirnickij 1991); *Treidler* (Lejn 1989); Белогвардеец (Belogvardeec) – *vitgardist* (Davidsson 1992); *hvitegardist* (Berkov 1994); *White Guardsman* (Ozieva et al. 1995); *White Guard* (Smirnickij 1991); *Weißgardist*, *Konterrevolutionär* (Lejn 1989).

The absence of information in bilingual dictionaries often precludes a non-native speaker from understanding Russian texts in a correct way. Consider the following example:

(1) Создана цепкая система переливающихся сосудов – личных связей, взаимной выручки, сплетенных между собой интересов. Гражданский и Уголовный Кодексы не стоят и гривенника, если раздастся звонок из *горкома или райкома.* Либо из некоей грозной конторы.
(Зорин Л. Трезвенник, Знамя. 2001(2))

Such texts raise questions non-native speakers cannot find answers to in dictionaries. So, from the definition of the word горком or its translation (*bykomité* – town committee (Berkov 1994)), it is not clear why the integrity of the law collapses when a call comes from the town committee.

We have a similar example with one other type of realia of the Soviet times, namely the magazine Юность (Junost'), which features in N. Baranskaja's novel "Week after Week", translated into Swedish by K. Hansson. The translator translates the name of the magazine Юность as *tidskriften "Junostj"* without explaining the role this magazine played for a Soviet reader (N. Baranskaja wrote her novel in 1969):

(2) Когда мы утрясаемся немного, мне удается вытащить из сумки "Юность". Читаю давно всеми прочитанную повесть. (Baranskaja 1981)
Jag läser Aksjonovs berättelse "Tomma tunnor", *som varenda människa läst för länge sedan.* (*article's author emphasis*). (Swedish translation by Hansson)

The meaning of the original sentence remains unclear for a non-native speaker. Literally, the Swedish translation says: "I am reading Aksenov's novel "Zatovarennaya bochkotara" ("Overstocked cask") (*in Swedish literally "Empty casks"*), which has already been read by everybody".

The translator uses the author's name and the name of the novel in the text because of the role the magazine Юность played in the social and cultural life of those times and because of the importance of the year in which Aksenov's novel emerged (i.e. 1968). But this does not reveal the meaning the magazine and the novel have for a Russian native-speaker and moreover for a contemporary of N. Baranskaja. For example, it is difficult to explain to a non-native speaker that even mentioning Aksenov's name could cause problems with publishing a novel. Creating a system of methods of lexicographic analysis demands a theoretical approach to the problem of representation of culture elements of this kind in bilingual dictionaries, at the same time the volume of this information must obviously be kept at a minimum.

The selection of cultural information to be included in a bilingual dictionary should be guided by the following cognitive and contrastive principle: a bilingual translation dictionary ought to explain the things that a native speaker knows, but not to explain the things that a dictionary user (i.e. a non-native speaker) knows.

[4] SWEDISH-ENGLISH-RUSSIAN EXAMPLES

The Swedish word "dagmamma" means a kindergarten teacher, while "dagbarn" can be defined, according to (Malmström et al. 1994), as a child ("dagbarn" literally is a "day child") left in another family for the day under the supervision of a teacher while its parents are at work. For a Russian native-speaker, it is not clear what the terms are which regulate the child's stays in the family and how much the family is paid for doing that. While searching for an English translation, one can find out that "dagmamma" (eng. baby-minder, child-minder) is a person paid for looking after a child more often at his/her home while both parents are at work. In Great Britain, designated services make sure that the premises are acceptable and the child-minder meets the necessary standards (DELC 1992). The dictionary (Apresjan & Mednikova 1999) suggests that "baby-minder" has the following meaning: приходящая няня (часто школьница или студентка), остающаяся с детьми за плату. The meaning of the word "child-minder" is: приходящая няня или 1) няня в яслях; 2) воспитательница (детского сада).

So, for example, for an English native-speaker "dagmamma" is a familiar concept and can be translated, though it is necessary to specify that "dagmamma" is a nurse (private or municipal). A municipal nurse is a woman employed by a municipality to organize a family kindergarten ("familjedaghem" in Swedish). The word "kindergarten" has a different meaning here since a "familjedaghem" is organized for children from 6 months to 12 years during the time the parents are at work or are studying and the children are not at school (*Juridik till vardags* 1993: 410–411). For a Russian-speaking dictionary user (there is no such word in Russian (Milanova 1992)), the word belongs to the group of realia without equivalents. It can hardly be called culture-specific because working mothers in Russia have always been helped by grandmothers or nurses who either came home to a child (so-called "coming babysitters"), or a child was taken home to them because of the reduced quantity of kindergartens. Thus the Swedish "dagmamma" cannot be considered as a culture-specific word or a word with background meaning (Denisova 1978). Distinctive features of realia should be given as comments taking into account the level of a non-native speaker's knowledge.

In this paper, the contrastive principle of material selection assumes that any information concerning realia will contain nationally specific data and their specificity in a bilingual dictionary should be determined by comparing such data to another language or another culture.

In the works on translation theory, one can observe that the typological analysis of the lexicon on the basis of the componential analysis is a widely used research method when dealing with the "cultural" lexicon. The method compares semes present in one equivalent and absent in another, a method which was established in the works of N.I. Tolstoy, see (Tolstoj 1997, 21).

In our approach, we also take into account the background knowledge of an average native speaker, i.e. a native-speaker of a given language and given culture with secondary education.

In lexicographic theory connected to translation and teaching of foreign languages, the concept of reconstruction of background knowledge has a different meaning. The main task of bilingual lexicography with regard to the culture-specific lexicon consists in selecting the minimum of information (provided as a comment) necessary to understand and use the words of another language. The problem of defining the content of such a comment and consequently of its minimization is in fact the problem of choosing the distinctive features that should be mentioned in a bilingual dictionary.

It is obvious that only the features that distinguish the given object or that are very important for the public life should be represented in a bilingual dictionary. Not everything that a native speaker knows about an object or that is associated for him/her with this object is relevant for a non-native speaker.

It is to be emphasized that, for example, definitions of a Russian word in a Russian-German dictionary will be in many cases different when compared to the definition of the same word in a Russian-Polish or a Russian-Vietnamese dictionary (Sternin 1992, 215).

[5] RUSSIAN-SWEDISH-NORWEGIAN-CZECH EXAMPLES

Our analysis of the Russian-Norwegian (Berkov 1994), Russian-Swedish (Davidsson 1992) and Russian-Czech (Vlchek 1985) dictionaries has shown that a Russian culture-specific word requires different comments in different dictionaries, i.e. various explanations to translation should be present in a dictionary definition.

For example, Baba-Yaga (Баба-яга) as a character from fairy tales is translated as Baba-Yga, which requires certain comments for Swedish and Norwegian native-speakers similar to the one given in (Berkov 1994): "gammel heks i slavisk mytologi; bortfører og fortærer barn, flyr i en morter, utvisker sine spor med en kost, bor i urskogen i en "hytte på hønebein"" (an old witch in Slavic mythology who steals and eats children, flies on a mortar, wipes out her tracks with a broom, lives in the woods in a "hut on chicken legs"). Compare this to (Davidsson 1992): "trollgumma (häxa) i ryska folksagor" (a witch in Russian fairy tales), where the comment does not provide sufficient information for a Swedish native-speaker. A Swedish native-speaker sees "häxa" (witch) as a woman with supernatural abilities which she uses to harm people.

For a Russian native-speaker, Baba Yaga is, according to a standard definition, rather the following: безобразная старуха-колдунья, передвигавшаяся в ступе и заметающая след помелом (хозяйка леса, повелительница его обитателей, вещая старуха, страж входа в царство смерти, живущая в дремучем лесу в избушке на курьих ногах), помощница героя. Thus in a Russian-Swedish dictionary, the comment on translation should be, in our opinion, as follows: "an ugly old woman who reigns over witches and other evil spirits, lives in the woods in a hut on chicken's legs, flies on a mortar, and wipes out her tracks with a broom; she lures heroes of fairy tales (especially children) to her hut where she roasts them in a stove by throwing them into it with a spade".

Other comments are needed for a Czech native-speaker. Knowing that Баба-яга is not only a character in Russian fairy tales but also a hero in Slavic mythology, we can assume that the comment directed at Swedish native-speakers will be redundant for a Czech native-speaker. The only thing that distinguishes the Czech Baba Yaga from the Russian one is that she is wicked and never helps the heroes. So, the comments to the Czech translation can be as follows: "in Russian fairy tales, Baba Yaga is sometimes a hero's helpmate; she favours heros and gives them advice."

[6] CONCLUSION

A contrastive definition analysis, i.e. the comparison of definitions of the same words in explanatory dictionaries of different languages, helps us to select the elements of cultural information necessary in a bilingual dictionary for a certain pair of languages. The analysis is aimed to show how a mastered word, as a sign of realia, reflects the knowledge of native speakers about a foreign culture element.

Culture-specific differences between words can be revealed by means of other methods as well (i.e. sociolinguistic methods, like interrogations; associative experiments; introspection (if the author of the dictionary is a native-speaker of a given language); through analysis of fiction, newspaper and magazine articles).

Realia features can vary greatly, and the selection of concept features can sometimes be subjective. It cannot be explained either by the dictionary type, or by the vocabulary category. Which features of realia should be represented in the comments in a bilingual dictionary? According to V.P. Berkov, a dictionary entry should contain two main components: the description of REALIA and description of its FUNCTION. Sometimes the SYMBOLIC MEANING of realia must also be represented (Berkov 2004, 163).

Taking into account the cultural value of a feature we select, the following elements of realia description should be, in our opinion, present in a bilingual dictionary: a) attributes (appearance, components, traditions); b) historical markedness; c) social status (functions); d) function (purpose, role); e) popularity / unpopularity of realia; f) symbolic meaning.

REFERENCES

Apresjan, Ju.D. & Je.M. Mednikova. 1999. *Novyj bol'shoj anglo-russkij slovar': v 3 t.; 3-e izd.* Moskva.

Berkov, V.P. 1975. Slovar' i kul'tura naroda. In *Masterstvo perevoda*, vol. 10, Moskva.

Berkov, V.P. 1994. *Russisk-norsk ordbok: 2. utgave*. Oslo.

Berkov, V.P. 2004. *Dvujazychnaja leksikografija: Uchebnik; 2-e izd.* Moskva.

Davidsson, K. 1992. *Russko-shvedskij slovar'*. Stockholm.

DELC. 1992. *Longman Dictionary of English Language and Culture*.

Denisova, M.A. 1978. *Lingvostranovedcheskij slovar'*. Moskva: Narodnoe obrazovanie v SSSR.

Gak, V.G. 1995. Dvujazychnyj slovar' aktivnogo tipa. In *Jazyk – sistema. Jazyk – tekst. Jazyk – sposobnost'*, Moskva.

Lejn, K. 1989. *Russko-nemeckij slovar' (osnovnoj): 10-e izd.* Moskva.

Malmström, S., I. Györki & P. Sjögren. 1994. *Bonniers svenska ordbok*. Stockholm.

Milanova, D.Je. 1992. *Shvedsko-russkij slovar'*. Stockholm.

Ozieva, A., O. Stott & M. Hepburn. 1995. *Collins Russian-English English-Russian Dictionary*.

Shcherba, L.V. & M.I. Matusevich. 1993. *Russko-francuzskij slovar': 14-e izd.* Moskva.

Smirnickij, A.I. 1991. *Russko-anglijskij slovar': 16-e izd.* Moskva.

Sternin, I.A. 1992. Nacional'naja specifika znachenija slova i leksikografija. In V.V. Dubichinskogo, S.V. Vakulenko & A.I. Danilenko (eds.), *Sovremennye problemy leksikografii*, Har'kov.

Tolstoj, N.I. 1997. *Slavjanskaja leksikologija i semasiologija, izbr. tr. t.1*. Moskva.

Vlchek, J. 1985. *Russko-cheshskij slovar': 3-e izd.* Moskva.

AUTHOR CONTACT INFORMATION

Olga Ivanishcheva
Department of Intercultural Communications and Language Theory
Murmansk State Pedagogical University
Russia
oivanishcheva@gmail.com

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 347–364. (ISSN 1890-9639)
http://www.journals.uio.no/osla

THE PHONE MAKES US SCREAM: CORPUS STUDY OF ENGLISH AND RUSSIAN

JULIA KUZNETSOVA
University of Tromsø

ABSTRACT

This study explores English and Russian speech verbs with phone prepositional phrases (PPs). It investigates two hypotheses: 1) A phone PP produces an independent construction and 2) A phone PP can be freely added to any speech verb. Two constructions in English and two constructions in Russian are used as the material for the analysis. In both languages I explore the most generally used phone PP and compare it with a PP meaning 'speak into the phone'. I present a new method – statistical profiling, that explores which words occur in a slot of a construction most frequently and how that frequency list for a slot is changed if another slot is filled. This paper shows that English *on the phone* phrase can freely be added to any speech and sound verb, while other phone PPs produce different phone constructions.

[1] INTRODUCTION

This study shows how one small and supposedly insignificant PP can completely change the distribution of the verbs used with it. I explore this question on the example of speech and sound verbs that can be used with phone PPs in English and Russian. In both languages I investigate the most generally used PPs (*on the phone* in English (1) and *po telefonu* 'speak on phone-DAT' in Russian (2)) and compare them with the PPs with preposition *into* (*into the phone* in English (3) and *v telefon* 'into the phone-ACC' in Russian (4)). The structure and examples of those PPs are given in Table 1 on the following page. For the purposes of this paper I consider a verb to be a speech verb if it means a sound that can come out of a person's mouth. Hence, I explore verbs like *breathe* or *sigh* that are usually not considered to be speech verbs. Verbs that denote an act of communication such as *say, speak* or *talk* are referred to as neutral speech verbs and are opposed to the verbs that introduce some additional information about the character of communication such as English *shout* and *whisper* or Russian *zagovorit'* 'start talking'.

(1) It feels wrong to sit in my pajamas TALKING ON THE PHONE with a U.S. attorney in D.C., sounding tough about a criminal he's trying to put away. [Huston, James W. Marine One (2009)]

PP	Examples
1) on + NP	speak on the phone
2) into + NP	speak into the phone
3) po + NPdat	govorit' po telefonu 'speak on phone-DAT'
4) v + NPacc	govorit' v telefon 'speak into phone-ACC' govorit' v trubku 'speak into receiver-ACC'

TABLE 1: Phone PPs in Russian and English

(2) Prezident SŠA Buš 12 minut GOVORIL PO TELEFONU s
president USA Bush 12 minutes talked on the-phone-DAT.SG with
prem'er-ministrom Slovakii Mikulašem Dzurindoj.
prime minister Slovakia Mikuláš Dzurinda
'President of USA Bush for 12 minutes TALKED ON THE PHONE with the prime minister of Slovakia Mikuláš Dzurinda.'
[Janina Sokolovskaja. Ljaščuk idët v Irak. Ukrainskie voennye gotovy k otpravke v Persidskij zaliv (2003) //«Izvestija», 2003.02.26]

(3) One guy called up and just SCREAMED INTO THE PHONE, no contaminated blood!
[One-Horned Unicorn Deer Found in Italy. The Bryant Park Project 8:00-9:00 AM. (2008)]

(4) Počemu ty togda ne skazala? — ZAKRIČAL ON V TRUBKU.
why you then not say? — shouted he into receiver
Why didn't you say it then? — SHOUTED HE INTO THE RECEIVER.
[Olga Zueva. Skaži čto ja tebe nužna . . . // «Daša», Nr. 10, 2004]

It seems that almost any speech or sound verb can be used with phone PP. Such uses raise the interesting theoretical question of whether this PP can be freely added to any speech verb. On the one hand this PP is not always used when a speech verb is used: usually use of a verb like *govorit'* 'talk' does not imply speaking into a phone. This argument appears only if the situation described by the sentence is suitable, i.e. includes a phone. On the other hand when it does appear it is semantically connected to the speech verb – it describes the channel for the movement of sound. Thus it remains unclear whether this PP can be freely added to a speech verb or these are realizations of a special phone construction. These two hypotheses will be evaluated based on statistical profiling.

[2] STATISTICAL PROFILING

Statistical profiling uses Construction Grammar (Fillmore 1989; Goldberg 1995, 2006; Tomasello 2003; Fried & Boas 2005) as its theoretical background. Construction Grammar is a theoretical approach that aims to account for various language phenomena using constructions. A construction is a pairing of a form and a meaning; a construction consists of several elements and has semantic restrictions on them. This approach can be illustrated with the example of the *there*-construction discussed in a recent book by Kuno & Takami (2004). Their study offers a constructional account for several phenomena in English that are usually considered to be connected with Unaccusativity. The authors show that the Construction Grammar approach gives better predictions about the data. For example Kuno & Takami (2004, 58) propose the following list of functional restrictions on the *there*-construction: "The *there*-construction is acceptable to the extent that the string to the left of its logical subject is interpretable as denoting existence, absence, appearance, or non-appearance of the logical subject referent. In addition, when the construction has a presentational force, the existence, absence, appearance, or non-appearance that the construction represents must be observable to the speaker (or the person whose point of view the speaker is representing)."

These restrictions allow the authors to explain some uses of *there*-sentences that contradict the Unaccusativity approach. First, it becomes possible to explain why transitive verbs can be used in the *there*-construction, see (5). Even though the verb *cross* is transitive, *cross someone's mind* denotes an event of appearing. Second, it explains why *there*-sentences with some unaccusative verbs are not grammatical (6), however a slight change in the sentence makes them grammatical (7). Addition of a locative phrase transforms how the situation is observed. The locative phrase and the verb together serve to denote the existence of the referent. Third, it explains why sometimes a change in a grammatical form affects the grammaticality of a *there*-sentence, see (8). "[T]he progressive form, since it describes an on-going action or event, establishes the speaker as a spectator of the action or event, and this fact in turn contributes to the 'existence' interpretation of the string to the left of the logical subject." [ibid: 53].

(5) There crossed her mind a most horrible thought.
(Kuno & Takami 2004, 21b from Kayne (1979))

(6) *There smoldered a flag in a corner of the room.
(Kuno & Takami 2004, 22a)

(7) In a corner of the room there smoldered a flag that some angry patriot had torn down and ignited. (Kuno & Takami 2004, 23a)

(8) a. *There swam in the river a young girl with a red headband.
(Kuno & Takami 2004, 46b)

b. There was swimming in the river a young girl with a red headband. (Kuno & Takami 2004, 45b)

Thus we see that the constructional approach in the case of *there*-sentences has an explanatory advantage that the unaccusativity approach lacks.

In investigating a construction, the relevant questions are what is the form of the construction, what is the meaning of the construction and what semantic restrictions does a construction have on its slots. While the first two questions are often investigated in the literature on construction grammar, the issue of the semantic restrictions on a slot is less studied. However, the restrictions posed on the whole construction and on its elements are an important part of a construction, because without knowing what restrictions a construction has we cannot explain grammatical and ungrammatical uses of the construction. This paper offers an objective method to find such restrictions using statistical methods – statistical profiling.

Statistical profiling is not the first attempt to apply statistical methods in construction grammar. S. Gries and A. Stefanowitsch have developed a statistical approach called collostructional analysis, which measures the attraction and repulsion of a lexeme for a slot of a construction (Stefanowitsch & Gries 2005, 2003; Gries & Stefanowitsch 2004). For example, Stefanowitsch & Gries (2005) discuss which lexemes are attracted and repulsed in the causative *into*-construction. Using the frequencies of two lexemes filling different slots in the construction (for example *fool* and *thinking*), they can predict what frequency their pairing would have if these events were independent. Comparing that prediction with the actual frequency of the pair, the authors make a conclusion about the attraction or repulsion of the two lexemes in the construction. Two lexemes are attracted if the actual frequency is higher than the prediction. Two lexemes are repelled if the actual frequency is lower than the prediction. For example, Stefanowitch & Gries show that *fool into thinking* occurs much more frequently than *fool into V-ing* and *V into thinking* would predict.

The semantics of some frames coincides with the semantics of the construction and elements of such frames are attracted to the construction, while some pairs of verbs do not constitute a suitable frame and as a result are repulsed from a construction. Tricking somebody into believing into something is a well formed idea in the mind of the speaker of English and therefore the instances of this frame such as *fool into thinking* or *mislead into believing* appear at the top Stefanowitch & Gries' list of attracted lexemes. On the other hand, physical aggression is an ineffective way to change someone's mind, and as a result we see that items reflecting this frame such as *force into thinking* or *bully into believing* are repulsed from the construction. Thus a collostructional analysis uncovers the semantic structure of the examples of a construction.

However collostructional analysis has several disadvantages. First, this method has a strong preference towards idiomatic use, for example in the *of* -construction the sure winner is *cup of tea* which definitely is an example of idiomatic use, and therefore does not provide much information about semantic restrictions on a slot. Second, the most frequent words such as *do, talk* or *walk* disappear from the list of attracted constructions, since they are usually not attracted to a construction with a specific meaning, such as for example causation. For instance, the verb *talk* is not in the list of the verbs attracted to a verbal slot of the *V on the phone* construction. This is a minus because even though these frequent verbs are not attracted to a slot, among all examples of a construction they appear frequently due to their overall frequency. As we know from experimental studies conducted by Goldberg (2006), the items that appear in a slot frequently contribute to our understanding of a construction. The most frequent items appearing in a slot give us information about the most neutral possible filler for the slot, thus, cutting these verbs we lose important information about semantic restrictions on the slot.

Statistical profiling, like collostructional analysis, investigates correlations between lexical items occurring in two different slots of a grammatical construction. Yet, statistical profiling concentrated on finding semantic restrictions on a slot solves both problems mentioned above: it is not skewed toward idiomatic use, actually idioms never appear in the results of the statistical profiling, and statistical profiling does not exclude the most frequent items, it only measures if these items are repulsed from a slot of a construction. Statistical profiling of the construction is based on the idea that the distribution of the elements in the slot reflects the semantic requirement on that slot. This predicts that if the distribution of elements in $slot_1$ is changed significantly when we fill $slot_2$, we are dealing with an independent construction. To use this method we need to explore which words occur in a slot of a construction most frequently and how the frequency list for $slot_1$ is changed if $slot_2$ is filled.

For example, coming back to the phone PP used with speech verbs, statistical profiling predicts that the phone construction should have specific semantic requirements on its elements and particularly on the verb in it. As a result of the semantic requirement, the distribution of verbs possible in the construction has to be different (and the difference is statistically significant) from the distribution of those verbs in general in the corpus, i.e. after filling $slot_2$ with the phone the distribution of the verbs in $slot_1$ is changed. On the other hand if these PPs can be added freely to a speech verb, then the distribution of the verbs with the phone PP should be similar to the distribution of the verbs without the phone PP, i.e. filling $slot_2$ with the phone does not affect distribution of elements in $slot_1$. The case studies below show the use of this approach to the speech verbs with the phone PPs. For each of the phone PPs there will be a choice between two alternative hypotheses: 1) A phone PP produces an independent construction and 2) A

phone PP can be freely added to any speech verb.

[3] DATA

English and Russian data for this study is collected from corpora. English data and examples for this study are collected from the Corpus of Contemporary American English (CoCA[1]), which consists of 385 million words. Russian data and examples for this study are collected from the Russian National Corpus (RNC[2]), which consists of 140 million words. Table 2 shows how many occurrences of each phone PP were found in the corpus. A speech verb that appears with a PP five or more than five times is included in the list of top speech verbs for that PP. Table 2 also shows how many top speech verbs are found for every phone PP. It can be seen that usually there are eight or nine top speech verbs for a PP, but the PP *v telefon* 'into the phone' has noticeably less top speech verbs – only five. To make the data for the Russian PP v+NPacc more comparable with data for other phone PPs I explored an additional variant of this PP: *v trubku* 'into the receiver', which has twelve top speech verbs.

PP	All occurences	Top speech verbs
on the phone	7230	8
into the phone	507	9
po telefonu 'on the phone'	2049	9
v telefon 'into the phone'	193	5
v trubku 'into the receiver'	272	12

TABLE 2: Top speech verbs with phone PPs

[3.1] *On the phone*

This section applies statistical profiling to the speech verbs with phone PP *on the phone*. All verbs that appeared in the context of PP *on the phone* in the CoCA are collected (7230 examples). Table 3 on the next page shows the eight verbs that appeared more than four times in this small subcorpus. The column labeled CORPUS shows how many examples of this verb are found in the corpus. The column labeled PREDICTED gives us the number of examples that would occur before the PP, if that distribution were similar to the distribution in corpus. The column labeled OBSERVED shows how many examples of that verb are found in the context of PP *on the phone*. The numbers in the column PREDICTED are calculated using the following mechanism. The eight speech verbs are used most frequently with *into*

[1] http://www.americancorpus.org
[2] http://www.ruscorpora.ru

the phone are taken (990 examples). The number of occurrences for the same verbs in the corpus is calculated (390 370 examples). For each verb the percentage of its occurrences in the corpus is calculated. For example, for the verb *talk*, which occurs in 256 892 examples in the corpus, this percentage is 65%, since 256 892 is 65% of 390 370. Thus if the distribution with *on the phone* were the same as in the corpus it would occur in 651 examples (65% of 990 examples), however it actually occurs in 741 examples, as can be seen from the column labeled OBSERVED.

VERB	CORPUS	ON THE PHONE -PREDICTED	ON THE PHONE -OBSERVED
talk	256 829	651	741
speak	80 590	205	140
cry	23 139	59	33
sound	14 929	38	29
chat	7 594	19	26
gab	109	0	9
whisper	2 714	7	6
yell	4 466	11	6
TOTAL	390 370	990	990

TABLE 3: Top speech verbs with the PP *on the phone*

The semantic field of speech and sound in the CoCA is dominated by two verbs: *talk* and *speak*, as the left pie chart of Figure 1 on the following page shows. The same verbs dominate with phone PP *on the phone*, as can be seen from the right pie chart of Figure 1. The two charts in Figure 1 show that the distribution of the speech and sound verbs with *on the phone* is similar to the distribution of these verbs in the corpus. While the chi-square test shows that the difference is statistically significant[3] ($\chi^2 = 51.6, 6d_f, P = 2E-09$), the effect size index[4] $w = 0.22$ shows that the size of the effect is small. Thus the second hypothesis is confirmed: English *on the phone* phrases can attach freely to any speech verb. The most frequent speech verbs are frequent in this construction and vice versa. Therefore the PP *on the phone* does not add a lot of specific information and does not pose additional semantic requirements on the verbs used with it.

[3] For this test and all the tests below, both the chi-square test and the calculation of the effect size effect index w are performed using only those verbs which have more than 5 predicted occurrences

[4] The effect size index w for goodness-of-fit chi-square test is discussed in (Cohen 1988/1977), $w = 0.1$ is considered small, $w = 0.3$ medium and $w = 0.5$ large effect size. Thus the effect size with the index $w = 0.22$ can be characterized as medium to small. However, we will see below that the noticeable differences result in the effect size being higher than $w = 0.5$, therefore the effect size with the index $w = 0.22$ might be counted as insignificant.

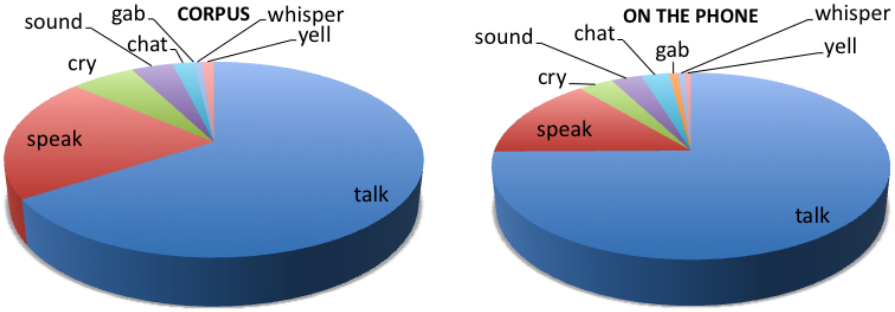


FIGURE 1: Top speech verbs with the PP *on the phone*

[3.2] *Into the phone*

If we look at a similar table for the PP *into the phone*, we see that the distribution in the corpus is as in the previous case dominated by neutral speech and sound verbs: *say, speak* and *talk*. Thus the distribution in the CoCA predicts that these verbs should dominate the distribution with the PP *into the phone*: the column labeled PREDICTION gives a prediction of 209 occurrences of *say*, thirteen occurrences of *speak*, thirty occurrences of *talk* and three or less occurrences of other verbs. However, the distribution of the top speech and sound verbs with PP *into the phone* is noticeably different. Neutral speech verbs such as *say* and *talk* appear less frequently than predicted, while verbs of shouting (*scream, shout, yell, bark*) and verbs of whispering (*whisper, sigh, breathe*) appear more frequently than the corpus predicts.

VERB	CORPUS	INTO THE PHONE −PREDICTED	INTO THE PHONE −OBSERVED
say	1 845 675	209	96
speak	112 668	13	43
scream	21 312	2	29
whisper	18 640	2	22
shout	19 045	2	19
yell	13 531	2	18
talk	262 293	30	17
sigh	13 433	1	9
breathe	23 673	3	7
bark	6 675	1	5
TOTAL	2 336 945	265	265

TABLE 4: Top speech verbs with the PP *into the phone*

The differences between the pie chart on the left and on the right of Figure 2 show that the distribution of the speech and sound verbs with *into the phone* is different from the distribution of the same verbs in the corpus. This difference is statistically significant ($\chi^2 = 135.9, 2d_f, P = 2E - 30$), and the size of the effect is large $w = 0.71$. Thus for this PP the first hypothesis is confirmed: it produces a special phone construction with specific semantic requirements on the verb that can be used in it.

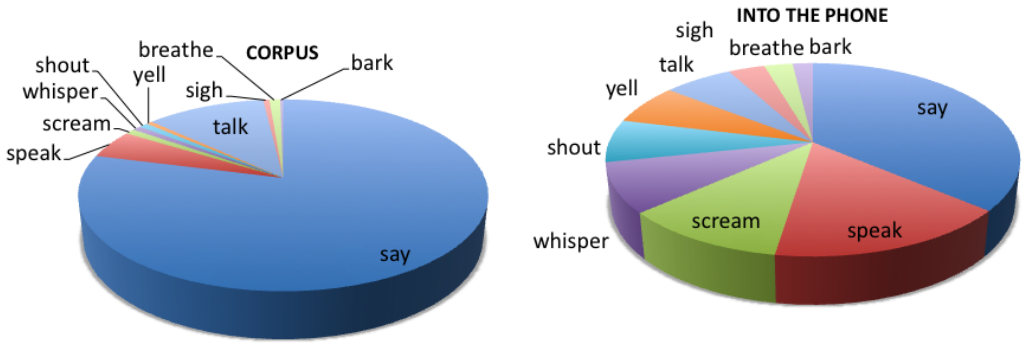


FIGURE 2: Top speech verbs with the PP *into the phone*

[3.3] *Po telefonu 'On the phone'*

If we turn to the Russian analogue of the English PP *on the phone*, we see that the top speech verbs with the PP *po telefonu* 'on the phone' mostly belong to the neutral speech verbs, see Table 5 on the following page. We have seen that in English the PP *on the phone* can freely be added to any speech and sound verbs. Here, the observed numbers of occurrences are noticeably different from what is predicted.

Figure 3 on page 357 shows that the distribution of the speech and sound verbs with *po telefonu* 'on the phone' is different from the distribution of the same verbs in the corpus. This difference is statistically significant ($\chi^2 = 1174.1, 5d_f, P = 1E - 251$), and the size of the effect is tremendous $w = 1.37$. The main difference is in the distribution of most neutral verbs *govorit'* 'talk' and *skazat'* 'say'. While the verb *skazat'* dominates in the corpus, the top verbs with with the PP *po telefonu* are dominated by *govorit'* 'talk'.

This difference is affected by punctuality vs. durativity of an event. The nature of the situation of speaking on the phone presupposes that the situation lasts over a period of time. As a result the verbs that denote protracted, "durative" events such as *govorit'* 'talk' are preferred by this PP, while instantaneous and "punctual" events such as *skazat'* 'say' are dispreferred. Because of this preference all imperfective speech verbs (*govorit'* 'talk', *razgovarivat'* 'converse', *sprašivat'* 'ask', *boltat'* 'chatter', *rasskazyvat'* 'tell', *orat'* 'yell') are used with the PP *po telefonu* 'on

VERB	GLOSS	CORPUS	PO TELEFONU -PREDICTED	PO TELEFONU -OBSERVED
govorit'	talk	44 477	190	286
razgovarivat'	converse	4 587	20	153
skazat'	say	76 397	327	57
pogovorit'	talk for a while	6 248	27	41
sprašivat'	ask	330	1	27
boltat'	chatter	920	4	25
rasskazyvat'	tell	6 656	28	17
vyzvat'	send for	4 224	18	7
orat'	yell	726	3	5
TOTAL	144 565	618	618	

TABLE 5: Top speech verbs with the PP *po telefonu* 'on the phone'

the phone' more frequently than the corpus predicts. Between the two perfective verbs the verb *pogovorit'* 'talk for a while', which has a reference to the period of time added by the prefix *po-* is used more frequently than the corpus predicts and the only punctual perfective verb in the list *skazat'* 'say' is used less frequently than overall. Thus the PP *po telefonu* 'on the phone' produces a new independent phone construction sensitive to the durativity of the event.

[3.4] *V telefon 'into the phone' and v trubku 'into the receiver'*

There are only five top speech verbs with PP *v telefon* 'into the phone'. However even such a small list shows preferences similar to those we observed for its English analogue *into the phone*. While neutral speech verbs such as *govorit'* 'talk' and *skazat'* 'say' show a decrease compared to the prediction, shouting verbs such as *kričat'* 'shout' and *orat'* 'yell' occur with the PP *v telefon* 'into the phone' more frequently than the corpus predicts.

VERB	GLOSS	CORPUS	PREDICTED	OBSERVED
kričat'	shout	28 993	3	25
govorit'	talk	344 097	30	17
skazat'	say	421 203	37	16
otvetit'	answer	2 836	0	7
orat'	yell	4 983	1	6
TOTAL	802 112	71	71	

TABLE 6: Top speech verbs with the PP *v telefon* 'into the phone'

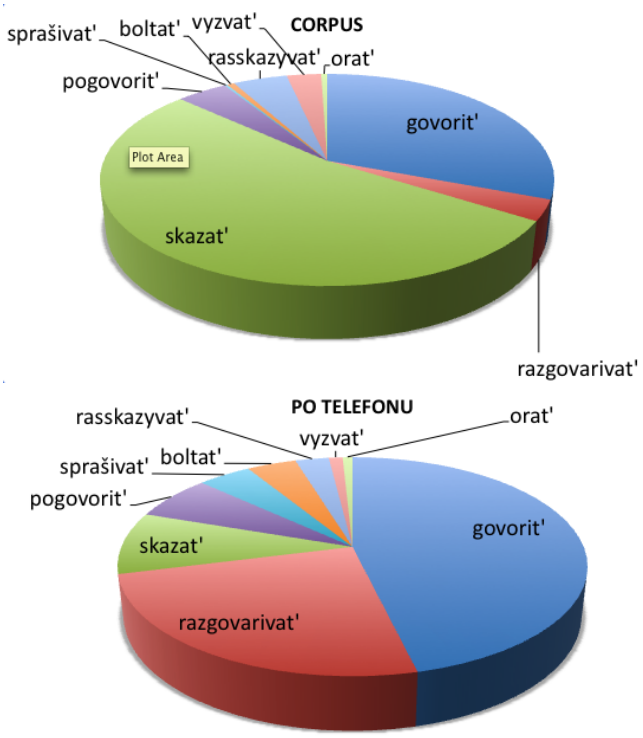


FIGURE 3: Top speech verbs with the PP *po telefonu* 'on the phone'

Even though the five top speech verbs give us some indications about the behavior of this phone PP, there is, as I mentioned above, not enough data for comparison. To make the data on the phone PP with the preposition *v* more representable I investigate an additional example of that PP using *trubka* 'receiver' as a filler for the NP slot. The PP *v trubku* 'into the receiver' has twelve top speech verbs, and clearly presents an expansion of the list of top speech verbs with the PP *v telefon* 'into the phone'. Table 7 on the following page represents the speech verbs that occur with the PP *v trubku* 'into the receiver' more than four times.

It can be seen from Table 7 that there are only eight different roots from which the twelve top speech verbs with the PP *v trubku* 'into the receiver' are derived: *burk-* 'mutter', *govor-* 'talk', *krik-* 'shout', *molk-* 'remain silent', *or-* 'yell', *otvet-* 'answer', *šept-* 'whisper' and *skaz-* 'say'. Six verbs are produced using a bare verb root and verb ending and six other verbs are derived using a prefix or a suffix. Among the suffixes we see *za-* which has an ingressive meaning and is glossed as 'start V-ing' (see Sokolova (2009) and references therein) and *pro-* which has the meaning of producing a quantum and is glossed as 'V something' (see Krongauz (1998) and references therein). The only suffix present in the data is *-nu-*, which has a semelfactive meaning (see Janda & Makarova (2009) and references therein). However, as a speech verb its meaning is close to the quantum verbs:

VERB	GLOSS	CORPUS	PREDICTION	OBSERVED
kričat'	shout	28 993	6	69
skazat'	say	421 203	89	36
govorit'	talk	344 097	72	29
zakričat'	start shouting	13 212	3	16
otvečat'	answer	110 143	23	12
prokričat'	shout something	1 488	0	10
orat'	yell	4 983	1	7
progovorit'	talk about something	483	0	7
prošeptat'	whisper something	5 047	1	6
burknut'	mutter something	1 510	0	5
zagovorit'	start talking	13 505	3	5
molčat'[5]	remain silent	40 581	9	5
TOTAL		908 245	207	207

TABLE 7: Top speech verbs with PP *v trubku* 'into the receiver'

'mutter once' means 'mutter a quantum of information', so in section 4 examining the semantics of these verbs the verb *burknut'* 'mutter once' is grouped with the verbs that mean 'produce a quantum of information'.

The semantic field of speech and sound in the RNC is dominated by three verbs: *govorit'* 'talk', *skazat'* 'say' and *otvečat'* 'answer'. The first pie chart of Figure 4 on the next page shows those verbs which dominate the chart. In contrast, the environment of *v trubku* 'into the receiver' is dominated by a different verb — *kričat'* 'shout'. We can see from Figure 4 that *kričat'* 'shout' occurs with the PP *v trubku* 'into the phone' ten times more often than the corpus predicts. The distribution in the corpus predicts that *kričat' v trubku* 'shout into the reciever' should appear six times, while it is actually found in sixty-nine occurrences.

The two charts in Figure 4 clearly show that the distribution of the speech and sound verbs with 'into the phone' is different from the distribution of the same verbs in the corpus. This difference is statistically significant ($\chi^2 = 725.78, 4d_f, P = 9E - 156$), and the effect size is gigantic $w = 1.87$. Thus the first hypothesis is confirmed: these are examples of use of a specific phone construction with specific semantic requirements on the verb that can be used in it. The most frequent speech verbs are not frequent in this construction and the verbs that are most frequent in this construction are not frequent overall. We see that both Russian and English show tendency for shouting and whispering speech verbs, when the preposition meaning 'into' is involved in phone PP. The interesting question arises as to why these verbs are preferred by such PPs.

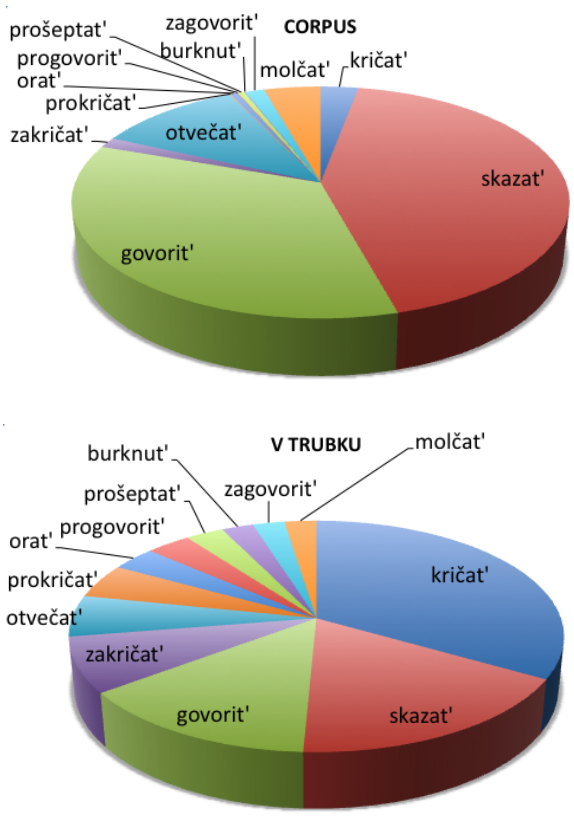


FIGURE 4: Top speech verbs with the PP *v trubku* 'into the receiver'

[4] WHY DO WE SCREAM INTO THE PHONE?

Most neutral verbs from the list of the top speech verbs of the PP *into the phone* – *say* and *talk* – show a decrease compared to the prediction. Therefore, the PP *into the phone* repels neutral speech verbs,[6] see Table 8.

VERB	INTO THE PHONE -PREDICTED	INTO THE PHONE -OBSERVED
say	209	96
speak	13	43
talk	30	17

TABLE 8: Neutral verbs with the PP *into the phone*

Among the verbs that can be used with *into the phone* the most prominent items

[6] Yet, it is interesting, that the neutral speech verb *speak* occurs in this construction more frequently than the corpus predicts. The reason for that should be clarified in future studies.

are non-neutral speech verbs: shouting verbs (like *scream*, *shout*, *yell*, *bark*, see Table 9) and whispering verbs (like *whisper, sigh* and *breathe*, see Table 10).

VERB	INTO THE PHONE -PREDICTED	INTO THE PHONE -OBSERVED
scream	2	29
shout	2	19
yell	2	18
bark	1	5

TABLE 9: Shouting verbs with the PP *into the phone*

VERB	INTO THE PHONE -PREDICTED	INTO THE PHONE -OBSERVED
whisper	2	22
sigh	1	9
breathe	3	7

TABLE 10: Whispering verbs with the PP *into the phone*

For Russian, I have investigated in more details the PP *v trubku* 'into the receiver', which has more top speech verbs and therefore gives us more material for comparison. Most neutral verbs from the list – *skazat'* 'talk', *govorit'* 'say', *otvečat'* 'answer' and *molčat'* 'remain silent' – show a decrease compared to the prediction. Therefore, the Russian PP *v trubku* 'into the receiver' like its English analogue repels neutral speech verbs, see Table 11.

VERB	GLOSS	PREDICTION	OBSERVED
skazat'	say	89	36
govorit'	talk	72	29
otvečat'	answer	23	12
molčat'	remain silent	9	5

TABLE 11: Neutral verbs with the PP *v trubku* 'into the receiver'

Among the verbs that can be used with *v trubku* 'into the receiver' the most prominent items are also non-neutral speech verbs: shouting verbs, whispering verbs and quantization verbs. Shouting verbs are presented by verbs like *kričat'* 'shout', *zakričat'* 'start shouting' or *prokričat'* 'shout something' and whispering

verbs in Russian are presented by only one verb *prošeptat'* 'whisper something', see Table 12.

VERB	GLOSS	PREDICTION	OBSERVED
kričat'	shout	6	69
zakričat'	start shouting	3	16
prokričat'	shout something	0	10
orat'	yell	1	7
prošeptat'	whisper something	1	6

TABLE 12: Shouting and whispering verbs with the PP *v trubku* 'into the receiver'

Quantization verbs can be divided into two classes: first, verbs that mean to 'produce a quantum of information' (like *progovorit'* 'talk about something' or *prošeptat'* 'whisper something') and second, those that mean to 'start speaking' (like *zakričat'* 'start shouting' or *zagovorit'* 'start talking'), see Table 13. Three verbs *prokričat'* 'shout something', *zakričat'* 'start shouting' and *prošeptat'* 'whisper something' belong both in the shouting and whispering class and in the quantization class, and therefore appear both in Table 12 and Table 13. We are not able to see the class of quantization verbs in the English list of the top speech verbs with *into the phone*, because in Russian the quantization meaning is introduced by verbal prefixes and English lacks mechanisms parallel to Russian prefixation.

VERB	GLOSS	PREDICTION	OBSERVED
zakričat'	start shouting	3	16
prokričat'	shout something	0	10
progovorit'	talk about something	0	7
prošeptat'	whisper something	1	6
burknut'	mutter something	0	5
zagovorit'	start talking	3	5

TABLE 13: Quantization verbs with the PP *v trubku* 'into the receiver'

The preference for these verbs can be explained by the nature of the situation of talking into the phone. It is important to note that the person who utters a sentence like (9) is not a participant in the communication, but an observer. That person cannot be the destination point for the message said into the phone. On the contrary, this person is located near the participant who is the source of information in the communication into the phone and observes him or her speaking.

(9) V sosednem kupe poslyšalsja golos, KRIČAŠČIJ V TELEFON.
in next compartment heard voice shouting into phone
In the next compartment there was heard a voice SHOUTING INTO THE PHONE.
[V.P. Kataev. Vremja, vpered! (1931-1932)]

There are three possibilities for the observer to participate in the communication. First, the speaker is talking to the listener, but the speaker is talking too loud and the observer hears it even though he or she might not be interested, as it happens in (9). This type of situation explains the raise in frequency for shouting verbs, such as *shout* or *yell*. Second, the speaker might be aware of the observer and intentionally might want to exclude the observer from the communication on the phone. In this case the speaker would speak in a low voice. This type of situation explains why whispering verbs occur frequently with PPs meaning 'into the phone'. Third, the speaker can participate in two communications at the same time: one with the listener on the phone and one with the observer. In this case it is not clear for whom the pronounced sentence is intended: for the speaker or for the observer. Thus such cases need disambiguation of the channel of communication. However, such disambiguation is not always necessary. If we are dealing with a continuing communication then channel disambiguation is not needed, but if the communication has just started or there has been produced a quantum of communication, then the channel needs to be chosen, because for this new piece of information the intended addressee is not clear. Thus when we use verbs like *govorit'* 'talk' or *skazat'* 'say', which refer to continuous communication, we do not need to mention if that was into the phone or not. However, if we use verbs like *progovorit'* 'talk about something' or *zagovorit'* 'start talking', which denote quantized communication, then we need to specify which channel was used for this communication. As a result, the verbs which mean 'start talking' or 'say a quantum' occur more frequently with PPs *v trubku* 'into the receiver', which is a way to choose the channel.

Thus, the verbs attracted to the English *into the phone* and the Russian *v telefon* 'into the phone' or *v trubku* 'into the receiver' are shouting verbs, whispering verbs and quantization verbs. The preference for such verbs reflect the nature of the situation of communicating into the phone.

[5] CONCLUSIONS

Summing up it can be concluded that English *on the phone* can be added freely to any speech verb. Russian *po telefonu, v telefon (v trubku)* and English *into the phone* cannot be added freely to a neutral speech verb and produce independent constructions. Russian *po telefonu* 'on the phone' has a preference for durative speech verbs. 'Into the phone' in both languages is used as an element of a phone construction that has a preference for shouting and whispering verbs both in Russian

and English, and quantization verbs in Russian.

Statistical profiling used in this study explores the idea that if a slot is filled and that changes the distribution of elements in another slot significantly, then we are dealing with a new construction. Based on the construction grammar approaches, it can be assumed that significant change in the filling of a slot results from any restriction posed on that slot, therefore such change signals that we are facing a new construction that is characterized with new restrictions. Statistical profiling provides a measure of how far the construction has moved on the scale of the syntax vs. lexicon continuum (Croft 2001, 17). In addition it demonstrates what kind of restrictions on the variable the new construction has.

ACKNOWLEDGEMENTS

I would like to thank the anonymous reviewers, whose comments helped to improve this paper.

REFERENCES

Cohen, J. 1988/1977. *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Croft, W. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Fillmore, C. J. 1989. Constructions. A Construction Grammar Approach to Argument Structure. In S. Axmaker, A. Jaisser & H. Singmaster (eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 35–55. University of California, Berkeley: Berkeley Linguistics Society.

Fried, M. & H.C. Boas. 2005. *Grammatical constructions: Back to the roots*. Amsterdam/ Philadelphia: John Benjamins.

Goldberg, A. E. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.

Goldberg, A. E. 2006. *Constructions at Work: the nature of generalization in language.* Oxford: Oxford University Press.

Gries, S. Th. & A. Stefanowitsch. 2004. Co-varying collexemes in the into-causative. In M. Achard & S. Kemmer (eds.), *Language, Culture, and Mind*, Stanford, CA: CSLI Publications.

Janda, L. A. & A. Makarova. 2009. Do It Once: A Case Study of the Russian НУ-Semelfactives. *Scando-Slavica* 66. 78–99.

Kayne, R. 1979. Rightward NP movement in French and English. *Linguistic Inquiry* 10. 710–719.

Krongauz, M.A. 1998. *Pristavki i glagoly v russkom jazyke: semantičeskaja grammatika*. Moscow: Jazyki russkoj kul'tury.

Kuno, S. & K. Takami. 2004. *Functional Constraints in Grammar. On the unergative-unaccusative distinction*. Amsterdam/Philadelphia: John Benjamins.

Sokolova, S. 2009. Ingressive usage of the Russian verbal prefixes *za-* and *po-*: Any spatial features? In *Workshop on Spatial Expressions*, University of Tromsø. May 2009.

Stefanowitsch, A. & S. Th. Gries. 2003. Collostructions: investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.

Stefanowitsch, A. & S. Th. Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1(1). 1–43.

Tomasello, M. 2003. *Constructing a language: a usage based theory of language acquisition*. Cambridge: Harvard University Press.

AUTHOR CONTACT INFORMATION

Julia Kuznetsova
University of Tromsø
Department of Language and Linguistics
NO-9037 Tromsø
Norway
julia.kuznetsova@uit.no

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 365–391. (ISSN 1890-9639)
http://www.journals.uio.no/osla

CONSTRUCTIONAL PROFILE OF THE VERBAL PREFIX ZA-: A COMPARATIVE STUDY OF RUSSIAN AND POLISH

SVETLANA SOKOLOVA & WOJCIECH LEWANDOWSKI
University of Tromsø, Universitat Autònoma de Barcelona

ABSTRACT

In this paper we investigate the semantic and syntactic properties of the prefix ZA- in Russian and Polish against the background of the Locative Alternation. We provide lists of alternating verbs for Russian and Polish and examine their occurrences with the prefix ZA- as presented in the Russian National Corpus and the Polish National Corpus. While the literature on the prefix ZA- mainly offers semantic analysis, we look at the "constructional profile" of ZA-, i.e. the relative frequency distribution of the two locative constructions for each ZA-verb. Our data suggest that there are both similarities and differences in the syntactic behavior of the corresponding ZA-verbs in Russian and Polish, although in general there is a strong tendency for ZA- to be used in the Goal-Object construction. We provide evidence that both constructions are possible with verbs like 'hang' and 'pour', which, according to lexical approaches, should not alternate. Our study fortifies the idea that prefixes and constructions correlate and that the Locative Alternation is construction-driven. The data analyzed also indirectly supports the classification of alternating verbs as "manner", "path" and "hybrid".

[1] INTRODUCTION

In this paper we present empirical data related to several theoretical issues. The first issue involves the semantics of the prefix ZA- in two related languages, Russian and Polish, which are often considered to be similar in their use of verbal prefixes. We investigate the semantics of the prefix ZA- against the background of the second issue, namely the "Locative Alternation", a current issue in research on argument structure.The Locative Alternation phenomenon manifests itself in the use of a verb in two different structures, such as *John loaded the hay onto the truck* (Theme-Object construction) vs. *John loaded the truck with hay* (Goal-Object construction). The question of what motivates such variation still puzzles linguists. Given the list of alternating verbs developed for other European languages, we provide lists of corresponding verbs for Russian and Polish and examine their occurrences with the prefix ZA-. While the literature on the prefix ZA- mainly of-

fers semantic analysis, we focus on both semantic and syntactic properties of ZA- by looking at its "constructional profile" (Janda & Solovyev 2009), i.e. the relative frequency distribution of the two locative constructions for each ZA-verb. The data are extracted from the Russian National Corpus (www.ruscorpora.ru, henceforth RNC) and the Polish National Corpus (www.nkjp.pl, PNC) respectively. Our data suggest that there are both similarities and differences in the constructional behavior of the corresponding verbs with ZA- in Russian and Polish, although in general there is a strong tendency for ZA- to be used in the Goal-Object construction. We provide evidence that both constructions are possible with verbs like 'hang' and 'pour', which, according to lexical approaches, should not alternate. This way our study fortifies the idea that prefixes and constructions correlate and that the Locative Alternation is "construction-driven" (Goldberg 1995, 2006). Moreover, the data analyzed also support Lewandowski's (2009) classification of alternating verbs as "manner", "path" and "hybrid".

We start with a brief overview [2] of the main theoretical issues, namely the semantics of the prefix ZA- in Russian and Polish [2.1] and the Locative Alternation [2.2], placing specific focus on the classification of alternating verbs [2.3]. Our objectives are to test the similarities and differences in the distribution of alternating ZA-verbs in Polish and Russian, and in this way deduce information about the semantics of ZA- in both languages. We present an empirical study of our data extracted from the corpora [3.1] using the method of constructional profiling [3.2]. The analysis [4] addresses the relationship between the semantics of the verb, the prefix and the Locative Alternation. Conclusions are offered in [5].

[2] THEORETICAL ISSUES

Russian and Polish are considered to be similar in the way their verbal prefixes function, especially when opposed to languages that do not have derivational means for expressing aspectual relations. The similarities between the two languages are also emphasized by Dickey's 2000; 2005 classification of aspectual patterns. This section provides the reader with a general outline of the scholarly literature on the semantics of the prefix ZA- in Slavic languages and the Locative Alternation.

[2.1] *The semantics of the prefix za-*

The semantic properties of the prefix ZA- in Russian and Polish are much debated among Slavic linguists. Researchers are concerned with how the prefix is used as both a resultative prefix, referring to an accomplished action, and as an inceptive prefix, marking the beginning of an action or a state (Golovin 1964; Ovčinnikova 1979; Sokolova 1982; Wróbel 1984; Janda 1985; Śmiech 1986; Dickey 2000; Tabakowska 2003; Zaliznjak 2006), etc.). Another property of ZA- is its productivity and frequency in word formation (Pavlova 1988; Čertkova 1996; Łaziński 2008).

The majority of works dealing with the semantics of ZA- in Russian follow either a lexicographical tradition (Golovin 1964; Švedova et al. 1980) or a structuralist approach (Ovčinnikova 1979; Sokolova 1982). Lexicographers describe the derivational types of verbs with the prefix ZA-, some of which are related to each other (Golovin 1964), while structuralists view the semantics of ZA- as a smaller set of unrelated senses, or distinctive features (Ovčinnikova 1979; Sokolova 1982). Similarly to the situation in Russian, the Polish prefix ZA- has traditionally been analyzed as a list of different senses, which can be characterized as either contributing to the verb some kind of Aktionsart information or deriving verbs with a new lexical meaning (Wróbel 1984; Śmiech 1986). The main senses of ZA- in Russian and Polish most frequently distinguished in the literature on verbal prefixation are summarized in Table 1 on the following page.

As can be seen from the overview, the basic meanings of Russian and Polish ZA- appear to be very similar. Additionally, with some predicates ZA- has been considered a pure marker of perfectivity, where the only difference between the unprefixed verb and its prefixed counterpart with ZA- (or Natural Perfective, as in Janda's terminology) is that of imperfective vs. perfective. For Russian, this applies to the resultative use of ZA- as in (6) and some examples of "covering" (3b) and "filling" (4b) (Ožegov & Švedova 2001). In Polish, this is the case of verbs like *zaśpiewać* 'sing-PFV' (14) where, according to (Kurzowa 1997, 17) ZA- is devoid of semantic content.

The structuralist approach was significantly revised in a more recent work by Zaliznjak (2006), where the prefix ZA- is characterized by a certain set of semantic features (or a "conceptual schema"), such as primary BEHIND, IN, EDGE, UP, DEVIA, FAR, and derived COVER, HIDE (from BEHIND) and BECOME, BEGIN, FIX (from IN), which in different combinations appear in verbs with ZA- (Zaliznjak 2006, 311). The semantic type of the ZA-verb is calculated on the basis of semantic features of the prefix, the properties of the unprefixed base verb, and the argument structure of the unprefixed verb. Importantly, Zaliznjak emphasizes the role of the argument structure in defining the semantic type of the ZA-verb, although she does not discuss whether ZA- affects the constructions in which the verb is used. In section [4], we illustrate that not only can ZA- change the basic constructional properties of the unprefixed verb, but it can also provide conditions for alternation.

An alternative approach is presented by Janda (1985, 1986) for Russian and by Tabakowska (2003) for Polish, who describe different meanings of ZA- in terms of a cognitive radial network. Janda points out that although a prefix appears to be semantically fractured, certain submeanings are related to each other and can be presented as a set of configurations (or spatial image-schemas). A configuration consists of a landmark and a trajector which moves in relation to it. Thus, the central configuration for ZA- can be described as the trajector transgressing the

TABLE 1: The main meanings of the prefix ZA- in Russian and Polish.

Meanings		**Russian**	**Polish**
Lexical meanings	moving behind something	(1) *zabežat'* (za ugol) za-run (behind corner-ACC) 'run around the corner'	(9) *zajechać* (za róg) *za-drive* (behind corner- ACC) 'drive around the corner'
	deflection	(2) *zajti* (k drugu) za-walk (to friend-DAT) 'drop by a friend's house'	(10) zajść (do kolegi) za-walk (to friend-GEN) 'drop by a friend's house'
	covering	(3) a. *zastroit'* za-build 'build up an area with new blocks or flats' b. *zakrasit'* za-paint 'cover with paint'	(11) a. *zalesić* za-forest 'plant with trees' b. *zamalować* za-paint 'cover with paint'
	filling	(4) a. *zapolnit'* za-fill 'fill' b. *zaplombirovat'* za-fill a tooth 'fill a tooth'	(12) a. *zapełnić* za-fill 'fill' b. *zaplombować* za-fill a tooth 'fill a tooth'
	getting possession of something	(5) *zaxvatit'* za-seize 'seize, take over, capture'	(13) *zawojować* za-wage war 'conquer'
	reaching natural endpoint	(6) *zaregistrirovat'* za-register 'register'	(14) *zaśpiewać* za-sing 'sing'
Aktionsart meanings	inceptive	(7) a. *zapet'* za-sing 'begin singing' b. *zacvesti* za-blossom 'begin blossoming'	(15) a. *zapłonąć* za-burn 'begin burning' b. *zakwitnąć* za-blossom 'begin blossoming'
	intensive	(8) a. *zaxvalit'* za-praise 'overpraise, to lionize' b. *zagovorit'sja* za-talk Refl 'get excessively absorbed in talking'	(16) a. *zaspać* za-sleep 'oversleep, sleep too long' b. *zagadać się* za-talk Refl 'get excessively absorbed in talking'
	terminative		(17) a. *zajechać* (do wsi) za-go into village 'go into (the village)' b. b. *zaszybować* (do miasta) za-glide into town 'glide into (the town)'

boundary of the landmark and passing into the area outside the landmark (Janda 1985, 29).

In terms of such interaction between the trajector and the landmark it is possible to account for various seemingly unrelated senses of ZA-. Tabakowska shows that the inceptive sense of ZA-, illustrated in (15), is a metaphorical extension of the central spatial sense, in which the landmark is conceptualized as an abstract boundary separating "non-action" from "action" (Tabakowska 2003, 168). The apparently paradoxical terminative sense is based, as Tabakowska demonstrates, on the same notion of "passable borderline" with the difference that this borderline is part of the landmark, expressed as part of the prepositional phrase. On the other hand, the covering meaning relies, according to Tabakowska's analysis, on the so-called sense of curtain (cf. Weinsberg (1973)): the trajector covers the landmark, so that the latter is hidden behind the former and cannot be seen, accessed, etc. If the action of covering extends to the entire landmark, ZA- implies "filling", as in (12).

Quite importantly, both in the traditional approaches (cf. (Golovin 1964; Ovčinnikova 1979; Śmiech 1986)), as well as (although to a lesser extent) in more recent accounts (Tabakowska 2003; Zaliznjak 2006), it is argued that different senses of a given prefix are verb-class specific, i.e. they are compatible with different semantic classes of verbs. Here, it will be shown that the different senses of ZA- are not only verb class-specific, but also construction-specific. In particular, it will be shown that each variant of the locative alternation is associated with a different meaning of ZA-. The next two sections provide the relevant information on the Locative Alternation and its relation to verbal roots.

[2.2] *Different approaches to alternating verbs*

The Locative Alternation is a phenomenon attested in many languages, where a given verb can occur in two alternative constructions (see (18), (19)), both of which deliver approximately the same information:

(18) John loaded the hay onto the truck. (Theme-Object construction)

(19) John loaded the truck with hay. (Goal-Object construction)

The terminology which is used to denote the two constructions is diverse. In this work, we follow Brinkmann (1997); Nichols (2008) and Sokolova et al. (forthcoming) in naming the constructions Theme-Object and Goal-Object as above. The *hay* item stands for the theme and the *truck* item for the goal, while "object" refers to the direct object, which in Russian and Polish is consistently coded with the Accusative case in both constructions.

The works on the Locative Alternation can be divided into three major groups, according to the approach they use: (1) SYNTACTIC/LEXICAL (Rappaport Hovav &

Levin 1988, 2005, 2008; Pinker 1989; Levin 1993; Brinkmann 1997; Dowty 2000) (2) FRAME (Fillmore 1968, 1977, 2008; Boas 2001, 2006) (3) CONSTRUCTIONAL (Goldberg 1995, 2006; Michaelis & Ruppenhofer 2001; Iwata 2005, 2008). Each approach in its own way addresses the question of what motivates the Locative Alternation. The syntactic/lexical approach focuses on the meaning of the verb, treating the syntactic options as secondary to the intrinsic properties of the verb ("content-oriented" or "container-oriented") (Pinker 1989, 125–127). Within existing classifications of verbal roots, it is problematic to account for all the occurrences of the Locative Alternation, since some verbs which are classified as 'non-alternating' can alternate under certain circumstances (see Boas (2006) for some counterexamples). The frame approach takes the syntactic construction as the point of departure, showing how different constructions are related to each other within bigger frames. Words like *load* are split into separate lexical units, depending on the semantic frames they evoke. Yet, the focus on the frame makes it less evident why a single verb alternates between constructions (for more detail, see Sokolova et al. (forthcoming).

According to the constructional approach, the Locative Alternation is an epiphenomenon of the compatibility between the verbal meaning and two independently existing constructions (Goldberg 2006, 40). Our data support Goldberg (1995, 2002, 2006) and Michaelis & Ruppenhofer (2001) in that the Locative Alternation is construction-driven with a reference to verbal semantics. In the next section we will discuss different ways to classify verbal roots with regard to the Locative Alternation.

[2.3] *Different approaches to verbal roots*

Manner vs. Result

It has been widely assumed that verbs from various lexical fields can be classified as lexicalizing manner (e.g. *wipe, float*) or result/path (e.g. *clean, enter*). Levin & Hovav (1991) postulate that manner and result/path are in complementary distribution: a verb can codify either one or the other meaning component, but not both at the same time. Even if some verbs refer to results brought about using a conventionally associated manner or, analogically, some manner verbs specify actions performed to bring about a conventionally associated result, only one of these semantic components is codified in the verbal root. The other can only be expressed outside the verb. For instance, although the action of "wiping" is usually used to clean a surface, the verb *wipe* only denotes a manner of motion (which can be characterized as "surface contact"), whereas the final state of the surface is codified in a separate linguistic unit, i.e. the adjectival phrase *clean* (20):

(20) Pat wiped the table clean.

On the other hand, Talmy (1985, 2000) shows that the manner/path dichotomy is relevant for characterizing crosslinguistic lexicalization patterns. He argues that languages can be categorized as either VERB-FRAMED, such as Romance or Turkish, or SATELLITE-FRAMED, such as Germanic and Slavic. Whereas the former lexicalize the path of motion in the verb and express the manner, if specified, in a secondary element (e.g. a prepositional phrase or a Gerund), the latter codify the manner of motion in the verb, with the Path being relegated to a secondary element, commonly a preposition or prefix, cf. (21) and (22).

(21) La botella entró en la cueva (flotando). [Spanish]
the bottle entered in the cave floating
'The bottle entered the cave (floating).'

(22) The bottle floated into the cave.

Manner and path distinction as a gradient phenomenon

Although it is an uncontroversial fact that there are prototypical manner or path/-result verbs, as Levin & Hovav (1991); Rappaport Hovav & Levin (1998) and Talmy (1985, 2000) convincingly demonstrate, it has been postulated that a strict dichotomy is empirically not justifiable. Lewandowski (2009) classifies verbs entering into the locative alternation in Polish and Spanish as "manner", "path" and "hybrid". Manner verbs, such as Polish *chlapać*, Spanish *salpicar* 'splatter' provide information about how the action denoted by the verb is performed (in this particular case, the liquid is distributed in a wide-spread fashion; cf. Pinker (1989). Path verbs, such as Polish *wieszać*, Spanish *colgar* 'hang', clearly imply displacement of the theme. In contrast, hybrid verbs lexicalize both manner of motion and path. For instance, Polish *ładować*, Spanish *cargar* 'load' implies that a large quantity of items is displaced, but since items are usually loaded into a container, this verb also evokes the trajectory "outside-inside". Lewandowski (2009, forthcoming) shows that the distribution of the alternating verbs in one or the other pattern of the locative alternation is statistically different depending on the root type. Overall, manner verbs tend to appear more often in the Goal-Object construction, and path verbs in the Theme-Object construction, while the distribution of hybrid verbs is similar in both constructions.

In this paper, we follow Lewandowski's classification of alternating verbs and provide evidence that the hybrid nature of verbal roots is also relevant for the grammatical profile of prefixes.

[3] DATA AND METHODOLOGY

Our study examines the constructional profiles of the alternating ZA-verbs in Russian and Polish as evidenced by data from the RNC and the PNC. We first describe how our data were extracted and coded and then present our methodology.

[3.1] *Data. Alternating za-verbs in Russian and Polish*

The list of alternating verbs for Russian and Polish

Following the classification of Lewandowski (2009), we compose a list of alternating verbs for Russian and Polish, which comprise three major groups: "manner", "path" and "hybrid". The lists of such verbs are given below:

Two groups of MANNER VERBS:[1].

a) Wide-spread or undirected distribution of a liquid

Russian:	Polish:
bryzgat' 'splatter'	*bryzgać* 'splatter'
pryskat' 'spray'	*pryskać* 'spray'

b) Contact of a mass against a surface

Russian:	Polish:
mazat' 'daub, smear'	*mazać* 'daub'
	smarować 'smear'

"Manner" verbs do not codify path: you can smear something up, down, to the left, to the right, etc. Thus, path is underspecified in their verbal meaning and they are more likely to be used in the Goal-Object construction.

PATH VERBS:

Russian:	Polish:
klast' 'lay'[2]	*kłaść* 'lay'
vešat' 'hang'	*wieszać* 'hang'
stavit' 'stand'	*stawiać* 'stand'

"Path" verbs imply path, i.e. they denote movement of the theme from one place to another. The prediction will be that they favor the Theme-Object construction.

HYBRID VERBS:

Russian:	Polish:
gruzit' 'load'	*ładować* 'load'
pakovat' 'pack'	*pakować* 'pack'
pixat' 'stuff'	*pchać* 'stuff'
lit' 'pour'	*lać* 'pour'
sypat' 'strew'	*sypać* 'strew'

[1] The characterization of the manner component is taken from Pinker (1989), who claims that alternation does not extend to verbs of "pure manner of motion" such as *pour,* verbs of force exertion *(push, drag, pull, tug, yank)* or verbs of positioning *(lay, place, position, put)* since there is no way to predict on the basis of the verb meaning alone what the effect on the goal argument will be (Pinker 1989, 80).

"Hybrid" verbs are associated with a particular manner of displacing things, but at the same time all of them evoke a container, involving the directionality "outside-inside", cf. Section [2.2]. These verbs can put the focus on both participants and this way have the potential to occur in both constructions.

When used with prefixes, the verbs listed above can behave in three different ways:

(a) Some verbs can alternate when unprefixed. For Russian, this is the case with *bryzgat'* 'splatter', *mazat'* 'smear, daub', *gruzit'* 'load', *pakovat'* 'pack', whereas for Polish it is the case of all the verbs, except *pchać* 'stuff' and the positional verbs *kłaść* 'lay', *wieszać* 'hang' and *stawiać* 'stand'.

(b) Other verbs do not alternate without a prefix and can be used either in the Theme-Object or Goal-Object construction depending on the prefix. For instance, the Russian unprefixed verb *stavit* 'put, place', as well as its Natural Perfective with PO- (*postavit'*), are used in Theme-Object construction, while its perfectives with ZA- and OB- choose the Goal-Object construction (*zastavit'* 'cover something with standing objects'; *obstavit* 'furnish'). In Polish, we find a similar effect with the verbs *pchać* 'stuff', *kłaść* 'lay', *wieszać* 'hang' and *stawiać* 'stand', which appear in the Theme-Object construction without a prefix but favor the Goal-Object construction when used with a resultative prefix like ZA- or OB- (*zapchać* 'stuff, choke', *obłożyć* 'cover by putting things on a surface', *zawiesić* 'cover by hanging things on a surface', *zastawić* 'cover by standing things on a surface, block access').

(c) Finally, some Russian verbs do not alternate when unprefixed but can be used in both constructions with certain prefixes. This is the case of Russian *zalit'* 'pour', *zasypat'* 'strew', *zavešat'/zavesit'* 'hang', *založit* 'lay' with the prefix ZA-.

Examples considered in this study

To the Russian and Polish verbs from the list above, we add the prefix ZA- and investigate how such prefixed versions of the base verbs are represented in the corpora. For the purpose of this study, we used the Modern subcorpus (2000-2009) of the RNC[3], which contains 53 million words, and the PNC, containing 350 million words. For both Russian and Polish, we extracted all occurrences of each ZA-verb from the corpora, excluding passive participles[4], and manually coded the examples as Theme-Object vs. Goal-Object.

[3] We used the modern subcorpus of the RNC since it is most equivalent with the PNC, which includes only modern texts.

[4] The contexts with passive participles require a separate investigation since they present examples of the Locative Alternation where the focus on one of the participants is greater than in the contexts with non-passive forms. However, the general assumption is that the distribution of passive forms between the two constructions resembles the situation with non-passive forms, as far as the choice of the main construction, see Sokolova et al. (forthcoming).

The Theme-Object and Goal-Object constructions differ in which of the participants is marked as the direct object: the Theme (i.e. elements like *hay*), or the Goal (i.e. elements like *truck*). In both constructions in Russian and Polish the direct object is consistently coded with the Accusative case, while the second participant can be expressed via different forms. The Theme-Object construction encodes the Goal via a prepositional phrase with a noun in the Accusative case, as illustrated in (23) and (24). In the Goal-Object construction the second participant is coded by the Instrumental case without a preposition in (25) and (26):

THE THEME-OBJECT CONSTRUCTION

(23) Voditel' zagruzil paket v bagažnik [Russian]
Driver-NOM loaded bag-ACC in trunk-ACC
'The driver loaded the bag into the trunk'

(24) Smuga natychmiast polecił załadować zapasy jarzyn na
Smuga at once ordered za-load provisions vegetables-GEN on
łodzie [Polish]
boats-ACC
'Smuga ordered that the vegetables be loaded onto the boats immediately'

THE GOAL-OBJECT CONSTRUCTION

(25) Tam krasnoarmejcy zagruzili kuzov jaščikami [Ru]
There Red-Army-soldiers-NOM za-loaded truck-bed-ACC boxes-INS
'There the Red Army soldiers loaded the truck bed with boxes'

(26) Chciałem załadować armatę pociskiem (...) i dobić go. [Polish]
I wanted za-load tank-ACC projectile-INS (...) and finish him
'I wanted to load the tank with a projectile and kill him'

The Theme-Object construction focuses on the Theme and on the change of its location. For instance, (23) concentrates on what happens to the bag, i.e. the Theme. On the other hand, the Goal-Object construction marks the change of the state of the Goal, as in (25), which is about the truck bed and how it is loaded.

The two constructions of the Locative Alternation can be represented via full constructions where both participants (Theme and Goal) are overtly expressed, as well as via "reduced constructions", where one of the participants is missing. Most cases with the omitted Theme or Goal argument are instances of ellipsis since the missing participant is perceived from the context. Example (27) below illustrates a Theme-Object construction with a missing Goal:

(27) Tol'ko zagruzit' ugol' budet problematično, poskol'ku iz-za
Just load coal-ACC will-be problematic, since due-to
moroza on prevratilsja v glyby. [Russian]
frost-GEN it-NOM turned-into into blocks-ACC.
'Just getting the coal loaded will be problematic since due to the cold it has turned into blocks.'

In our research we aggregate data from both the full constructions and the reduced constructions. The only examples extracted from our analysis are constructions which can be characterized as hybrid. This means that one of the components of the classical locative construction is expressed by a different form, as in (28) below:

(28) Italija, ešče vcera otkazavšajasja ukryt' u sebja
Italy-NOM only yesterday refuse-ACT NOM to hide by self-GEN
prestupnikov ... segodnja rešila "zamazat'" v nravstvennom
criminals-ACC ... today decided to daub in moral
prestuplenii drugie strany Evropy. [Russian]
crime-LOC other countries-ACC Europe-GEN
'Italy, which only yesterday refused to give shelter to the criminals, today has decided to accuse other European countries of moral crime'

In (28), the first participant is marked by Accusative case, while the second element is in the Locative case. Thus, it is not obvious whether the first element stands for the Theme or the Goal.

[3.2] *Methodology*

The method used in this study is "constructional profiling", i.e. "the frequency distribution of the constructions that a word appears in", based on corpus data (Janda & Solovyev 2009, 367). This methodology is inspired by construction grammar, as it treats the construction as the relevant unit of linguistic analysis (Goldberg 1995, 2006) and implies that speakers are sensitive to the frequency of words in constructions (Goldberg 2006, 46, 62). On the one hand, the constructional profile is a more focused version of the behavioral profile, which involves a wide range of factors (collocational, morphosyntactic, syntactic, and semantic) to investigate synonymy (Divjak 2006; Divjak & Gries 2006). On the other hand, constructional profiling takes the word as the point of departure and in this sense is the inverse of the collostructional methodology, which starts with a construction and investigates what words can occur in it (Stefanowitsch & Gries 2003, 2005).

The constructional profiles of the alternating ZA-verbs in Russian and Polish presented in this study show in which sense the two languages can be treated as similar and in which way they differ.

[4] ANALYSIS

First, we aim to investigate the similarities and the differences in the distribution of the alternating ZA-verbs in Russian and Polish. Second, by comparing the constructional profiles of the ZA-verbs, we analyze the semantics of ZA- in the two languages and the relation between the semantics of the verb, the prefix and the locative alternation.

The data frequencies, collected as described in section [3], are presented in Tables 2 and 3 on the facing page.

Verb	Gloss	Theme-Object construction		Goal-Object construction		Total
		raw frequency	relative frequency	raw frequency	relative frequency	
zapryskat'	'spray'	0	0%	0	0%	0
zabryzgat'	'splatter'	0	0%	28	100%	28
zamazat'	'daub, smear'	0	0%	56	100%	56
zagruzit'	'load'	85	47.8%	93	52.2%	178
zapakovat'	'pack'	13	100%	0	0%	13
zapixat'	'stuff'	63	100%	0	0%	63
zasypat'	'strew'	36	17%	176	83%	212
zalit'	'pour'	59	11.8%	440	88.2%	499
zavešat'/ zavesit'	'hang'	2	8.7%	21	91.3%	23
zastavit'	'stand'	0	0%	5	100%	5
založit'	'lay'	238	96.7%	8	3.3%	246

TABLE 2: Locative Alternation among non-passive forms of Russian alternating verbs.

Given the relative frequencies in Tables 2 and 3 on the facing page , we end up with the following constructional profiles of the alternating ZA-verbs for Russian and Polish (see Figures 1 and 2 on page 378).

As can be seen from Figures 1 and 2, the verbs that favor the Goal-Object construction in both Russian and Polish are: 'splatter' (Russian *zabryzgat'*), 'smear' and 'daub' (Russian *zamazat'* and Polish *zasmarować*, *zamazać*), 'strew' (Russian *zasypat'*, Polish *zasypać*), 'pour' (Russian *zalit'*, Polish *zalać*) and 'stand' (Russian *zastavit'*, Polish *zastawić*). On the other hand, the verbs for 'pack' (Russian *zapakovat'*, Polish *zapakować*) and 'lay' (Russian *založit'*, Polish *założyć*) show a preference

[5] Since only one example of *zabryzgać* 'splatter' was attested in the corpus we will treat it as 0.

Verb	Gloss	Theme-Object construction		Goal-Object construction		Total
		raw frequency	relative frequency	raw frequency	relative frequency	
zapryskać	'spray'	0	0%	0	0%	0
zabryzgać	'splatter'	0	0%	1	(100%)[5]	1
zasmarować	'smear'	0	0%	21	100%	21
zamazać	'daub'	0	0%	33	100%	33
załadować	'load'	471	81.8%	105	18.2%	576
zapakować	'pack'	507	94.9%	27	5.1%	534
zapchać	'stuff'	0	0%	69	100%	69
zasypać	'strew'	1	0.2%	572	99.8%	573
zalać	'pour'	2	0.5%	383	99.5%	385
zawiesić	'hang'	578	99.3%	4	0.7%	582
zastawić	'stand'	0	0%	241	100%	241
założyć	'lay'	789	100%	0	0%	789

TABLE 3: Locative Alternation among non-passive forms of Polish alternating verbs.

for the Theme-Object construction. The differences between the two languages concern the verbs for 'load' (a preference towards the Theme-Object construction in Polish and an even distribution between the two constructions in Russian), the verbs for 'stuff' (Theme-Object construction in Russian and the Goal-Object construction in Polish), and the verbs for 'hang' (the Russian verb favors the Goal-Object construction, while the Polish one is only attested in the Theme-Object construction). It can also be seen that Russian and Polish are characterized by a different number of verbs that show a variation between the two constructions.

The similarities and differences between the alternating ZA-verbs in Russian and Polish are discussed in the subsections below.

[4.1] *General tendencies in the alternating za-verbs in Russian and Polish*

The analysis of the alternating verbs in Russian and Polish shows that the semantics of the prefix is construction-specific. When the Goal-object construction is "headed" by ZA-, the prefix bears the meaning of "covering" (cf. Russian *zabryzgat'*, *zamazat'*, *zastavit'*; Polish *zasmarować*, *zamazać*, *zasypać*, *zalać*, *zastawić*) or "filling" (as in Russian *zagruzit'*; Polish *zapchać*). On the other hand, when ZA- appears in the Theme-object construction, it has the meaning of reaching a natural endpoint (Russian *zagruzit'*, *zapakovat'*; Polish *załadować, zapakować*) or "placing" (Russian *zapixat'*, *založit'*; Polish *założyć*).

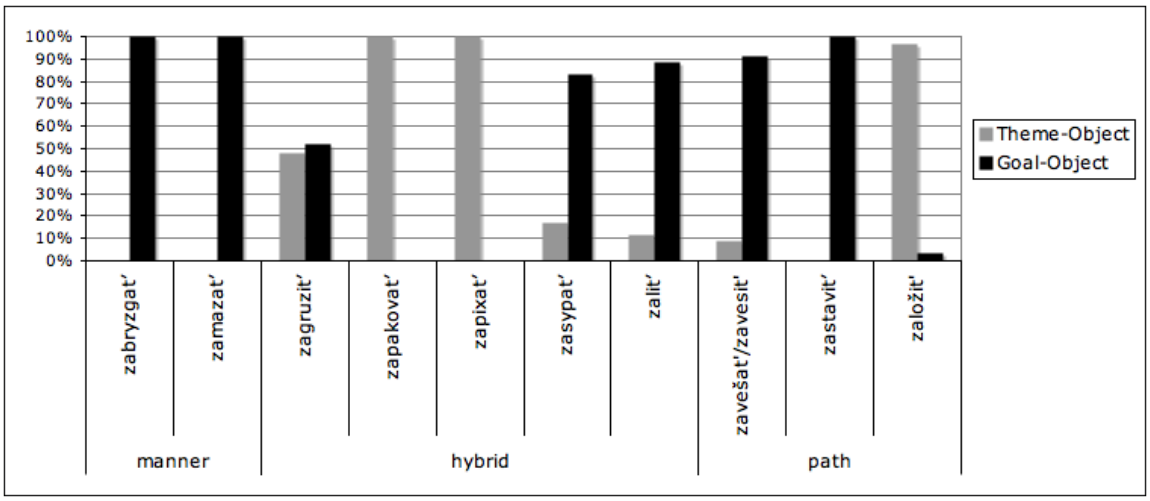


FIGURE 1: Constructional profile of the prefix ZA- in Russian.

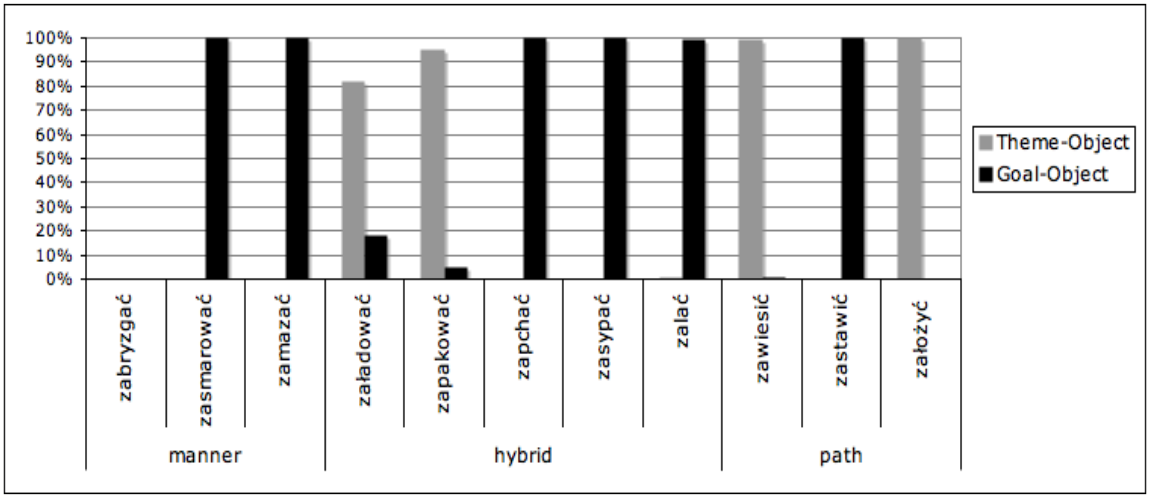


FIGURE 2: Constructional profile of the prefix ZA- in Polish.

As can be seen from Figures 1 and 2 above, in both Russian and Polish, "manner" verbs are clearly associated with the "change of state" meaning. This could be due to the general tendency for verbs with a strong manner component to avoid constructions describing displacement than "hybrid" or "path" verbs. In other words, the manner component "clashes" to a certain degree with the displacement or "change of location" meaning. For instance, in Polish, "manner" verbs, unlike "hybrid" verbs, systematically reject spatial prefixes like W- 'in', DO- 'into' (Lewandowski forthcoming):

(29) "Hybrid" verbs: *wpakować* 'in-pack', *włożyć* 'in-lay', *wlać* 'in-pour', *wsypać* 'in-strew'; *dopakować* 'into-pack', *dołożyć* 'into-lay', *dolać* 'into-pour', *dosypać* 'into-strew'

> "Manner" verbs: **wpryskać* 'in-spray', **wbryzgać* 'in-splatter', **wmazać* 'in-daub'; **dopryskać* 'into-spray', **dobryzgać* 'into-splatter', **domazać* 'into-daub'.

On the other hand, "path" and "hybrid" verbs differ from "manner" verbs in that most of them appear in both constructions (Russian *zagruzit'*, *zasypat'*, *zalit'*, *zavešat'/zavesit'*, *založit'*; Polish *załadować*, *zapakować* and very sporadically *zawiesić*, *zasypać*, *zalać*, where the attested frequency for one of the constructions is below 1%)). The proportion between the Theme-Object and the Goal-Object constructions for "hybrid" verbs depends on the idiosyncratic properties of individual verbs, first of all on the degree to which the Theme is specified. For instance, *zalit'* 'pour' indicates the Theme as a liquid, so the focus is usually placed on the Goal, which is modified. The verb *zasypat'* 'strew' presents the Theme as a dry substance (small objects that are perceived as a mass) with a somewhat higher frequency of the Theme-Object construction than in the case of *zalit'*. The Theme of the verb *gruzit'* 'load' is a single heavy object or a number of separate objects (which are not presented as a mass), which gives more opportunities for profiling either the Theme or the Goal. Thus, the closer the Theme of the verb is to a count noun, the more even the distribution between the constructions appears to be, since the focus can be placed on both participants. However, this hypothesis requires testing on a larger amount of data.

Another common effect for both Russian and Polish is that ZA- seems to be more frequent in the Goal-Object construction, which indicates that there is a certain correlation between the construction and the prefix. In Russian, uprefixed "manner" verbs usually occur in the Goal-Object construction, while "hybrid" and "path" verbs favor the Theme-Object construction. In addition, the "manner" verbs *bryzgat'* 'splatter', *mazat'* 'smear, daub' and the "hybrid" verbs *gruzit'* 'load', *pakovat'* 'pack' can alternate between the two constructions, still showing a stronger preference towards either the Goal-Object construction (in the case of "manner" verbs) or the Theme-Object construction (in the case of "hybrid" verbs). The choice for the Theme-Object or Goal-Object construction within Polish unprefixed "manner" and "hybrid" verbs is similar to Russian, with the only difference that in Polish all unprefixed "manner" and "hybrid" verbs, except for *pchać* 'stuff', potentially alternate (the verb *pchać* 'stuff' and the positional verbs are categorically excluded from the Goal-Object frame when unprefixed). The correlation of the prefix ZA- with the Goal-Object construction reveals itself in two ways: first, ZA- eliminates alternation within the "manner" verbs, strengthening the status of the Goal-Object construction, second, it shifts many "hybrid" and "path" verbs from the Theme-Object construction to the Goal-Object construc-

tion (cf. the verbs *zagruzit'*, *zasypat'*, *zalit'*, *zavešat'/zavesit'*, *zastavit'* in Figure 1 on page 378 and *zapchać*, *zasypać*, *zalać* and *zastawić* in Figure 2 on page 378).

The use of each ZA-verb is generally skewed in favor of one of the locative constructions. The only exception is the Russian verb *zagruzit'* 'load', where the distribution between the Theme-Object and the Goal-Object constructions is almost even (Theme-Object: 47.8%; Goal-Object: 52.2%). A more elaborate analysis of the examples indicates that this could be due to the number of additional metaphorical uses that this verb has in the Goal-Object construction. As shown in Sokolova (2010), of the three prefixed counterparts to the verb *gruzit'* 'load' (with prefixes ZA-, NA- and PO-), *zagruzit'* is more often used metaphorically: *zagruzit'* is characterized by 39% of metaphorical uses, while *nagruzit'* and *pogruzit'* have 25% and 11% respectivly. The major metaphorical extensions of *zagruzit'* involve a "person" (Goal), who serves as the metaphorical CONTAINER, and "information" or "work" (Theme), which represent metaphorical CONTENTS, as shown in (30)-(31):

(30) Ah, vam interesny podrobnosti iz žizni
Oh, you-DAT are-interesting particulars-NOM from life-GEN
zvezd? Radi boga, Andrej Maksimov "zagruzit"
pop-stars-GEN? For god, Andrej Maksimov-NOM za-load-FUT
vas ètoj informaciej. [Russian]
you-ACC this-INS information-INS
'Oh, you are interested in the details of the life of our pop stars? No problem, Andrej Maksimov will provide you with this information.'

(31) Zasedanie Gossoveta po kul'ture zagruzit rabotoj
Meeting-NOM State-Council-GEN on culture-DAT za-load-FUT work-INS
sotrudnikov Minsterstva kul'tury na bližajšie neskol'ko
members-ACC Ministry-GEN Culture-GEN for nearest-ACC few-ACC
let. [Russian]
years-GEN
'The agenda of the State Council on Culture will keep the members of the Ministry of culture busy for several years.'

It is remarkable that in non-metaphorical uses, zagruzit' favors the Theme-Object construction (68.7%), resembling the situation in Polish (see Table 4 on the facing page). However, in metaphorical contexts, it is skewed towards the Goal-Object construction (70.5%).

As can bee seen from Figure 3 on page 382, *zagruzit'* 'load' is the only Russian alternating verb where metaphorical contexts have a crucial affect on the general distribution between the Theme-Object and the Goal-Object constructions because the two contexts behave so differently. It is also the verb that has the

Verb	Extension	Theme-Object construction		Goal-Object construction		Total
		raw frequency	relative frequency	raw frequency	relative frequency	
zagruzit' 'load'	non-metaphorical	57	68.7%	26	31.3%	83
	metaphorical	28	29.5%	67	70.5%	95
zasypat' 'strew'	non-metaphorical	35	20.5%	136	79.5%	171
	metaphorical	1	2.4%	40	97.6%	41
zalit' 'pour'	non-metaphorical	50	11.6%	381	88.4%	431
	metaphorical	9	13.2%	59	86.8%	68
zavešat'/ zavesit' 'hang'	non-metaphorical	2	10%	19	90%	21
	metaphorical	0	0%	2	100%	2
založit' 'lay'	non-metaphorical	115	94%	7	6%	122
	metaphorical	123	99.2%	1	0.8%	124

TABLE 4: The distribution of the two locative constructions among metaphorical and non-metaphorical contexts within Russian alternating verbs that are attested in both constructions.

highest percentage of metaphorical contexts: for *zagruzit'*, metaphorical extensions constitute 53.4% of all the contexts; for *založit'* this number amounts to 50%; *zasypat'* has 19.3% of metaphorical uses, while *zalit'* and *zavesit'* show only 13.6% and 8.7%. The percentage of the metaphorical contexts for the verb *založit'* is also high since most of its Theme-Object usages are represented by frequently used collocations like *založit' fundament/ osnovu* 'lay the foundation' (36.6% of all uses). Cf. example (32):

(32) Vmeste oni zalozili osnovu novogo stilja
Together they-NOM laid foundation-ACC new-MASC.GEN style-GEN
nacional'noj muzyki. [Russian]
national-FEM.GEN music-GEN
'Together, they established a new style for national music.'

Thus, the overall tendency is that alternating verbs prefer one construction over the other, given that there are no other factors interfering, like metaphor.

[4.2] *Differences between the alternating za-verbs in Russian and Polish*

As shown in the previous subsection, the Russian and Polish alternating verbs have a lot in common. Yet, our analysis also points to some differences between the two languages. Firstly, in Russian, more verbs with the prefix ZA- are attested

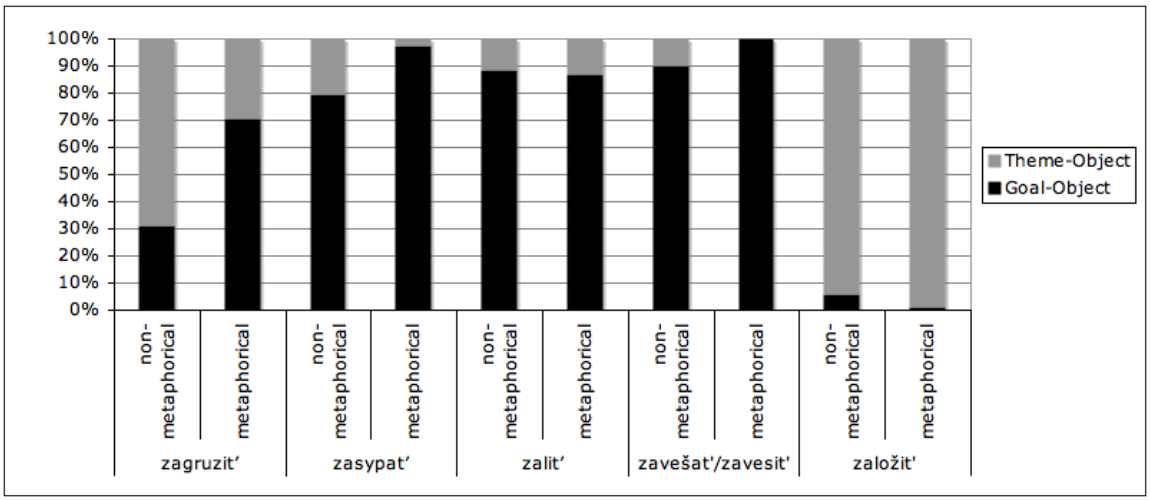


FIGURE 3: The distribution of the two locative constructions among metaphorical and non-metaphorical contexts within Russian alternating verbs that are attested in both constructions.

in both constructions. Leaving aside the Russian verb *zavešat'/zavesit'* 'hang' for which we have little data, we are still left with three verbs (*zagruzit'* 'load', *zasypat'* 'strew', *zalit'* 'pour'), where the frequency of the secondary construction is above 10%. For Polish this is the case only for *załadować* 'load'. The Polish ZA- therefore shows a stronger preference for one of the constructions, depending on the semantics of the verbal root, while in Russian there is a whole group of ZA-verbs that can alternate. The arrangement of the group of verbs that show constructional variation can be summarized as follows: the Russian and Polish verbs for 'load' show the highest variation between the two constructions; other potentially alternating verbs are 'strew', 'pour' and 'hang', where a relatively high variation is attested only for Russian. Finally, the Polish data indicate variation for 'pack' whereas the Russian data has variation for 'lay'.

The Goal-Object construction takes only 5% of the total uses of 'pack' in Polish, as in examples like (33):

(33) Zapakujemy plecak piwem i będziemy się integrować z
we-pack-FUT satchel-ACC beer-INS and we-will REFL integrate with
gośćmi.
guests-INS
'We will pack the bag with beer and then we will mingle with the guests.'

In the case of Russian 'pack', we have too little data to claim that this verb can't alternate at all, although our examples can be perceived as indirect evidence that the dominant construction here is the Theme-Object construction, resembling

the situation in Polish. Internet pages contain examples like (34), which indicate that the behavior of the verbs for 'pack' in Russian and Polish is quite similar:

(34) Est' dva čemodana: odin ja xoču
Be two-NOM suitcases-GEN: one-ACC I-NOM want-PRES.1SG
zapakovat' ličnymi veščami, drugoj podarkami
za-pack-INF personal-INS things-INS, another-ACC presents-INS
'I have two suitcases: one of then I want to pack with my personal things and the other one with presents'
(chemodan.eu/news/2008/01/4373.html?Page=5)

The Russian example (34) profiles the Goal (suitcases) as the direct object and the Theme (personal things) as the noun phrase in the Instrumental case and is thus parallel to the Polish example (33), where the bag (Goal in the Accusative case) is packed with beer (Theme in the Instrumental case).

It is remarkable that the corresponding ZA-verbs in Russian and Polish usually prefer the same construction. However, our data record two opposing cases: the verbs for 'hang', where the Russian *zavesit'* favors the Goal-Object construction and the Polish *zawiesić* chooses the Theme-Object construction, and the verbs for 'stuff', with the Russian *zapixat'* attested only in the Theme-Object construction and the Polish *zapchać* attested only in the Goal-Object construction.

A mismatch between Russian and Polish in the case of 'hang' can be accounted for by tracing the differences in the prototype of ZA- in the two languages. It appears that the Russian ZA- is strongly associated with a container. In particular, this idea is supported by the distribution of ZA- with spatial prefixes NA 'on' and V 'in'. For instance, the unprefixed verb *gruzit'* 'load' does not set any restrictions on the type of the Goal (i.e. information whether the Goal is a container or a surface) and shows an even distribution between the container-oriented preposition V and the surface-oriented preposition NA that introduce the Goal in the Accusative case. However, its Natural Perfective with ZA- strongly prefers the container-oriented V (88% vs. 12% with NA), while NA-perfective favors the surface-oriented preposition NA, see (Sokolova et al. forthcoming). It is plausible that when added to the Russian alternating verbs, ZA- as a rule shifts the focus from the Theme to the Goal. Yet, a ZA-verb can be used with the Theme-Object construction if the Goal is a container. For instance, in (35)-(36), the Russian verb *zasypat'* 'strew' has the same Theme (the gravel). When the Goal is a surface, as in (35), only the Goal-Object construction is used.

(35) Kogda vse rastenija budut vysaženy, zasyp'te ploščadku
When all-NOM pants-NOM be-FUT planted, za-strew-IMP ground-ACC
graviem
gravel-INS

'Once the plants are planted out, strew the ground with gravel.'

Cf.

(35′) ??Zasyp'te gravij na ploščadku
za-strew-IMP gravel-ACC on ground-ACC
'Strew the gravel on the ground.'

At the same time, the Theme-Object construction is possible with the same Theme (the gravel) if the Goal represents a container, which is usually headed by the preposition V:

(36) Zasyp'te gravij v akvarium
za-strew-IMP gravel-ACC into aquarium-ACC
'Strew the gravel into the aquarium.'
(minibiohome.com/manual_aquasaurs.php)

Even when the Goal of the verb *zasypat'* is marked by the preposition NA, it still refers to a container:

(37) Na dno tranšei zasyp'te gravij
On bottom-ACC ditch-GEN za-strew-IMP gravel-ACC
'Strew the bottom of the ditch with gravel.'
(www.mukhin.ru/home/decoland/30.html)

However, if the Goal of a verb is almost never a container, a shift in the construction occurs, as in the case of *zamazat'* 'smear', which is attested only in the Goal-Object construction. The verb *zamazat'* gains an additional meaning of "covering" and "hiding" the Goal, making it inaccessible (38):

(38) ... a kraskoj zamaž'te nadpisi na stenax
... and paint-INS za-smear-IMP inscriptions-ACC on walls-LOC
pod"ezdov
entrances-GEN
'...and use the paint to cover up the messages on the walls in the entrances.'

We might assume that the Russian ZA-, which is associated with "crossing a boundary", presupposes that the trajector is always INSIDE or BEHIND the landmark. As a result, Russian sets a restriction on the use of *zavesit'* 'hang' in the Theme-Object construction since the Goal of 'hang' is always a surface. It seems that Polish doesn't have such restrictions, thus allowing for the use of *zawiesić* 'hang' in the Theme-Object construction.

The second reverse case, attested in the verbs for 'stuff', occurs due to a significant shift in meaning of the Polish verb *zapchać.* Unprefixed *pchać* usually means 'push', referring to the situations of 'pushing on a door', 'pushing a baby carriage' or 'pushing a person', which use the "change of location" pattern. When prefixed with ZA-, it gains the meaning 'stuff' or 'choke', placing the focus on the Goal and changing the construction:

(39) Rozpostarłem gazetę i (...) zapchałem usta
I spread-PAST newspaper-ACC and (...) I stuffed mouth-ACC
obeschniętym plackiem.
dry-INS biscuit-INS
'I opened the newspaper and (...) I stuffed my mouth with a dry biscuit. '

On the other hand, the Russian verbs *pixat'* and *zapixat',* despite certain functional differences, both share the meaning 'stuff', as can be seen in (40)-(41), and can be used in the same, Theme-Object, construction:

(40) Ded uspeval prosledit', skol'ko šokoladnyx
Grandfather managed-IPFV trace-INF, how-many chocolate-GEN
konfet ... zapixali sebe v rot deti.
sweets-GEN ... za-stuff-PAST.3PL refl-DAT in mouth-ACC children-NOM
'Grandfather was able to keep track of how many chocolates the children stuffed into their mouths.'

(41) ... ne vse pixajut sebe v rot po dva
... not all-NOM stuff refl-DAT in mouth-ACC by two-ACC
buterbroda srazu
sandwiches-GEN at-once
'Not everybody stuffs his mouth with two sandwiches at once'
(http://forum.privet.com/viewtopic.php?f=4&t=27310&start=125)

In addition to a comparative analysis of the way the alternating ZA-verbs function in Russian and Polish, our data also make a crucial contribution to the general discussion on the Locative Alternation with regard to verbal roots. The properties demonstrated by the Russian ZA-verbs like *zavešat'* 'hang', as well as the selection of the Goal-Object construction by the Russian verbs *zalit'* 'pour', *zastavit'* 'put in a standing position' and the corresponding Polish verbs *zalać*, *zastawić*, present counterexamples to Pinker's claim that verbs like 'hang', 'pour' and 'put' should not appear in the change of state pattern. Thus, the occurrence of the Locative Alternation with such verbs cannot be described in terms of verbal semantics only but should be rather treated as a complex interaction of the verbal root, the construction and the prefix, which is closely related to the construction.

[5] CONCLUSIONS

The ZA- verbs in Russian and Polish show both similarities and differences in their syntactic behavior. On the one hand, the semantics of the prefix ZA- in both languages is construction-specific: when used with the Goal-object construction, the prefix refers to "covering" or "filling", whereas in the Theme-object construction, it bears the meaning of "reaching a natural endpoint" or "placing". In both Russian and Polish, the ZA-verbs prefer the Goal-Object construction, supporting the idea that prefixes correlate with constructions. In particular, in both Russian and Polish, ZA- eliminates alternation within the "manner" verbs, making the Goal-Object construction more prominent, and shifts many "hybrid" and "path" verbs from the Theme-Object construction to the Goal-Object construction. The use of each ZA-verb is generally skewed in favor of one of the locative constructions. The one exception is the Russian verb *zagruzit'* 'load', where an almost even distribution between the Theme-Object and the Goal-Object constructions appears to be the result of additional metaphorical uses in the Goal-Object construction.

On the other hand, some Russian and Polish alternating ZA-verbs behave differently. Firstly, more Russian ZA-verbs alternate between the two constructions: cf. Russian *zagruzit'*, *zasypat'*, *zalit'*, *zavešat'/zavesit', založit'* vs. Polish *załadować*, *zapakować*. Another difference is that although the corresponding ZA-verbs in Russian and Polish usually prefer the same construction, two opposing cases exist: the verbs for 'hang' (where Russian *zavesit'* favors the Goal-Object construction and the Polish *zawiesić* the Theme-Object construction), and the verbs for 'stuff' (the Russian *zapixat'* is attested only in the Theme-Object construction and the Polish *zapchać* in the Goal-Object construction). A mismatch between Russian and Polish in the case of 'hang' is motivated by the fact that the Russian ZA- is strongly associated with a container. As a result, Russian sets a restriction on the use of *zavesit'* 'hang' in the Theme-Object construction since the Goal of 'hang' is always a surface. The second case can be accounted for via a significant shift in meaning, which occurs in the Polish verb *pchać* 'push'. When prefixed with ZA-, it gains the meaning 'stuff' or 'choke', placing the focus on the Goal and changing the construction.

Our analysis also contributes to the study of the Locative Alternation. First, our findings indirectly support the division of the alternating verbs into "manner", "path" and "hybrid". In both Russian and Polish, "manner" verbs are clearly associated with the "change of state" meaning and are not attested in the Theme-Object construction. On the other hand, "path" and "hybrid" verbs differ from "manner" verbs in that some of them appear in both locative constructions. The proportion of the Theme-Object and Goal-Object constructions for "hybrid" verbs depends on the idiosyncratic properties of individual verbs, first of all on the degree to which the Theme is specified: the closer the Theme of the verb is to a

count noun, the more even the distribution between the constructions appears to be.

Moreover, our case study suggests that the prefix ZA- is the prefix that allows verbs typically associated with the change of location pattern to appear in the change of state construction (cf. the Russian verb *zavešat'* 'hang'). This goes against Pinker (1989), who claims that verbs like 'hang' or 'pour' should not appear in the change of state pattern. As our data from Russian and Polish show, such a shift is possible when the corresponding verbs are prefixed with ZA-.

Further elaboration of this topic requires a thorough comparison of the data described in this article with the corresponding unprefixed alternating verbs. Some other issues for future investigation are metaphorical extensions of the Theme-Object and the Goal-Object constructions in both languages and hybrid constructions.

ACKNOWLEDGMENTS

The major part of this research was made possible due to a grant from the Norwegian Research Council in support of the "Exploring Emptiness", a research group at the University of Tromsø. It has aslo been supported by the grants HUM2006-13295-C02-02 of the Spanish Ministery of Science and Innovation, and SGR-00753-2005 of the Generalitat de Catalunya.

REFERENCES

Boas, H. C. 2001. A lexical-constructional account of the locative alternation. In L. Carmichael, C.-H. Huang & V. Samiian (eds.), *Proceedings of the 2001 Western Conference in Linguistics 13*, 27–42.

Boas, H. C. 2006. A frame-semantic approach to identifying syntactically relevant elements of meaning. In P. Steiner, H. C. Boas & S. Schierholz (eds.), *Contrastive Studies and Valency. Studies in Honor of Hans Ulrich Boas*, 119–149. Frankfurt.

Brinkmann, U. 1997. *The locative alternation in German. Its structure and acquisition*. Amsterdam: John Benjamins.

Dickey, S. M. 2000. *Parameters of Slavic aspect: A cognitive approach*. Stanford: CSLI.

Dickey, S. M. 2005. S-/z- and the grammaticalization of aspect in Slavic. *Slovene Linguistic Studies* (5).

Divjak, D. 2006. Ways of intending: Delineating and structuring near-synonyms. In S. Gries & A. Stefanowitsch (eds.), *Corpora in Cognitive Linguistics: The Syntax-lexis Interface*, vol. 2, 19–56. Berlin: Mouton de Gruyter.

Divjak, D. & S. T. Gries. 2006. Ways of trying in Russian: clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2. 23–60.

Dowty, D. 2000. The fallacy of 'argument alternation'. In Y. Ravin & C. Laecock (eds.), *Polysemy*, 111–128. Oxford: Oxford University Press.

Fillmore, C. F. 2008. Border Conflicts: FrameNet Meets Construction Grammar. [Presented as a plenary talk at Euralex 2008].

Fillmore, C. J. 1968. The Case for Case. In E. Bach & R. T. Harms (eds.), *Universals in linguistic theory*, 1–88. New York.

Fillmore, C. J. 1977. The Case for Case Reopened. In P. Cole (ed.), *Syntax and Semantics. Grammatical Relations*, vol. 8, 59–81. New York Academic Press.

Goldberg, A. E. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.

Goldberg, A. E. 2002. Surface Generalizations: an alternative to alternations. *Cognitive Linguistics* 13(4). 327–356.

Goldberg, A. E. 2006. *Constructions at Work: The Nature of Generalizations in Language*. Oxford: Oxford U Press.

Golovin, B. N. 1964. Slovoobrazovatel'nyje tipy glagolov s pristavkoj ZA-. Voprosy teorii i vuzovskogo prepodavanija russkogo jazyka. *Učenyje zapiski Gor'kovskogo gosudarstvennogo universiteta* 68. 47–70.

Iwata, S. 2005. Locative alternation and the two levels of verb meaning. *Cognitive Linguistics* 16(2). 355–407.

Iwata, S. 2008. *Locative Alternation: A lexical-constructional approach*. Amsterdam: John Benjamins.

Janda, L. A. 1985. The meaning of Russian verbal prefixes: Semantics and grammar. In *The scope of Slavic aspect*, 26–40.

Janda, L. A. 1986. *A Semantic Analysis of the Russian Verbal Prefixes ZA-, PERE-, DO- and OT*. Munich: Otto Sagner.

Janda, L. A. & V. Solovyev. 2009. What Constructional Profiles Reveal About Synonymy and Metaphor: A Case Study of Russian words for 'sadness'. *Cognitive Linguistics* 20(2). 367–393.

Kurzowa, Z. 1997. *Tackling Polish verbs*. Kraków: Baran & Suszczyński.

Łaziński, M. 2008. The prefix ZA- in contemporary Polish against the Slavic background. In *Abstracts of the Third International Conference Perspectives on Slavistics*, 28–31. Hamburg.

Levin, B. 1993. *English verb classes and alternations: a preliminary investigation*. Chicago: University of Chicago Press.

Levin, B. & M. Rappaport Hovav. 1991. Wiping the Slate Clean. *Cognition* 41. 123–151.

Lewandowski, W. 2009. A corpus-based analysis of the locative alternation in Polish and Spanish. In *Presentation at the Fourth Annual Meeting of the Slavic Linguistic Society*, University of Zadar, Croatia.

Lewandowski, W. forthcoming. *La alternancia locativa en castellano y polaco: un análisis tipológico-construccional*. Universitat Autònoma de Barcelona.

Michaelis, L. & J. Ruppenhofer. 2001. Valence Creation and the German Applicative: the Inherent Semantics of Linking Patterns. *Journal of Semantics* 17(4). 335–395.

Nichols, J. 2008. Prefixation and the Locative Alternation in Russian contact verbs. In *Presentation at the annual conference of the American Association of Teachers of Slavic and East European Languages*, San Francisco.

Ovčinnikova, A. V. 1979. Tipologija značenij glagol'noj pristavki ZA-. In *Materialy po russko-slavjanskomu jazykoznaniju*, vol. 7, 110–116. Voronež: Izdatel'stvo Voronežskogo universiteta.

Ožegov, S. I. & N. Ju Švedova. 2001. *Slovar' russkogo jazyka*. Moscow: Russkij jazyk.

Pavlova, N. I. 1988. Semantičeskij potencial prefiksov v russkom slovoobrazvanii (na materiale izučenija slovoobrazovatel'nogo polja s prefiksom ZA-). In *Edinstvo sistemnogo i funkcional'nogo analiza jazykovyx edinic*, Belgrad.

Pinker, S. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, Mass.: MIT Press.

Rappaport Hovav, M. & B. Levin. 1988. What to do with theta-roles. In W. Wilkins (ed.), *Syntax and semantics: Thematic relations*, vol. 21, 7–36. San Diego: Academic Press.

Rappaport Hovav, M. & B. Levin. 1998. Building Verb Meanings. In M. Butt & W. Geuder (eds.), *The Projection of Arguments*, 97–134. Stanford: CSLI Publications.

Rappaport Hovav, M. & B. Levin. 2005. Change of State Verbs: Implications for Theories of Argument Projection. In N. Erteschik-Shir & T. Rapoport (eds.), *The Syntax of Aspect*, 274–286. Oxford: Oxford University Press.

Rappaport Hovav, M. & B. Levin. 2008. Reflections on Manner/Result Complementarity.

Śmiech, W. 1986. *Derywacja prefiskalna czasowników polskich*. Wrocław: Ossolineum.

Sokolova, S., O. Lyashevskaya & L. Janda. forthcoming. The Locative Alternation and the Russian 'empty' prefixes: A case study of the verb GRUZIT' 'load'. In *Frequency effects in cognitive linguistics: what statistical effects can(not) explain*, vol. 2 Trends in Linguistics Series, Berlin: Mouton de Gruyter.

Sokolova, S. O. 1982. O differencial'nykh semantičeskikh priznakakh glagol'noj pristavki ZA- v russkom jazyke. *Problemy leksičeskoj i kategorial'noj semantiki* 2. 32–41.

Stefanowitsch, A. & S. T. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8. 209–243.

Stefanowitsch, A. & S. T. Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1. 1–43.

Tabakowska, E. 2003. Space and time in Polish: the preposition ZA and the verbal prefix ZA-. In H. Cuyckens et al. (eds.), *Motivation in Language. Studies in honor of Günter Radden*, 153–177. Amsterdam: John Benjamins.

Talmy, L. 1985. Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (ed.), *Language typology and syntactic description: Grammatical categories and the lexicon*, vol. 3, 57–149. Cambridge, UK: Cambridge University Press.

Talmy, L. 2000. *Towards a cognitive semantics II: Typology and process in concept structuring*. Cambridge, MA: MIT Press.

Čertkova, M. J. 1996. *Grammatičeskaja kategorija vida v sovremennom russkom jazyke*. Moscow: Moscow State University.

Švedova, N. Ju et al. 1980. *Russkaja grammatika*, vol. I. Moscow: Nauka.

Weinsberg, A. 1973. *Przyimki przestrzenne w języku polskim, niemieckim i rumuńskim*. Wrocław: Ossolineum.

Wróbel, H. 1984. Słowotwórstwo czasowników. In R. Grzegorczykowa et al. (eds.), *Gramatyka współczesnego języka polskiego*, 467–512. PWN, Warszawa.

Zaliznjak, A. A. 2006. *Mnogoznačnost' v jazyke i sposoby ee predstavlenija*. Moscow: Jazyki slavjanskikh kul'tur.

AUTHOR CONTACT INFORMATION

Svetlana Sokolova
Department of Language and Linguistics
University of Tromsø
Norway
Svetlana.Sokolova@uit.no

Wojciech Lewandowski
Centre de Lingüística Teòrica
Universitat Autònoma de Barcelona
Spain
Wojciech.Lewandowski@uab.cat

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 393–412. (ISSN 1890-9639)
http://www.journals.uio.no/osla

THE FUNCTION OF MITIGATION IN SPOKEN LANGUAGE. THE ANALYSIS OF "TAK SKAZAT'" (RUSSIAN) AND "DICIAMO" (ITALIAN)

ELIZAVETA KHACHATURYAN
University of Oslo

ABSTRACT

The present article is dedicated to the analysis of one Italian and one Russian discourse marker (DM), both of which are formed with the verbum dicendi *skazat'/dire* ("to say"): *tak skazat'* ("so to say") and *diciamo* ("let's say").

[1] INTRODUCTION

The present article is dedicated to the analysis of one Italian and one Russian discourse marker (DM)[1], both of which are formed with the verbum dicendi *skazat'/dire* ("to say"): *tak skazat'* ("so to say") and *diciamo* ("let's say").

Both words appear frequently in spoken language. In Italian a formal equivalent of *tak skazat'* exists – *per così dire* (*per* is the preposition "for," *così* = *tak* or "so," *dire* = *skazat'*, or "to say"), but it is characteristic of the written language. At the same time, Russian has the word *skažem* (1st person plural of the verb *skazat'*), which occurs less frequently in spoken language compared to *diciamo*. Moreover, in spite of their formal equivalence, these expressions appear in different contexts in Russian and Italian: *diciamo* does not correspond to *skažem*, and tak *skazat'* does not correspond to *per così dire*. This is very typical of DMs: in spite of their formal similarities, their semantics and functions often vary considerably.[2]

The verb "to say" is in many different languages at the foundation of DMs that regulate communication. The description and analysis of these markers, which seem to be similar but in fact operate differently, can help us, on the one hand, to arrive at general formulas about universal functions that regulate communication and, on the other hand, to identify essential forms that carry out the same

[1] We will not discuss here the properties of these expressions which allow to us to classify them as discourse markers (for a detailed typology of discourse markers see Paillard, 2009).

[2] The same phenomenon can be illustrated by other examples: *vidiš'*(in Russian) – *vedi* (in Italian) – *tu vois* (in French); *poslušaj* (in Russian) – senti (in Italian) – *écoute* (in French) (Khaciaturian 2005): *veramente* – *vraiment* (Khachaturyan & Vladimirska 2010), *infine* – *enfin* (Rossari 1994). All these elements have the same form but often very different contexts of use.

function in different languages. These results will be useful for understanding second-language acquisition and translation. Moreover, the present analysis will illustrate several theoretical problems concerning the semantics of DMs, their use in spoken language, and the semantics of words traditionally called "semantic primitives" – for example, the verb "to say."

[1.1] *Definition of discourse markers*

At present, there are two main tendencies in studies on DMs which also represent two main methodological and theoretical approaches to their description: the functional-pragmatic approach and the formal-syntactic or formal-semantic approach. This analysis is based on the formal-semantic approach as elaborated by a French semantic school which takes as its point of departure the ideas of A. Culioli (*théorie de l'énonciation*).

According to this approach, DMs are considered to be a class of words (together with, for example, nouns, verbs, adjectives, etc.) with their own formal and semantic properties. "A DM is a word which defines the discursive status of the sequence p corresponding to its scope. In forming the scope of the DM, the sequence p stands for a particular way to say a state of affairs Z." (Franckel & Paillard 2008, 255) This means that every DM has its own semantics that can be deduced through the analysis of contexts of use and formulated in terms of which kind of discourse status is given to the scope p. In the present description, my purpose will be to formulate the semantics of *tak skazat'* and *diciamo*.

[1.2] *"Empty words" in spoken language*

Tak skazat' and *diciamo* occur so frequently in spoken language that they are often considered "empty words".

Usually, the analysis of "empty words" in spoken language raises the problems of grammaticalization, semantic weakening, and pragmatic enrichment (see i.e. Andersen (2000)). The functions distinguished for these words are often very similar: they are used to fill in pauses, indicate hesitation, an inexact or mitigate/attenuate nomination, or to allow the speaker the time to search for a word. However, in reality, the words in different languages with the same function are not equivalent and could not be used in the same context, not even as translations of one another.

The definitions given to *tak skazat'* and *diciamo* are very similar: both are used to attenuate statements.

> Ožegov's *Slovar' russkogo jazyka*: tak skazat' употребляется как оговорка, смягчающая решительность какого-нибудь утверждения (*tak skazat'* is used like a slip of the tongue, mitigating the resoluteness/absoluteness of a statement).

> Evgenieva's dictionary (MAC): *tak skazat'* употребляется для смягчения формулировки, для указания на неточность сказанного (*tak skazat'* is used to mitigate the formulation or to show the inexactitude of what is said).
>
> Sabatini & Coletti's dictionary: *diciamo* è usato nel parlato come riempitivo o come segnale di correzione di un dato (*diciamo* is used in spoken language as a filler or a mark of correction for dates)
>
> Bazzanella (1995): *diciamo* segnalatore di incertezza o di difficoltà di formulazione, [...], segnale "di attenuazione, o di "cortesia" (a signal of incertitude or of difficulties in formulating, a mark of attenuation or of politeness),
>
> Hölker (2003): "Diciamo als Mitigator": Ausdrucksbesonderheiten, Korrekturen (diciamo as a mitigator, mark of particular expression, mark of correction).

In the definitions above, almost the same terms are used to describe both words: inexactitude, mitigation, correction. However, we will see below that the contexts of use are often different. My next aim will be to describe the function of mitigation by analyzing the underlying mechanism upon which it is based in the case of *tak skazat'* and *diciamo*.

[2] TAK SKAZAT'

In all contexts where *tak skazat'* is used, the main problem that arises with this marker is "how to name" the given reality or which words to use. The chosen word or expression – the scope p of *tak skazat'* – has a specific discourse status as announced by tak skazat'. Three types of context can be distinguished based on the different status of the scope in the context. We will see afterwards that the role of *tak skazat'* has similar features in all contexts of use.

[2.1] *Case 1: Detached use*

A priori, p, the scope of *tak skazat'*, could be considered adequate to denote the given situation. But *tak skazat'* marks the speaker's[3] distance from what is being said and suggests that for her/him p is not an appropriate denomination of the reality R. In the examples below, this detachment of the speaker is discussed in the context that follows.[4]

[3] I will refer to the one who pronounces the phrase with the analyzed DM as "the speaker," and "the hearer" will refer to his/her interlocutor.

[4] In this analysis I used the Tübingen Russian Corpora and the Russian National Corpus.

(1) И помог он вам? – Не помог – сам все сделал. Когда я два дня спустя после работы к нему поехала, как мы условились, рецензия была готова – напечатана в двух экземплярах, все честь по чести. Я его благодарить, а он головой покачал и сказал: Не надо, Валя, это я в своих интересах, чтобы нам с вами сегодня не работать, а шампанское пить . . . – И глаза у него были в тот момент необыкновенные – грустные и какие-то сияющие, я таких ни у кого еще не видела. Наверное, в тот момент я в него и влюбилась. Что ж, я – человек решительный. Прямо при нем сняла трубку и позвонила домой, что буду ночевать у подруги. – Любовь с первого взгляда, так сказать . . . – Это что, ирония? (А. Стругацкий)

. . . and did he help you? – He didn't only help me, but he did everything himself. When I went to him, as arranged, two days later after work, the review was ready – printed in duplicate. Everything is just as it should be. I thanked him but he shook his head and said: "There is no need to thank me, Valya. I did it in my own interest so that we don't have to work today but can just drink champagne . . ." and, in that moment, his eyes were extraordinary – sad and yet sort of shining, such as I have never seen in a person before. I probably fell in love with him at that very moment. Well, I am a decisive person. Right in front of him, I picked up the phone and called home to say that I would be staying with a friend that night. – Love at first sight, tak skazat' . . . – What's that, irony? (A. Strugatskij)

In (1), the situation described in the previous context can be tagged as "ljubov' s pervogo vzgljada" (love at first sight) (p). But here the speaker does not share this tag: p is introduced by *tak skazat'*. It causes the reaction of the hearer (*What's that: irony?*) who perceives the disengagement of the speaker and interprets it as an ironic comment: *ironija* (irony).

(2) Брат Потапова в прошлом году арестован и осужден за вредительство, он находился в связи с консулом одной из враждебных держав и получал задания от иностранной разведки. Во всем этом он сознался. Вот вам вторая и, так сказать, неожиданная сторона бригадира Потапова. Вы всего этого, конечно, не знали, – улыбнулся он. – Про брата знал, – сказал я неожиданно для самого себя. (Ю. Домбровский)

"Potapov's brother was arrested last year and convicted of sabotage. He was connected with the consul of one of the hostile states and received his instructions from the foreign secret service. He confessed everything. So here is a second – and, tak skazat', unexpected – side to your foreman Potapov. You, of course, knew nothing about all this", he smiled. "I knew about his brother", I found myself saying unexpectedly. (Ju. Dombrovskij)

In (2), the speaker defines the facts from Potapov's life as part of a "neožidannaja storona" (unexpected side) and adds: "of course, you knew nothing about it." In other words, the interlocutor presumes that this side of Potapov is unexpected to the first person narrator (the hearer), whereas he himself, as an experienced secret service official, did suspect something (as is indeed verified in this novel by Ju. Dombrovskij). The secret service officer uses *tak skazat'* to distance himself from what he is saying.

[2.2] *Case 2: Conventional use*

In this case, the term p a priori is not appropriate to denominate the situation. By contrast, *tak skazat'* indicates that p can be considered as a kind of figurative denomination of R: there is something similar between the described situation and the situation usually denominated by this word.

(3) Пусть вас не обманывает его должность. Он фигура, величина, три, а то и все четыре ромба, больше, чем его начальники в Канске, потому и форму не надевает. Был, между прочим, за границей, а попал сюда. Боюсь, он наш будущий, ТАК СКАЗАТЬ, коллега или сотоварищ. А может, и обратно выскочит, все зависит от каких-то высших, нам с вами неизвестных обстоятельств. (А. Рыбаков)

You shouldn't be fooled by his position. He is a prominent figure, a big name, three or four rhombs more than his superior in Kansk – that is why he doesn't wear a uniform. He was abroad, by the way, but then turned up here. I fear he will be our future, TAK SKAZAT', colleague or associate. But maybe he will jump out back – it all depends on some higher circumstances, unknown to us. (A. Rybakov)

The word *kollega* (colleague) indicates someone who works together with others. In (3), since it is a political prisoner who is speaking, the denomination "colleague" is not entirely adequate: the second word "sotovarišč" (associate, fellow, inmate) fits better.

(4) (о дедовщине в армии) . . . На вторую ночь нас начали по одному поднимать и, ТАК СКАЗАТЬ, знакомиться . . . На следующий день у меня заметили синяк . . . (газета «Коммерсант»)

(regarding violence against young conscripts in the army) . . . the second night they started to wake us up one by one and, TAK SKAZAT', make our acquaintance. The following day they noticed my bruise. . . (newspaper "Kommersant")

In (4) the combination "*tak skazat'*, znakomit'sja" (*tak skazat'*, make our acquaintance) indicates that there is something specific in this act of getting acquainted. Indeed, talking about violence in the army, the expression "to make smb's acquaintance" (i.e., to beat someone up!) is not completely adequate. Still, the expression can be interpreted as a particular way to describe the situation.

[2.3] *Case 3: Quotation*

In this case the detachment of the speaker can be explained by the status of the quotation itself: *tak skazat'* marks the words of another in the text of the speaker.

(5) – Не надо без мужа ходить по ресторанам, – начал Костя, – можно нарваться на неприятности. Будь ты со мной, к тебе бы никто не пристал, пошла без меня, вот и нарвалась. – До тебя, – ответила Варя, – когда у меня не было, так сказать, мужа, ко мне никто не приставал, никто меня не оскорблял. Эта особа оскорбила меня именно потому, что я была твоей женой, и посчитала меня тоже шлюхой. – Она психопатка, – возразил Костя, – она больная . . . (А. Рыбаков)

"You shouldn't go to restaurants *without your husband*", began Kostya "you may run into trouble. If you had been with me, nobody would have bothered you but you went without me and look what happened". "Before you were around", Varya answered "when I didn't have, так skazat', a husband, nobody bothered me, nobody insulted me. This individual insulted me precisely because I am your wife, even made me out to be a whore. "She's mad", retorted Kostya "she's sick". (A. Rybakov)

In (5) the speaker repeats the term *muž* (husband [italics are mine – E.K.]), which the other speaker had used to refer to himself, and marks her distance from the given tag through *tak skazat'* (her discovery that he was officially married to another woman makes the distance doubly apparent). So, for the female speaker the word introduced by *tak skazat'* is inadequate to describe the reality: the strangeness (to her) of Kostya's word choice is marked through the use of *tak skazat'*. Example (6) illustrates the point more fully.

(6) – Вот уж никогда бы не подумал, – пробормотал я, – что у Хинкуса есть друзья, которые согласны разделить с ним его одиночество. Хотя . . . почему бы и нет? Пуркуа па, так сказать . . . (бр. Стругацкие)

"I would never have thought", I muttered "that Hinkus has any friends, who are prepared to share his solitude. Although . . . why not?" Pourquoi pas, так skazat' . . . (Strugatsky brothers)

In (6), there is a formal reason for not attributing the highlighted words and their interpretation to the speaker: they are foreign words for which the speaker does not assume responsibility.

[2.4] Tak skazat'*: Summing up*

We can notice that all three types of context have some general features that can be considered a core (invariant) meaning of *tak skazat'*. It could be formulated in following way: *tak skazat'* p means that the words p used to say R are not completely adequate to speak about the world R. The speaker does not share the responsibility for what he is saying and marks by *tak skazat'* the "space of the words" – a kind of zone free from the speaker's engagement in what s/he is saying.[5]

It is interesting to notice that in the majority of contexts with *tak skazat'* it is possible to use inverted commas (or they are used by the author) – a typographic mark that indicates the "plurivocity of words" and "the infinity of interpretations" (Authuer-Revuz 1995, 141).

[2.5] Tak skazat' *in spoken language*

Tak skazat' has the same role in spoken language: it marks the incompatibility of the contextualized words with the context itself and indicates the disengagement of the speaker. In spoken language the incompatibility of p is often based on switching to another register. It explains the properties (typical for *tak skazat'*) of "poorly organized" discourse and the role of *tak skazat'* as an interrupting or correcting marker. Actually, it is the scope of *tak skazat'* that does not fit the context (not only for semantic, but also for syntactic or stylistic reasons) and as such interrupts/breaks the development of discourse.

Since the context in spoken language has less stable characteristics, in comparison to written texts, more formal criteria should be used to distinguish the uses of *tak skazat'*: these criteria should be first of all the description of prosodic features. In the present contrastive analysis, whose main purpose is to compare *tak skazat'* and *diciamo*, it will suffice to illustrate the use of *tak skazat'* in the spoken language by means of the following examples.

(7) [Бунич, муж] Да / но / тем не менее вот этот проект создания торгово-развлекательного центра на Красной площади под . . . ээ . . . так сказать вместо музеев он существует и кроме того там . . . ээ . . . какие-то люди роют . . . ээ . . . ямы / какие-то подкапываются под Красную площадь. [Программа «Диалог с Андреем Буничем» на телеканале РБК (2006)

(Bunič, male) Yes . . . but . . . nevertheless, take the project to create a re-

[5] The paradox is that disengagement can be also interpreted as engagement in a negative sense ((1) is a good illustration of this).

tail centre in Red Square under . . . hm . . . так skazat' instead of museums it exists and moreover . . . hm . . . some people dig holes there / some dig under Red Square. (Programme "Dialogue with Andrey Bunič" on the TV channel RBK)

In (7), the beginning of the phrase is interrupted by the word combination "vmesto muzeev" (instead of museums), which is adequate to describe the situation but can be considered too colloquial, or syntactically inappropriate: the name of the project could be "the creation of a retail and entertainment centre" but not "vmesto muzeev," so it is very improbable that in the formal definition the word "muzej" (museum) is present. The expression used by the speaker "vmesto muzeev" could be considered an unofficial definition of what is happening.

In (8), the speaker tries to say something difficult using simple words incommensurate with the style of the lecture. This simplification inappropriate to the situation explains the majority of the "lecturer's *tak skazat'*."

(8) [Н.К., муж] Повторяю / есть / значит / такое заболевание / как / значит / э / значит [нрзб] синдром / когда в случае заболевания мужской организм начинает перерождаться в женский / а женский в мужской / без каких бы то ни было / вот / так сказать / внешних толчков / эффектов внешней среды. [Лекция о мозге (2006)]

(NK, male) I repeat / there is / sort of / such a disease / as/ sort of / hm / sort of (incomprehensive) syndrome / when, in the event of illness, the male organism starts to re-develop into a female / the female into a male / without any / так skazat' / external stimulus / effects of the environment (Lecture about the brain, 2006)

The opposite situation is also common: the speaker uses a term characteristic of "high style." This is often typical of quotations, as in (9),(10), where the scope of *tak skazat'* includes the expression "prizvany okhranjat'" (9) and the word "vnimaju" (10), which are more refined in comparison to the colloquial style of the rest of the text.

(9) [Пясецкий, муж, 1925] Вот / допустим / приказ был такой / поскольку немцы оказывают сопротивление / вот / но танковые колонны / невзирая на это / должны идти вперед / а мы были / так сказать / призваны охранять эти танковые колонны. [радиопередача, 2005]

(Pjasetskij, male, 1925). Here / let us suppose / there was such an order / as soon as Germans put up resistance / here / but the tank convoys /in

spite of this / must move forward/ and we were / TAK SKAZAT' / called up to protect these tank convoys. (Radio programme, 2005)

(10) [No. 1, муж] Давайте все-таки вернемся к теме классической музыки. Хотя / конечно / мы не так далеки / тема / действительно / актуальная / но давайте поближе / господа! Евгений Львович / Ваше мнение . . . [Nr. 6, ?] Я вот сижу / слушаю / внимаю / ТАК СКАЗАТЬ / и . . . дивлюсь. Дивлюсь потому / что мы забыли о самом важном факторе / влияющем на формирование личности / и имя ему / семья. Семья / друзья мои! Семья! [Программа «Большие» на телеканале ТВЦ (2006)]

(Nr. 1, male) Let us however revert to the subject of classical music. Though /of course / we are not too far / the subject /really / is topical / but let us come closer / gentlemen! Evgenij L'vovič / your opinion . . . (Nr. 6, ?) I am sitting here / listening / paying close attention / TAK SKAZAT' / and . . . marvelling. Marvelling because / we forgot the most important factor that influences the formation of personality / namely / the family. Family / my friends! Family! ("Grown-ups" programme on TVC (2006)

[3] DICIAMO

The contexts in which *diciamo* appears have to do with the same problem of denomination: how to name and speak about a situation R. Diciamo p (like *tak skazat'*) means that the words p used to say R are not completely adequate to speak about the world R. The scope p of *diciamo* can also be marked by inverted commas. Three contexts of use can be distinguished.[6]

[3.1] *Case 1: Figurative denomination*

The word introduced by *diciamo* is not the word normally used to speak about R: this means that p is used for another situation R'. However, there is something similar between both situations that allows the same word to be used.

(11) Proposte di assoluzione non ce ne saranno molte, ma – ha anticipato il rappresentante dell' accusa – non mancheranno richieste di condanna a pene miti per quegli imputati, "DICIAMO di serie C", che *hanno finito per aderire* all' organizzazione per delinquere di don Raffaele Cutolo *per paura, alcuni addirittura lasciandosi affiliare in carcere per non rimanere tagliati fuori, per tutelare la propria incolumità, perché costretti a scegliere:* o con quelli della Nuova camorra organizzata o contro di loro.

There won't be many proposals for acquittal – anticipated the representa-

[6] In this analysis I used the Repubblica Corpus, LIP, C-oral-Rom.

> tive for the prosecution – but there won't be any lack of proposals for light sentences for those defendants – DICIAMO of the bottom league – who, for fear, have ended up being part of don Raffaele Cutolo's delinquency organisation, some even letting themselves be signed up in prison so as not to be left out and to safeguard their own safety because they have to choose; or with the new organised Camorra or against them.

The expression "di serie C" (of the bottom league) is used to denote the less important and less well-known soccer teams in the Italian championship. It is the less important and the less famous league. In (11), the expression "di serie C" is used to refer to the defendants in the Camorra process. The context that follows [my italics – E.K.] describes the defendants (considered neither famous nor important) of this group and in this way motivates the figurative use of the expression "di serie C."

The same explanation can be given to the following example (12):

(12) Se Alain Juppé dovesse dimettersi, voi preferireste un uomo come Philippe Séguin, neogollista, DICIAMO, di sinistra? Il presidente dell'Assemblea, nemico del trattato di Maastricht, è andato a dialogare coi ferrovieri e a distribuire pacche sulle loro spalle. (Il Corriere della Sera)

> If Alain Juppé were to resign, would you prefer someone like Philippe Séguin, a neo-Gaullist, DICIAMO, and left-wing? The president of the (National) Assembly and enemy of the Treaty of Maastricht, went to talk with the railway workers and give them pats on their backs.

The tag "di sinistra" (left-wing) would normally be considered the opposite of the preceding one – "neogollista" (neo-Gaullist) –referring to a right-wing party. However, it has to be interpreted in the context of this first definition, and in this case it becomes a particular denomination of a less radical group.

[3.2] *Case 2: Ambiguous denomination*

The scope p a priori is adequate to say R. If *diciamo* is removed nothing disturbs the acceptance of p. But p can have different interpretations according to the context, and both (or more than one interpretation) must be considered in the case of *diciamo* p. *Diciamo* marks the problematic status of p: it is an ambiguous denomination of R.

(13) Per parte nostra vorremmo esprimere un' impressione (che, ovviamente, potrebbe anche essere sbagliata). L' impressione è questa: la Chiesa italiana va a Loreto per la "riconciliazione", ma sembra andarci già con un peccato (DICIAMO tentazione) di orgoglio: quello di voler fare bella figura

di fronte alla società italiana, di cui essa stessa è parte.

On our part, we would like to convey an impression (which obviously could be wrong). Our impression is the following: the Italian Church goes to Loreto for "reconciliation", but it seems to go there already with the sin (DICIAMO temptation) of proudness; as if they want to look good in the eyes of the Italian society to which the Church itself belongs.

According to Sabatini & Coletti, the word "tentazione" (temptation) can be interpreted as either a desire reproachable from a moral point of view or a simple desire (cf. *La tentazione di rubare* "the temptation to steal," and *ho la tentazione di raggiungerti con il primo treno,* "I am tempted to join you with the first train"). In (13) both meanings of the word should be taken into consideration.

(14) Poiché cinema e Tv non sempre si possono distinguere come dimostra il fuori concorso che giovedì prossimo chiuderà il festival dopo le premiazioni : Siete meravigliosi di Giuseppe Bertolucci con Roberto Benigni, ripresa del suo show dello scorso anno che è dunque insieme cinema, teatro e tv (lo ha prodotto la Rai, ma non lo ha ancora mandato in onda per problemi, DICIAMO, lessicali).

Cinema and television are not always distinguishable as seen from the "out-of-competition" that will close the festival after the prize-giving next Thursday: Siete meravigliosi by Giuseppe Bertolucci starring Roberto Benigni, taken from his previous year's show and so is at the same time cinema, theatre and television (produced by RAI but still not aired for *diciamo* lexical problems).

In (14), "lessicale" (lexical) can be something that belongs to the lexis in general, or something connected with the use of the words (cf. *Sistema lessicale* "lexical system" – errore lessicale "lexical mistake"). This ambiguous interpretation is kept by *diciamo*.

In (15), in the context that follows the nomination "capolavoro" (masterpiece), two properties of the book defined as "capolavoro" are introduced, but one of them does not fit the book in question, since it was "never read." That's why the definition "capolavoro" is problematic.

(15) Infine, per rimanere dalle parti dell'India, ci sarebbe un libro scritto una sessantina di anni fa, DICIAMO un capolavoro: Passaggio in India, di E.M. Forster, molto citato e letto mai. Consigliamo di comprarlo prima che arrivi il film: così uno può seguire meglio la storia.

> Finally, always on the India theme, there is a book written about sixty years ago; DICIAMO a masterpiece, A Passage to India, by E. M. Forster, quoted often and never read. We suggest that you buy it before the film gets here so you will be able to follow the story better.

[3.3] *Case 3: Another possibility*

The term that corresponds to the scope p of *diciamo* is the interpretation or explanation of what is said in the previous context, but it could be considered as such only in this concrete situation. Often, p takes the form of an exact, singular and/or objective denomination (i.e. numbers or names), while the first denomination has a subjective character.

(16) Dante, mi sembra, è oggi voce morta per i giovani e per i meno giovani, DICIAMO per i quarantenni; forse parla ancora ai cinquantenni e oltre, che si sono avvicinati a lui in epoche eroiche, quando cercavamo nel sapere parole definitive [. . .]

> It seems to me that Dante in our day is a dead voice for the young as well as the less young DICIAMO for the forty-year-olds; maybe he is only talking to people fifty years old and more [. . .]

In (16), "i meno giovani" (the less young) – can be interpreted in different ways depending on the situation, but also on the speaker, "i quarantenni" (the forty-year-olds) – is an explanation appropriate in this concrete situation from the point of view of the speaker.

The same interpretation can be given to (17): p explains what "pochissime" (very few; the absolute superlative form of the adjective "poco") means in this context, but that does not mean that "mille su ventimila" (one thousand to twenty thousand) is the exact number: it is simply used to give a concrete illustration of the word "pochissime".

(17) E fa anche una statistica, cifre alla mano, delle "vere" donne francesi dei suoi tempi. Erano pochissime: DICIAMO, mille su ventimila prese in considerazione. Dice che devono essere oggetti gradevoli , ben vestiti , spirituali . . .

> And he presents also the statistical data, with numbers at hand, of the "true" French women of his day. They are very few: DICIAMO one thousand to twenty thousand that were considered. He says that they have to be pretty subjects, well dressed, spiritual . . .

[3.4] Diciamo*: Summing up*

So, *diciamo* p (like *tak skazat'*) indicates that the words p used to say R are not completely adequate to speak about the world R. But, in contrast to *tak skazat'*, *diciamo* introduces a somehow individual denomination. The speaker tries to involve the hearer (through the form of the 1st person plural) and invites him to accept conventionally and/or provisionally what is said.

This difference between the two words is very well illustrated by the last use (case 3 for both words): *diciamo* introduces an individual interpretation of what is said before; *tak skazat'* accompanies a quotation or the words of another for which the speaker does not assume responsibility.

The other two uses of *diciamo* and *tak skazat'* are very similar. But at the same time the form of the scope is different: *diciamo* introduces a metaphorical (in case 1) or an ambiguous (in case 2) denomination, while the scope of *tak skazat'* is often a generalized denomination or an idiomatic expression.

[3.5] Diciamo *in spoken language*

In spoken language, the similar contexts of use of *diciamo* can be distinguished.

(18) Senta colonello // un'ultima domanda // sia l'onorevole Bertinotti / sia / l'onorevole Buttiglione / da posizioni politiche molto diverse / vi hanno espresso / la loro simpatia / e i loro auguri // voi / vi sentite / appoggiati / dalla classe politica italiana? – beh / come sono andate le cose / DICIAMO / in questi giorni / sì / prima ci sentivamo un po' abbandonati / per tutte le problematiche / che ci sono state // parti / non parti // forse sì // forse no // vediamo // ecco / questi sono stati i nostri problemi

The last question colonel / both the Member of Parliament Bertinotti and the Member of Parliament Buttiglione / from different political positions / expressed their sympathy and best wishes // do you feel supported by Italian politics? / – well / what was happening / diciamo / in those days / yes / in the beginning we felt ourselves a little bit abandoned / because of all the problems / that we had / you have to leave / you don't need to leave // probably yes / probably no / we'll see / so / these were our problems

(19) l'investimento risale a / millenovecen // duemila . . . duemila // – duemila // quindi praticamente / lei s'è beccato tutto il periodo / DICIAMO / di massimo splendore dei mercati finanziari

the investment dates by / nineteen . . . // two thousand . . . two thousand // two thousand // it means that / you picked up the whole period / DICIAMO / of the highest grandeur of the financial market

In (18), the scope of diciamo is the speaker's personal reinterpretation of the question: "vi sentite appoggiati dalla classe politica" (do you feel supported) as "come sono andate le cose" (what was happening). In (19), p is an individual, a subjective characterization of the period.

[4] *tak skazat'* – *diciamo*: ANALOGIES AND DIFFERENCES

As we have seen *diciamo* and *tak skazat'* have several common features which can be defined as typical of mitigation/attenuation. Both:

- participate in the denomination: marking p as adequate or not to say R
- take into account others' possibilities of saying R
- take into account possible interpretations of p, i.e., other R^{n} that could be said by p.

At the same time, the attenuation/mitigation that stays in the definition of both DMs can have different interpretations. For both words it is based upon not assuming responsibility. But in case of *tak skazat'* this "non-assumption of responsibility" is an active removal of responsibility: the speaker disengages from what s/he is saying and leaves the words to perform on their own. In case of *diciamo* "non-assumption of responsibility" means "sharing it with the audience."

It is interesting to compare the words of the same form: *diciamo* and *skažem*, on the one hand, and *tak skazat'* and *per così dire,* on the other hand.

The DM *skažem*, in contrast to *diciamo*, is not involved in the process of denomination. It marks that the scope p could be adequate to say the situation R ("state of affairs") and proposes p as a point of departure for the development of the context that follows.

(20) Почему крестьянство поддержало революцию в центральных губерниях и не поддержало на окраинах, СКАЖЕМ, в Сибири? В центральных губерниях мужик видел помещика, дворянина, а в Сибири их не было. А когда появился дворянин Колчак, тогда сибирский мужик поддержал революцию. (А. Рыбаков)

Why did the peasantry support the revolution in the central provinces but not in the outlying districts, SKAŽEM, in Siberia? In the central provinces the mužik (peasant) could see the landowner, the nobleman, but in Siberia there were none. And when the nobleman Kolčak appeared there, then the Siberian mužik supported the revolution. (A. Rybakov)

The form *per così dire* is used in the written language and is very unusual in spoken language. In this case, the attention is focused on the ambiguous status of the

words that can be both adequate and inadequate to say R. In contrast, *tak skazat'* (as we have seen) shows the disengagement of the speaker based on the possibility of different interpretations of the words.

(21) Quest'opera di 'salvataggio' è da un lato resa più agevole dall'impressionante ricchezza della stampa italiana all'estero (. . .), e da una crescente attenzione nei confronti di testimonianze di tipo diaristico-autobiografico; dall'altro è penalizzata – PER COSÌ DIRE – dall'assenza di opere singole tali da venire assunte a simbolo rappresentativo di quella grande storia collettiva.

This last-minute 'rescue operation' is, on one hand, made easier by the impressionable wealth of the foreign Italian press (. . .) and by growing attention aimed at diarist-autobiographical like stories. On the other hand, it is penalised – PER COSÌ DIRE – by the absence of single works able to be considered emblematic of such a great collective history.

[5] THE VERBS *skazat'* – *dire*: CHARACTERISTIC FEATURES

I will argue in the last part of this paper that the differences between DMs could be explained by the semantic differences between the DMs' lexemes of origin.

The hypothesis concerning this link between the semantics of a discourse element and the form from which it derives was formulated in the theory of Culioli (1990, 2002); Culioli & Normand (2005) and was illustrated in different studies (i.e., Franckel & Paillard (2008); Paillard (1998, 2001, 2002, 2009); Khachatourian (2006); Khachaturyan (2008); Vladimirska (2008)).

In this paper I will analyze only a few contexts that can illustrate the semantic differences between the seemingly similar verbs *skazat'* and *dire* ("to say").

The Italian verb in several contexts synonymous with the verb *pensare* ("to think"). But it is impossible to use the Russian verb *skazat'* in the following examples:

(22) a. E dire che: E dire che non ha ancora 20 anni!
Dire INF that: And dire INF that he is even not 20 years old! (*litt.*)
b. Chi l'avrebbe detto!
Who dire PAST CONDITIONAL this! (*litt.*)
c. Si direbbe: Fa bel tempo. Si direbbe l'estate.
Si IMPERS PRONOUN dire PRESENT CONDITIONAL: The whether is nice. Si direbbe the summer. (*litt.*)

The idea of the exteriorization of thoughts via words is important for the Italian verb, while the Russian verb is focused on the interpretation that the hearer will give to the words in this concrete situation (see this argument in more detail in

Khachatourian (2006)). Both verbs in similar contexts will have different interpretations. E.g., the negation with the verb *dire* – "senza dire niente" (without saying anything) – can be glossed as "without 'vocalising' / without giving a voice to thoughts, actions, events". The situation is often on the order of: "smb knew smth but did not say it." This interpretation can be illustrated by the following three examples [my italics – E.K.].

(23) E ho coinvolto un altro dei miei miti, Hugo Pratt. Fa un commissario straniero che arriva, guarda, *capisce tutto* e se ne va SENZA DIRE NIENTE. Ha capito che i delitti non sono roba sua, è roba da psicanalista bravo. (La Repubblica)

So I introduced one of my others myths, Hugo Pratt. He is a foreign commissary who arrives, looks, understands everything and goes away SENZA DIRE NIENTE (without saying anything). He has understood that these crimes are not for him, but for an experienced psychoanalyst.

(24) Era ... strano. Dopo l'orrore, mi riempiva di gentilezze. Entrando alla dacia avevo guardato un albero di limoni. *Solo uno sguardo*, SENZA DIRE NIENTE: e subito mi fa trovare una cesta di limoni in tavola. (id.)

It was ... strange. After the horror that we survived, he was very kind to me. When I entered the dacha, I looked at the lemon tree. Only looked at it, SENZA DIRE NIENTE (without saying anything): and immediately he sent a basket of lemons to my table.

(25) *Una scuderia decide* di montare un motore di 3500, 4000, 5000 cc anziché di 3000 cc come da regolamento e così facendo vince alla grande SENZA DIRE NIENTE *al pilota*. (id.)

The team decides to put a motor of 3500, 4000, 5000 cc instead of 3000 cc established by the rules, and in this way the team wins SENZA DIRE NIENTE (without saying anything) to the pilot.

In (23), the eventual ("failed") speaker understands everything and goes away without saying what he has understood. In (24), the communication is based on glances that transmit the desire without words. In (25), the pilot is not warned about the changes made to the motor.

The Russian verb in the negative construction (*ni slova ne skazav*[7] "without saying a word") is used in a situation where there is a hearer who is waiting for

[7] The form of the construction is also interesting: usually there is the word "slovo" ("a word"), and less often "ničego" ("nothing") is used.

the words of the eventual speaker (as in (26)). The hearer is often introduced by a negative pronoun: "nikomu" ("to nobody").

(26) Случайно оказавшись зимой на берегу Десны, Алеша заметил, что в воду упал маленький мальчик (. . .). Алеша не только вытащил малыша на берег, сделал ему искусственное дыхание, но и остановил милицейскую машину, проезжавшую по шоссе, довез Валеру до больницы, а сам никому больше ни слова не сказав, отправился домой. (Огонек, 1997, 05)

Finding himself by chance on the bank of the Desna river in winter, Alyosha noticed a little boy fall into the water (. . .) Alyosha not only dragged the small child onto the bank and gave him mouth-to-mouth resuscitation but also stopped a passing police car, took Valera to hospital and left for home NI SLOVA NE SKAZAV (without saying a word) to anyone. (Ogonyok, 1997, 05)

We can compare with the construction containing the aspectual pair of *skazat'* (perfective form) the verb *govorit'* (imperfective form). The negative construction with *govorit'* will be interpreted as "keeping silence," as in (27), where any hearer is present.

(27) . . . молодой человек, не глядя ни вправо, ни влево, поднялся в свой двадцатый номер, бросил слуге фуражку и шпагу, а на расспросы лишь качнул головой. Привычный Маса понимающе поклонился и проворно расстелил на полу соломенную циновкую. Куцую шпажонку почтительно обернул шелком и положил на щифонер, сам же, ни слова не говоря, вышел в коридор и встал спиной к двери в позе грозного бога Фудоме, повелителя пламени. (В. Акунин)

. . . the young man, without looking right or left, went upstairs to his room, number 20, threw his service cap and sword to the servant and just shook his head in answer to questions. Masa, being accustomed to it, bowed to him knowingly and quickly spread out the straw mat on the floor. He respectfully wrapped the short sword in the silk and put it on top of the wardrobe, then, NI SLOVA NE GOVORJA (without saying a word), went out into the corridor and stood there with his back against the door in a pose of the menacing god Fudome, tribe sovereign. (B. Akunin)

The participation of the active hearer in the context with *skazat'* can be also illustrated by another example. The constructions like *skaži/skazal pravdu* (tell me the truth/(he) told the truth) are usually followed by the comments of the speaker

(who will become the hearer) who knows the truth, so he can judge whether what will be said (or was said) is the truth or not.

(28) – Значит, никто не пробегал? – спросила Алиса. Она уже поняла, что никто не пробегал. Иначе бы этот тихий уголок переполошился. – Нет, – сказала молоденькая мама в широких брюках. – А они должны побежать? СКАЖИ ПРАВДУ, тогда я отвезу малыша домой. Детям нельзя волноваться. (К. Булычев)

– So, nobody ran by? – asked Alice. She already realised that nobody had run by, otherwise this place would not be so quiet. – No, – said the very young mother in wide trousers. – And should they have run by? SKAŽI PRAVDU (Tell me the truth), and then I'll drive the child home. Children should never be made to worry. (K. Bulyčev)

The construction with the word *novost'* ("news") has a similar property: it is often followed by the interpretation given from the point of view of the hearer, as in (29), where it is clear that the news is unpleasant for the hearer.

(29) Не сердись, Миколя. Но СКАЖУ ТЕБЕ НЕПРИЯТНУЮ НОВОСТЬ . . . – Он сразу изменился в лице.

Don't be angry, Mikolya. But SKAŽU TEBE NEPRIJATNUJU NOVOST' (I'll tell you the unpleasant news) . . . He immediately changed his countenance

[6] CONCLUSION

In this article I have described two DMs, *tak skazat'* and *diciamo*, which occur quite frequently in spoken language and are often considered to be "empty words." Traditionally, in the written language, both markers are designated as mitigating/attenuating elements in the text. We have seen, however, that the function of mitigation, although in both cases it actually corresponds to the "non-assumption of responsibility," is based on different principles: *tak skazat'* marks the disengagement of the speaker from what is said, while *diciamo* indicates the union with the hearer that allows for shared responsibility. These differences could be explained not only by the form of the two DMs, but also by the different semantics of the verbs *skazat'* and *dire*. Both markers retain their semantics in spoken language. The idea of the non-assumption of responsibility brings us to the general problem which concerns the meaning of words and plurivocity (various interpretations) of the same word.

REFERENCES

Andersen, G. 2000. The role of the pragmatic marker *like* in utterance interpretation. In *Pragmatic markers and propositional attitude*, Amsterdam: Jonh Benjamins.

Authuer-Revuz, J. 1995. *Ces mots qui ne vont pas de soi: boucles réflexives et non-coïncidences du dire*. Paris: Larousse.

Bazzanella, C. 1995. I segnali discorsivi. In L. Renzi, G. Salvi & A. Cardinaletti (eds.), *Grande grammatica di consultazione*, vol. III, il Mulino.

Culioli, A. 1990. *Pour une linguistique de l'énonciation.*, vol. I. Ophrys.

Culioli, A. 2002. *Variations sur la linguistique. Entretiens avec Frédéric Fau.* Paris: Klincksieck.

Culioli, A. & C. Normand. 2005. *Onze rencontres sur le langage et les langues.* Ophrys.

Franckel, J.J. & D. Paillard. 2008. Mots du discours : point de vue et adéquation. Etude de réellement, en réalité, en effet, effectivement. In *Revista de Estudos Linguìstico*, vol. 2, 255–274. Lisboa: Colibri / CLUNL.

Hölker, K. 2003. It. *diciamo* als Mitigator. In G. Held (ed.), *Partikeln und Höflichkeit*, 131–153. Bern: Peter Lang.

Khachatourian, E. 2006. *Les mots du discours formés à partir des verbes* dire / skazat' *(en français et en russe)*: Université Paris 7 dissertation.

Khachaturyan, E. 2008. Les mots du discours de *dire*. Analyse contrastive. In *Revista de Estudos Linguìsticos* 2, 287–315. Lisboa: Colibri / CLUNL.

Khachaturyan, E. & E. Vladimirska. 2010. *Vraiment / veramente*: analogies et contrastes. In *«Marqueurs du discours dans les langues romanes: une approche contrastive Madrid»*, .

Khaciaturian, E. 2005. Sul segnale discorsivo *senti*. In *Studi di grammatica italiana*, vol. XX, 127–137. Florence: Le Lettere.

Paillard, D. 1998. Les mots du discours comme mots de la langue I. *Le Gré des langues* 14. 10–41.

Paillard, D. 2001. Les mots du discours comme mots de la langue II. *Le Gré des langues* 16. 99–115.

Paillard, D. 2002. Les mots du discours. Identité sémantique et principes de variation. *Cahiers Linguistiques de l'INALCO* 4. 31–47.

Paillard, D. 2009. Prise en charge, commitment, ou scène énonciative. *Langue française* 162. 109–128.

Rossari, C. 1994. *Les opérations de reformulation: (analyse du processus et des marques dans une perspective contrastive, français-italien)*. Berne: Peter Lang.

Vladimirska, E. 2008. *Vraiment*: l'identité sémantique et les variations discursives. In *Revista de Estudos Linguìstico*, vol. 2, 255–274. Lisboa: Colibri / CLUNL.

AUTHOR CONTACT INFORMATION

Elizaveta Khachaturyan
Department of Literature, Area Studies and European Languages
University of Oslo
Norway
elizaveta.khachaturyan@ilos.uio.no

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 413–432. (ISSN 1890-9639)
http://www.journals.uio.no/osla

A CONTRASTIVE ANALYSIS OF RUSSIAN AND NORWEGIAN UTTERANCE-INITIAL COORDINATING CONJUNCTIONS

MARGJE POST
University of Bergen

ABSTRACT

In the study presented here, the three Russian basic additive and contrastive coordinating conjunctions *i*, *a* and *no* were compared to their two Norwegian counterparts *og* and *men* when used in utterance-initial position. By means of a direct comparison of sentences from Russian and Norwegian novels and their translations, both differences between the languages and language-internal boundaries between the conjunctions were made apparent. As a result of the study, a core meaning was formulated for each of the five conjunctions. According to the analysis presented here, the Russian conjunction *a* connects in a way fundamentally different from *i* and *no*. Metaphorically speaking, *i* and *no* can be said to connect on a horizontal, or syntagmatic, line, whereas *a* connects elements on a vertical, or paradigmatic axis. Unlike *i* and *no*, the conjunction *a* is implicationally unmarked for linear, logical connections. In Norwegian, *og* simply has an additive meaning, whereas *men* signals the existence of an element of conflict. These core meanings account not only for the use of these connectives in general, but they can also explain certain specific qualities and conditions for pragmatic use in utterance-initial position. An analysis in terms of core meanings needs to be supplied by contrastive studies on the basis of corpora, which show actual use of the words in almost all possible contexts.

[1] INTRODUCTION

This article presents a contrastive analysis of the counterparts of the basic additive and contrastive coordinating conjunctions *and* and *but* in Russian and Norwegian — the Russian coordinating conjunctions *i*, *a* and *no* and Norwegian *og* and *men*. The aim of the study was to give a better understanding of the properties of each of these conjunctions, in particular when they are used as pragmatic connectives in utterance-initial position. In some pragmatic uses these words can have more in common with discourse particles than with conjunctions. For instance, turn-initial *a* is usually considered to be a particle, cf. (Vasilyeva 1972; Rejmankova 1975; Rathmayr 1985; Šimchuk & Ščur 1999). Much could be said about the

conjunction or particle status of these utterance-initial connectives, but this topic will not be discussed here. For practical reasons the term *conjunction* will be used in this article to cover all uses, on a par with *connective*, which is meant as a general term for connecting words from a broader range of word classes.

For the present study a contrastive approach was chosen, because a comparison of two language systems demonstrates differences directly, not only differences between languages, but also boundaries between words in a single language system. As remarked by Jasinskaja & Zeevat (2008), an important role in the distribution of functions over words is played by the systemic factor: in which contexts and in which functions a particular conjunction can be used depends in part on what other conjunctions are available in the system of that language. In order to find subtle differences in pragmatic use of these conjunctions, the study was supported by a comparison of the actual use of the words in sentences from novels with their translations in the other language.

One of the main goals of the current paper is to show what comparative studies based on corpora and other empirical data can contribute to the study of connectives. Most of this study was based on a limited set of sentences from four Russian and four Norwegian (*bokmål*)[1] novels and their translations.[2] From each novel 10-15 pages with relatively much dialogue were chosen. The resulting approx. 100 pages in each language contains a total of 460 Russian and 334 of the Norwegian utterance-initial conjunctions in question (Post 1997). The development of Russian monolingual and bilingual corpora, such as the Russian National Corpus (RNC)[3] and the RuN corpus, a new Russian-Norwegian parallel corpus under development,[4] makes the study of much larger amounts of data more feasible. Unfortunately, it has not been possible to include more than some preliminary results based on data from these corpora in this paper, but an attempt will be done to show the relevance of the use of corpora.

For Norwegians, with their dual system, the choice between the three conjunctions in Russian is not always easy.[5] From a Norwegian, or English, point

[1] The Norwegian texts were restricted to the largest of the two Norwegian written standards, *bokmål*, because of the limited number of translations written in *nynorsk*.

[2] The eight novels and their translations represent six modern authors (two Norwegian and four Russian) and seven different translators (cf. Post (1997); the references below include most of them).

[3] The *Russian National Corpus* (RNC; русский национальный корпус) is a corpus consisting of a number of subcorpora, including monolingual Russian corpora of mainly written, but also transcribed spoken language, and two bilingual corpora: Russian-English and Russian-German (http://www.ruscorpora.ru).

[4] The RuN corpus is a Russian-Norwegian parallel translation corpus, consisting of mainly fictional texts and their translations, developed at the University of Oslo within the RuN project (http://www.hf.uio.no/ilos/forskning/forskningsprosjekter/run/).

[5] As a matter of fact, Russian has a fourth conjunction — *da* — and Norwegian a third one —*enn*, but they were not included in this study because of their low frequency and restricted distribution. *Da* is stylistically marked (for an extended description cf. Mendoza (1996)). The conjunction *enn* can in most varieties of Norwegian not introduce clauses, but only phrases in constructions like *«Enn De, da, Levin?»* («А вы, Левин?»; Л. Толстой; RuN corpus)

of view, the conjunction *a* is the odd one out. Russian utterance-initial *i* usually corresponds to Norwegian *og*, and *no* to *men*, but the conjunction *a* is about as often translated with additive *og* as with contrastive *men*.[6] The exact boundaries between *a* and *i* and between *a* and *no* have been an object for discussion for a long time, e.g. in Krejdlin & Padučeva (1974a,b); Kručinina (1988); Sannikov (1989); Uryson (2002, 2004); Jasinskaja & Zeevat (2008).

In both languages the conjunctions are frequently used in initial position, much more frequent than their counterparts in, for instance, English and Dutch; cf. Perridon (1987) .[7] The utterance-initial conjunctions are often left untranslated. A sample of 400 question utterances starting with a in RNC's Russian-English corpus shows that less than 50 % of these questions have a corresponding conjunction in English Røsstad (2009):[8]

(1) «*А* Собакевича знаешь?»
«Then you are acquainted with Sobakevitch?» (RNC, from Н. Гоголь)

(2) — *А* что ж? — Да так-с!
«Well»? «Just so, sir,» he answered. (RNC, from М. Лермонтов)

(3) «Listen, Nick; let me tell you what I said when she was born.»
— *А* хочешь знать, что я сказала, когда она родилась, Ник?
(RNC, from F. S. Fitzgerald)

In many cases a conjunction is even added, such as in (1) above, where *a* turns up in the Russian translation of a sentence with no conjunction in the source language. Despite their widespread use and the fact that they are frequently not translated or not even translatable, utterance-initial *i*, *a*, *no*, *og* and *men* have received little attention in the literature[9] and the conditions for their pragmatic use in this position have not yet been described in detail.

[2] WHAT IS SPECIAL ABOUT UTTERANCE-INITIAL USE?

I, *a*, *no*, *og* and *men* are coordinating conjunctions, i.e. in their prototypical use they connect two syntactically equal linguistic expressions to each other – words,

[6] This holds true only for the utterance-initial conjunctions, both in the sentences from the Russian novels that were gathered for the present study (see below) and in the RuN corpus. In general, *a* corresponds more often to *og* than to *men* (RuN corpus).

[7] In two translations from Norwegian to Dutch of a novel by Knut Hamsun a substantial number of instances of initial *og* were omitted in Dutch, despite the general tendency of translators to keep close to the original (Cannegieter 1997). The use of utterance-initial *og* (*och*) and *men* appears to be even more frequent in Danish and Swedish than in Norwegian; cf. (Hulthén 1947, 176–178) and (Øyslebø 1978, 224). Perridon (1987) cites a passage from August Strindberg in which *och* introduces almost every sentence.

[8] The reason why Røsstad chose to compare Russian to English, and not to Norweian, is that the RuN corpus of Russian and Norwegian prose is still under development and does not yet allow searching for question utterances.

[9] An exception is Mendoza's monograph on pragmatic use of *i*, *a* and *da* (Mendoza 1996).

phrases or clauses. Utterance-initial use shows that this is not always the case: the first constituent can be expressed much earlier than the second constituent, or not be expressed at all, but only presupposed. Use as the first word in an utterance means that the word connects not on the level of the sentence, but that it is a pragmatic connective on discourse level.[10] It also means that the connection is indirect: it is not immediately obvious which are the two elements that are being connected.

It is assumed that, on a mental level, coordinating conjunctions always express a link between two ideas or propositions.[11] The term *conjunct* is used to denote the two mental units that are connected. Utterance-initial use of coordinating conjunctions shows that not necessarily both ideas are represented by a linguistic expression. Only a part of the information that the speaker (or writer) intends to convey has an explicit linguistic expression. This means that the first conjunct can be non-verbal, such as in the next example, where *Men* signals a contrast between the actual behaviour of the addressee and his/her expected behaviour:

(4) Med uvant klossethet brøt han inn i avskjeden idet moren sto og trykket damene i hånden. [...] Det lyktes ham å få forkortet seremonien til det minst mulige.
«*Men* Lillelord, jeg tror du er gal,» sa moren, »Det er de voksne som skal si adjø først, det vet du meget godt.» (Borgen 1955, 108)
С необычной для него неловкостью он вторгся между ними как раз в ту минуту, когда мать прощалась за руку с сестрами Воллкварц. [...] Ему удалось по возможности сократить церемонию.
— Маленький Лорд, что с тобой, ты не в своем уме! — сказала мать. — Ты прекрасно знаешь, что сначала прощаются взрослые.
(Borgen 1979, 121)[12]

The coordinating conjunctions can even introduce a poem, a story, or even the title of a novel; the literature refers to several poems by Anna Achmatova (Kručinina 1988; Uryson 2002, 166) and to a short story by Kjell Askildsen (Øyslebø 1978). The title of one of the novels used for the current research actually starts with

[10] According to (Mendoza 1996, 57–58) all sentence-initial connectives, connecting across sentence boundaries, are pragmatic by definition, since every sentence consists of at least one speech act, and a connection is considered pragmatic if both conjuncts have their own illocution: «Eine Verknüpfung wird als pragmatisch bezeichnet, wenn beide Konjunkte eine eigene Illokution haben» (1996, 57).

[11] The reason for mentioning ideas in addition to propositions is that the connection triggering the use of the conjunctions can be associative and have little to do with propositional logic.

[12] Dutch has developed the same use of the contrastive conjunction as Norwegian: «*Maar* Kleine Lord, je bent niet wijs [...]» (Borgen [1983], 116)

i: *И дольше века длится день* (Ajtmatov 1983a), translated with *og* in Norwegian (Ajtmatov 1983a).[13]

In certain contexts the connectives have a pragmatic function quite distinct from their basic function as an additive or contrastive conjunction of two clauses. They are closer to pragmatic particles and other discourse markers than to conjunctions, fulfilling pragmatic functions, thus guiding the hearer in how to link the utterances to the linguistic and extra-linguistic context. In certain pragmatic uses they are called 'pleonastic'[14] or 'void of content'[15] in traditional reference books, such as the following construction in Norwegian, in (5), turning up in a translation from Russian (cf. section [6] below):

(5) *А* я-то думал [...] (Rasputin 1984, 82)
Og jeg som trodde [...] (Rasputin 1978, 91)[16]

However, according to the analysis presented here, they are not void of content. Even in such cases they can be said to have an invariant, core meaning, as described in the next section, and make a contribution to the utterance, although not necessarily on a lexico-semantic level. Some examples of such pragmatic use will be discussed in section [6].

[3] CORE MEANINGS

As a result of the study (Post 1997, 1999b,a), a core meaning was formulated for each of the five conjunctions. These core meanings, or basic properties, apply for the conjunctions in general, so they are not confined to utterance-intitial conjunctions. Core meanings have also been formulated in earlier literature. Among the core meanings proposed for Russian *a* are *juxtaposition* (*sopostavlenie*; AG 1980) and *distribution* (*a raspredelitel´noe*; Kručinina 1988); Sannikov (1989) describes *no* as expressing «nesootvetstvie norme, narušenie normal´nogo choda sobytij» (1989, 153)[17]. However, in most descriptions a polysemy point of view is taken, and if

[13] As pointed out by the anonymous reviewer, the title of this novel is a citation from a poem by Boris Pasternak. In the original text, this verse does not introduce the poem, but is actually used only in the nineteenth line. Interestingly, the Norwegian translation has kept the conjunction in place, but the English and Dutch translations lack a conjunction: *The Day Lasts More than a Hundred Years* (Aitmatov 1980); *De dag die langer duurde dan een eeuw* (Ajtmatov 1995). Possibly, only the Norwegian translator was aware of the fact that this was a citation, but it could also be a sign of the lesser frequency of utterance-initial use of coordinating conjunctions in English and Dutch as opposed to Russian and Norwegian.

[14] Beito (1986) refers to Norwegian *og* 'and' when starting a verse in folk songs or modern poetry as pleonastic ('eit pleonastisk *og* byrjar ofte ei line i folkeviser, stundom òg i nyare lyrikk'; (Beito 1986, 315)).

[15] The dictionary *Nynorskordboka* (NNO) on *og* 'som innhaldstom innleiing i visse utrop' on an example similar to (5). The counterpart dictionary of *bokmål*, Bokmålsordboka (BMO 2010) contains a very similar entry, but without the expression 'void of content'. In fact, *og* cannot be pleonastic and void of content in all respects, since the connective cannot be left out in the *og jeg som*-construction; cf. section [6] below.

[16] The Dutch translation contains the additive connective *en* in a full sentence: «*En* ik dacht nog wel [...]» (Raspoetin 1982, 111)

[17] I.e. *no* expresses a «divergence from the norm, a breach from the normal course of events».

core meanings are given, they play a minor role. This accounts especially for Russian *a*, which is almost always divided in three meanings, which are best known as (1) non-correspondence (*a nesootvetstvija*), (2) juxtaposition (*a sopostavlenija*) and (3) linking (*a prisoedinitel´noe*) (Krejdlin & Padučeva 1974a). An exception is the monosemy approach in Jasinskaja & Zeevat (2008).

The description given below was based on my own analysis and to a varying extent inspired by previous descriptions of coordinating conjunctions, including Prijatkina (1970); Krejdlin & Padučeva (1974a,b); Girke (1978); Yokoyama (1981); Sannikov (1989) and others, and it is closest to Mendoza's description of *i* and *a* (1996), which is the only work focussing on pragmatic use. The description is also compatible with Jasinskaja & Zeevat (2008)), although they take a different perspective.[18] For an extensive overview of previous literature on the Russian conjunctions, see Mendoza (1996). More recent descriptions with interesting viewpoints about the boundaries between sentence-internal *i*, *a* and *no* include Padučeva (1997); Uryson (2002, 2004); Zaliznjak & Mikaeljan (2009); Jasinskaja & Zeevat (2008).

The core meanings that were formulated as a result of the study (Post 1997, 1999b,a) are the following:

NORWEGIAN: In Norwegian, *og* simply has an additive meaning, whereas *men* signals the existence of an element of conflict.

RUSSIAN: In Russian, the conjunction *a* is different from *i* and *no* in the way the connection should be interpreted. Metaphorically speaking, *i* and *no* can be said to add a proposition on a linear, horizontal, or syntagmatic line, whereas *a* connects elements on a vertical, or paradigmatic axis. When *i* and *no* are used, the relation between the two elements connected by the conjunction is interpreted cumulatively, usually on a time line or as a causal relation – with *i* adding an element that is interpreted positively, and *no* introducing a deviation from the expected continuation. The conjunction *a*, however, is implicationally unmarked for such linear, logical connections (cf. Mendoza (1996) for an analysis of *i* and *a* along similar lines).[19] The latter simply introduces a new member to an existing set, about

[18] Jasinskaja & Zeevat (2008) give an interesting account of Russian *i*, *a* and *no* and English *and* and *but* by presenting the connectives as topic management devices, introducing answers to different kinds of (implicit and explicit) questions under discussion.

[19] Mendoza describes *i* as having a parallelisation effect (*Parallellisierungseffekt*): it can only connect conjuncts on the same semantic, thematic, argumentative or illocutive level (1996, 145, 236). The same appears to apply for *no*, with the additional meaning that the second part represents a divergence from the norm. Furthermore, *i* blocks an anti-iconic reading. This means that, if we call the structure «X *i* Y» and X and Y are arguments of the structure, then Y cannot precede X neither temporally nor causally (1996, 108; cf. Sannikov (1989) on *no*). «*I* und *no* haben [...] gemeinsam, daß sie die Konjunkte als abhängig repräsentieren (*но*) bzw. eine solche Interpretation nicht ausschließen (*и*). A hingegen initiiert keine Schlußfolgerungen über das logische Verhältnis der Konjunkte oder der in ihnen dargestellten Ereignisse.» (Mendoza 1996, 167)

which something new is expressed or implied, without implying any causal link with the preceding context.

In the literature, the difference between *a* and *no* in sentences where both can be used is often considered a difference in stylistics, with *a* representing a colloquial style. This difference can be explained by the proposed core meanings: logical, argumentative reasoning, strengthened by the use of *no*, is typical of prepared written texts, whereas associative thinking, supported by *a*, is more common in spontaneous speech.

The difference of *a* from *i* and *no* is reflected in sentence-internal information structure. In terms of the theory of actual sentence perspective (Krejdlin & Padučeva 1974a,b), only *i* and *no* can introduce *rhemes* (*R*); □ is typically followed by a pitch accented element that can be described as a new *theme* (*T*).[20] The connective a cannot be used to introduce rhematic predicates, coreferring to the same subject:

(6) *T* *R*$_1$ *R*$_2$ *R*$_3$
[Он] [встал], [подошел ко мне], *u/*a* [улыбнулся].
'He stood up, walked towards me, *i/*a* smiled' (Yokoyama 1981)

A typical example of *a* is the use of *a* between two sets of theme-rheme pairs that are put up against each other; cf. (7), from Berkov's Russian-Norwegian dictionary (2007):

(7) a. *T*$_1$ *R*$_1$ *T*$_2$ *R*$_2$
RU [Ты] [молод], *a* [он] [стар].

b. *T* *R*$_1$ *T*$_2$ R$_2$
NO [Du] er [ung], *men {og}* [han] er [gammel].

The conjunction *a* is also structurally different from *i* and *no*. According to Prijatkina (1970), *a* cannot normally express a connection between two words by itself, without the help of another word or expression (1970, 190–191). This other word or expression is either an adverb, or adverbial expression, or another emphasised element, when *a* connects two elements of the same category, as in the theme-rheme structures given above (Prijatkina 1970, 192–194). Thus, *a* is strongly connected to the word or expression following immediately after it, which car-

[20] *T = theme*; *R = rheme*. Roughly speaking, the *theme* marks what the sentence is about; the *rheme* expresses what is being said about this theme and represents the most important communicative part of the sentence (Švedova et al. 1980, 91).

ries a pitch-accent and often emphasis. In contrast, *a* itself is unstressed and can normally not be followed by a pause.

[4] CORE MEANINGS ACCOUNT FOR UTTERANCE-INITIAL CONJUNCTIONS

The small data set of sentences from Russian and Norwegian novels shows that the proposed core meanings are valid even for the utterance-initial conjunctions. They can even explain many of the specific conditions for and constraints on pragmatic use in utterance-initial position, as will be argued below.

The following examples of the connectives introducing additional questions show a clear parallel to sentence-internal information structure. In (8), the waiter asks, after receiving a first order of four sweet pastries,[21] what else the guests would like to order:

(8) «*Og* sjokolade med krem?» (Borgen 1955, 109)
— *И* шоколад со сбитыми сливками? (Borgen 1979, 122)

The appropriate conjunction to introduce this additional question is *i*, because *i* introduces a new element that corresponds to an addition to the rheme in the sentence «The guests (= theme T) would like to order four sweet pastries (= rheme R_1) *and* the guests (= same theme T) would like to order hot chocolate with cream (= rheme R_2)».

In the wh-question in (9), however, the waiter's question introduces the second theme and asks for the content of the second rheme in a double theme-rheme pair «T_1 orders R_1, and T_2 (= the young man) orders R_2 (= unknown)»:

(9) «*Og* til den unge herre?» (Borgen 1955, 114)
— *А* для молодого человека? (Borgen 1979, 128)

The additive connective *i* can generally be regarded as introducing a positive answer to a (real or presupposed) question, and *no* starts a negative answer, cf. Girke (1978); Jasinskaja & Zeevat (2008). In the domain of narrative structure, *i* is often used to start a positive (10) and *no* a negative continuation of the subtopic (11):

(10) «Kanskje vi skulle spise noe.»
Og det gjør de. (Borgen 1970, 54)
— Наверно, пора нам что-нибудь поесть.
И они едят. (Borgen 1979, 610)

(11) — Андрей, знаешь что?
— Что?
Но она передумала. — Ладно, потом. (Rasputin 1984, 80)
«Andrej, vet du hva?»

[21] «Fire wienerbrød?» (Borgen 1955, 109)

«Hva?»
Men hun betenkte seg. «Det var ingenting, senere.» (Rasputin 1978, 89)

On the other hand, *a* introduces not an answer, but something else, e.g. a change of perspective to a new subtopic (Mendoza 1996). In the next example, a introduces not a reaction, but a simple change of perspective:

(12) Ах, Сарала, ах, старина, мой славный конь, неужто жизнь так прекрасна, что даже в свой последний срок любить можно ...
А Сарала шагал дорожным ходом, пофыркивал, спеша домой, чтобы ногам дать отдых. (Ajtmatov 1983a, 439)
Akk, Sarala, akk du gamle, stolte hest, er livet virkelig så vidunderlig at man endog på sine siste dager kan få elske slik?...
Og Sarala gikk fram i skritt, prustet og skyndte seg hjem for å hvile bena. (Ajtmatov 1983b, 282-283)[22]

Similar to pragmatic particles (Foolen 2003 [1996]), the additive and contrastive conjunctions can signal the background or foreground status of a stretch of discourse. In argumentative discourse, *i* typically introduces a new argument, *no* introduces counterarguments, whereas *a* introduces something else, typically, a digression from the line of reasoning. In narrative structure, *i* and *no* add continuations to the current topic, whereas *a* is used to introduce digressions from the current topic, (sub)topic shifts, or returns to a previous topic (cf. (Mendoza 1996)). *I* can also be used to return to the main story line after a short digression (13a), where the digression itself can be introduced by *a* (13b). Metaphorically speaking, *i* continues the current line, whereas a introduces a transition to a new line:

(13) А третий был то что у меня отец-то был межевой инженер//

a. *А* межевой инженер и геодезист это очень близко //

b. *И* его же знакомый преподавал / однокашник по межевому институту преподавал геодезиу там в лесном институте (Mendoza 1996, 179, cited from a text collection of transcribed spontaneous speech)

In this section most attention has been paid to the Russian conjunctions, because the Russian system is more complex. This does by no means mean that the Norwegian counterparts are too boring to be investigated. Norwegian *og* and *men* can

[22] Contrary to the Norwegian translation, *a* was not translated in the Dutch translation (Ajtmatov 1995); cf. note (13).

fulfill discourse structuring functions similar to Russian *a* and *i*. *Og* can continue the story line, even at some distance in a new paragraph,[23] and a return to the main story line can be introduced by *men*. *Men* is also used at topic shifts. However, this does not mean that the Russian and Norwegian connectives always can be translated with each other whenever they fulfill similar functions. To show this, some examples will be discussed in section [6] below. A detailed contrastive analysis, however, with a precise demarcation of the pragmatic uses of the connectives, remains a task for future research.

[5] NORWEGIAN AND RUSSIAN UTTERANCE-INITIAL USE COMPARED

When comparing the Russian tripartite system to Norwegian, we expect that the distribution of *i*, *a* and *no* over *og* and *men* is very simple: as a rule, *men* is chosen when there is an element of conflict that deserves attention, such as a contrast or a denial of expectation, whereas *og* is used to express simple addition. The difference between Russian *a* and Norwegian *og* and *men* is shown in expressions where two sets of theme-rheme pairs are compared, as in (7), repeated below as (14). Unlike a speaker of Russian, a Norwegian language user has to express whether or not the difference between the pairs is in conflict with some expectation:

(14) a. T_1 R_1 T_2 R_2
[Ты] [молод], *а* [он] [стар].
b. T R_1 T_2 R_2
No [Du] er [ung], *men {og}* [han] er [gammel]. (Berkov 2007)

Below are two examples from the Russian and Norwegian novels: (15) introduces a positive reaction, (16) a negative one:

(15) Я [...] говорю: «Нюрка, это же генерал». *А* он мне: «Да, говорит, сынок, я и есть, говорит, генерал». (Pristavkin 1995, 11)
Jeg [...] sa: «Njurka, det er jo en general». *Og* han sa til meg: «Ja, gutten min, det er det jeg er, jeg er general». (Pristavkin 1991, 14)

(16) Он сказал подавленно: — Иди первый! Ты умный!
А Сашка ответил: — Ты тоже не дурак! Чего это я пойду?
(Pristavkin 1995, 166)
Han sa slukkøret: «Gå først! Du er flinkest!»
Men Sasjka svarte: «Du er ikke dum, du heller! Hvorfor skal jeg gå!»
(Pristavkin 1991, 208)

[23] Meaning no. 8 of *och* in the large dictionary of the Swedish Academy (SAOB 1898–) describes the use of *och* as introducing a new paragraph to mark that the first sentence is a direct continuation of the main presentation (topic), regardless of what has been mentioned in the immediately preceding context. Sometimes *och* changes into a general, comparatively neutral initial particle or initial interjection.

In the excerpts from the Russian and Norwegian novels and their translations used for this study (see section [1] above), the two Norwegian connectives are a little less frequent than their three Russian counterparts. The least frequent utterance-initial connective is *no*, as could be expected, since this word has the most specific meaning. Russian *a* corresponds as often to *og* as to *men*, but in more than half of the utterances *a* does not correspond to *og* or *men* in the Norwegian text. Norwegian *og* corresponds to *i* in 45 % of the sentences, and to *a* in 31 %, the remaining 24 % having no corresponding coordinating conjunction in the Russian texts.

In the RuN corpus the distribution is somewhat different: Russian *A* is much less frequent and *No* much more frequent than in the novel excerpts.[24] This is not surprising, because the excerpts in my own mini-corpus contain relatively much dialogue, in which utterance-initial *a* is frequent, especially turn-initially.

The large number of non-correspondences in the translations is the topic of the next section.

[6] BILINGUAL CORPORA REVEAL SPECIFIC USES

Corpora can be used not only to find empirical support for predefined hypotheses. Just like cross-linguistic studies, they can also reveal unexpected properties of words that cannot be predicted from a general theory and that could remain unremarked in studies based on introspection. Corpus studies can reveal minor differences in the pragmatic uses of words, such as differences in possible contexts and in frequency.

In the mini-corpus of Russian and Norwegian novels, 35% of the sentence-initial conjunctions did not correspond to a conjunction in the other language.[25] So, in these cases, the translators deliberately chose not to translate the connectives with one of its counterparts in the target language, or they chose to add one in the translation. Part of these cases can be ascribed to the liberty of a translator of fiction to refrain from direct translations, but in other cases they reflect subtle pragmatic differences. In other cases, the use of a conjunction would even be infelicitous in the other language.

The core meanings described above cannot predict the exact conditions of use of pragmatic *i*, *a*, *no*, *og* and *men*. Pragmatic uses in specific contexts are often

[24] As of April 2010, the RuN corpus, utterance-initial *i* is used 4138 times, *no* – 3502 times and *a* – 2850 times. The numbers for Norwegian are 5006 (*men*) and 4298 (*og*). The largest differences with the mini-corpus are that Norwegian *og* corresponds to *a* in only 17% of its occurrences, but to *i* in 60%, and that *men* corresponds to *no* far more often (58%) than to *a* (13%), but even here the number of non-correspondences is substantial: 29%. The RuN corpus gives the same proportions between utterance-initial *a* and its Norwegian counterparts (see above): 23% □ –*men*, 25% *a* – *og* and 51% non-correspondence.

[25] The numbers for the individual conjunctions lacking a corresponding conjunction in the other language were as follows: 50% of *i* (including the small group of occurrences of the focus particle *i*), 53% of *a*, 15% of *no*, 24% of *og* and 21% of *men* in initial position.

conventionally based uses that are language-specific and have to be learnt by the language user. However, the proposed basic properties can help explain these specific uses, as shown by some examples in the remainder of this section.

In the following example, *a* has been added to a question in the translation (cf. (1), (2), (3) above):

(17) Han sa: «Har du vært i Bretagne? Vi reiser til Bretagne.»
(Borgen 1970, 111)
Он вдруг спросил:
— *А* ты бывала когда-нибудь в Бретани? Едем в Бретань!
(Borgen 1979, 658)

A high frequency of initial *a* in questions can be expected because *a* can introduce a new member to an existing set of questions, independent of temporal or causal links. Therefore, it can introduce almost any question. A possible explanation for the much lower frequency of *og* and *men* in questions is that the Norwegian connectives are too easily associated with temporal and causal links.

Even *og* and *men* remain untranslated in a substantial part of the sentences. In the following examples, *men* has discourse structuring functions. In (18), this word introduces a return to a previous topic after a long digression; in (19) it introduces a topic shift in a new paragraph:

(18) *Men* nå lurte han på hvem [...] (Holt 1975, 122-123)
Так вот, он теперь решил разузнать [...] (Cholt 1982, 282)

(19) *Men* nå skal du høre på meg. (Holt 1975, 126)
Теперь послушай, что еще скажу. (Cholt 1982, 285)

Although the Russian conjunctions can have similar functions, as we saw in section [4], and translations with *i* in (18) and *a* in (19) could actually have been possible, the translators chose for some reason it was better to leave out the conjunctions. In both cases the digressions were comparably long. Possibly, Russian *i* allows a shorter digression than Norwegian *men* to be able to signal a return to the main topic. Other possible reasons for omitting *men* include the strong link of *i* and *no* to temporal and causal relations, while *a* can give too strong associations to comparisons between two situations if the context is favourable for such an interpretation, like here in (19), where *a* would have been followed by the adverb *teper´* 'now', as a translation of Norwegian *Men nå*.[26] The connective *a* cannot in itself suggest temporal and causal relations (Mendoza 1996, 167). This feature of *a* restricts its ability to connect over larger distances, if the connection is of a

[26] This does not mean that *a* can never be used after long pauses; cf. (17) «А ты бывала когда-нибудь в Бретани?» In this case, however, the context does not suggest temporal, causal or comparative relations.

mainly temporal or causal nature. As remarked in section [5], this topic has not yet been studied thoroughly.

A trait of Norwegian that could remain unnoticed without consulting bilingual corpora is the frequent use of the answering particles *ja*, *nei* and *jo* in this language. Below are two examples, where the answering particles were added to the Norwegian text in translations from Russian:

(20) «*А* по-вашему что??»
«*Ja*, hva mener De?» (RuN corpus; from Б. Пильняк)

(21) «Однако пора уже»
«*Nei, men* nå er det sandelig på tide å reise» (RuN corpus; from Л. Толстой)

Ja, *jo* and *nei* often appear in combination with *men*. *Ja* and *nei* have even merged with *men* and developed into the discourse particles *jammen*, *neimen* (BMO). Of course, also in Russian the conjunctions in question have developed into idiomatic expressions in combination with other words, such as *а то*, described in Uryson (2008). These new idioms have developed their own meanings and functions. Therefore, they deserve separate studies.

Another example of the subtle differences between Russian and Norwegian shown in the novels is the frequent use of *men* in addresses in Norwegian, functioning as surprised reactions, of which (4) is an example («Men Lillelord!»). In these cases, *men* has no counterpart in the Russian parallel texts from the novel excerpts (22), nor in the translation from the example from Ibsen mentioned in a dictionary (RMO):

(22) — *Men* kjære deg, det er da bare uhyggelig. (Borgen 1955, 112)
— Милый, но ведь это ужасное место! (Borgen 1979, 125)

(23) brændt! *Men* du godeste Gud –! Nej, nej, nej, dette er rent umuligt!
(Ibsen: Gabl., from RMO)
Сожгла! Боже милосердный! .. Нет, нет, нет! Это невозможно!
(Ибсен 1972, 616)

However, the construction occurred only three times in the mini-corpus,[27] so larger data bases were required to find more attestations of this construction. The RuN corpus contains no more than a few examples. There were no occurrences of the conjunction *а*. Some examples contain *но*, but only in translations from Norwegian and invariably followed by a comma.[28]

[27] It also contained an example of «*Ja men* + [name]».

[28] Berkov's dictionary actually includes the addressing construction:
но, дорогой мой! *men* kjære Dem deg da! (Berkov 2007)
The RuN corpus indeed contains a few examples in the translations from Norwegian, e.g. «Но, Дина! Его же повесят!», translated from «Men Dina! De skal henge en mann!» (RuN corpus; from H. Wassmo).

A check in RNC reveals absence of the conjunction *a* and a very low frequency of *no* in addresses containing personal names, and in none of the cases of *a*, this word could be interpreted as a conjunction in an addressing construction. *A* does occur in addresses, but in these cases, *a* is not related to the conjunction and it has a different prosody; cf. the following two examples from RNC:

(24) «Наташ, а Наташ!» (RNC; from Ю. Трифонов)

(25) «А, Федя! Дома Хорь?» (RNC; from И. Тургенев).

In the few occurrences starting with *no*, this connective was followed by a comma.[29] This suggests that the name used as an address was syntactically not part of the sentence introduced by *no* and that *no* probably was followed by a prosodic boundary, in which case *no* introduced not the name, but a sentence expressing a negation of an expected continuation. In Norwegian (as in Dutch) there is no prosodic boundary between the conjunction and the name; a boundary occurs only after the name.[30]

The low occurrence of *no* in these constructions could be related to the strong connection with causality of *no*. Interestingly, the translation in (23) does contain *no*, but only after the address, introducing the expression of contrast itself. *Men* is first of all used to underline the speaker's surprise (cf. NNO 2010), and not necessarily to introduce the linguistic expression of a second, contrastive, conjunct.

A final example of a conventionally based use of an utterance-initial connective is the Norwegian contruction *Og jeg som* [...], literally 'And I who [...]', given earlier in (5), and in (26) below, both in translations from Russian:

(26) «*Og* jeg som ventet på deg til klokken var to.»
—*А* я тебя ждал до двух часов. (RuN corpus; from Л. Толстой)

It is a rather frequent syntactic elliptic construction with left dislocation of the subject, which is used to express surprise or indignation over a contrast between two situations or facts.[31] This is obtained through juxtaposition of the situation involving the speaker with a previously activated situation by connecting them

[29] There is in fact a single example of *no* + a name, used as an address form without a preceding comma, but it was in free direct speech: Я глупо повторял про себя: Гонец из Пизы, из Ганы. Я был взбешен. Где же был этот хваленый рирайт, для чего, собственно, он нужен, как не для того, чтобы именно такие описки и исправлять! *Но Оля*! Что же вы-то, Оля, сплоховали с этим самым Константином? (RNC; from Н. Климонтович).

[30] However, as remarked by the anonymous reviewer, the presence of a comma does not necessarily imply the presence of a prosodic boundary in all of the Russian examples, since Russian punctuation is rather prescriptive.

[31] This *og jeg som*-construction is, of course, not restricted to the first person personal pronoun, but can also be used with other nouns and pronouns referring to persons. It is common in Swedish as well: SAOB describes it as introducing an ellipic sentence that consists of a noun or pronoun, determined by a following relative clause: "särskilt inledande en elliptisk sats som består av sbst. l. pron., bestämt av en följande relativsats: *Och jag, som letat så!*" SAOB (1898–).

with the additive connective *og*. At the current stage, the RuN corpus contains nine examples of this construction, all of which correspond to *а я* [...] in Russian, often supported by a particle emphasising the pronoun. In Russian, what follows is not the Norwegian elliptic construction of a noun phrase with a subordinate clause, but a syntactically complete sentence. The occurrence of *а* is expected, since the construction is comparative, putting up to pairs of alternatives against each other, comparable to two theme-rheme pairs.

[7] CONCLUSION

In the study presented above, the three Russian basic additive and contrastive coordinating conjunctions *i*, *а* and *no* were compared to their two Norwegian counterparts *og* and *men* when used in utterance-initial position. A core meaning was formulated for each of the five conjunctions. According to the analysis presented here, the Russian conjunction *а* connects in a way fundamentally different from *i* and *no*. Metaphorically speaking, *i* and *no* can be said to connect on a horizontal, or syntagmatic, line, whereas *а* connects elements on a vertical, or paradigmatic axis. Unlike *i* and *no*, the conjunction *а* is implicationally unmarked for linear, logical connections. In Norwegian, *og* simply has an additive meaning, whereas *men* signals the existence of an element of conflict.

The small corpus of sentences from Russian and Norwegian novels used for this study shows that there are clear parallels between the basic, intrasentential use of the additive and contrastive conjunctions in Russian and Norwegian and their pragmatic use in utterance-initial position. Their basic properties account not only for their use in general, but they can also explain certain specific qualities and conditions for pragmatic use in utterance-initial position. The core meanings, or basic properties, of Russian *i*, *а* and *no* and Norwegian *og* and *men* as proposed in this study help to explain why not all conjunctions are used in the same context with the same function. However, the formulation of a core meaning alone is insufficient for a good description of a word's meaning and functions, since a core meaning, which by necessity must be rather abstract, cannot predict the precise boundaries between two similar words, nor the exact conditions of use of a word. Therefore, an analysis of basic properties needs to be supplied by contrastive studies on the basis of corpora, which show actual use of the words in almost all possible contexts. These data show, for instance, that both the Russian and the Norwegian conjunctions can be used in the domain of narrative structure, introducing a continuation of the topic (*i*, *og*) or a return to a previous topic after a short digression (*а*, *men*), but that their conditions for use in these functions do not fully coincide. The data from novels and their translations also show language-specific developments. For instance, Norwegian *men* is often combined with an address form, unlike its Russian counterparts, and in Norwegian, answering particles can be used as discourse particles, turning up in combination with

og and *men*. The recent appearance of annotated corpora and the development of a Russian-Norwegian corpus will enable quick searches in large amounts of data to find empirical support for hypotheses and findings, and to reveal unexpected properties. Monolingual corpora, which for natural reasons are usually much larger than bilingual corpora, can supply evidence on lesser used words and constructions and answer questions whether certain contexts or constructions actually occur or not.

Many open questions remain to be addressed in future research. An interesting study would be a contrastive analysis of the Russian conjunctions and its counterparts in other Slavonic languages, which have different sets of additive and contrastive conjunctions (cf. Rejmankova 1975; Freidhof 1991; Gvozdanovič 1996), or with a comparison with Old Russian, with its frequent sentence-initial use of *a* and *i*, for instance in the birch-bark letters (Zaliznjak 2004).

The number of occurrences analysed for the present study is too low to allow for a detailed analysis explaining all cases of pragmatic use. Therefore, the results presented in this article can only be considered preliminary. However, the analysis of the mini-corpus and the first results from the study of the larger corpora reveal some interesting first results, showing the usefulness of corpora for this kind of linguistic studies.

ACKNOWLEDGMENTS

The author wishes to thank the anonymous reviewer for useful comments.

[8] PRIMARY SOURCES

Aitmatov, Ch. 1980. *The Day Lasts More than a Hundred Years.* Translated into English by F. J. French. Bloomington: Indiana University Press

Ajtmatov, Č. 1983a. *I dol´še veka dlitsja den´.* In Sobranie sočinenij v trech tomach. Moscow: Molodaja gvardija

Ajtmatov, T. 1983b. *Og dagen varer lenger enn et århundre.* Translated from Russian into Norwegian by E. Kolstad. Oslo: Tiden

Ajtmatov, T. 1995. *De dag die langer duurde dan een eeuw.* Translated from Russian into Dutch by M. Weijers. Breda: De Geus

Borgen, J. 1955. *Lillelord.* Roman. Oslo: Gyldendal

Borgen, J. 1970. *Vi har ham nå.* [s.l.]: Den norske bokklubben

Borgen Ju. 1979. *Malen´kij Lord. Temnye istočniki. Teper´ emu ne ujti.* Translated from Norwegian into Russian by Ja. Jachnina and S. Tarchanova. Moscow: Progress

Borgen, J. [1983]. *De Kleine Lord.* Translated from Norwegian into Dutch by R. Schoorl. Weesp: Agathon

Cholt, K. 1982. *Morskoj geroj.* Translated from Norwegian into Russian by L. Ždanova. In Trizna po ženščine. Morskoj geroj: Romany. Moscow: Raduga Holt, K. 1975. Sjøhelten. Oslo: Gyldendal

Ibsen, G. 1972. *Dramy. Stichotvorenija.* Plays translated from Norwegian into Russian by A. and P. Ganzen. Moscow: Chudožestvennaja literatura

Pristavkin, A. 1991. *Nattely for en gyllen sky.* Translated from Russian into Norwegian by E. Egeberg. Oslo: Gyldendal

Pristavkin, A. 1995. *Nočevala tučka zolotaja.* In Izbrannaja proza. Moscow: Kvadrat

Raspoetin, V. 1982. *Vlucht naar de wouden.* Translation from Russian into Dutch by A. Pries. Amsterdam: De arbeiderspers

Rasputin, V. 1978. *Aldri glemme.* Translated from Russian into Norwegian by G. Opeide. Oslo: Gyldendal

Rasputin, V. 1984. *Živi i pomni.* In Izbrannye proizvedenija v dvuch tomach. Tom 2. Moscow: Molodaja gvardija

RNC = Russian National Corpus (Русский национальный корпус), www.ruscorpora.ru

RuN corpus = Russian-Norwegian Parallel Corpus, http://www.tekstlab.uio.no/glossa/html/index_dev.php?corpus=run

REFERENCES

Beito, O. T. 1986. *Nynorsk grammatikk. lyd- og ordlære*. Oslo: Det norske samlaget.

BMO. 2010. *Bokmålsordboka: definisjons- og rettskrivningsordbok*. Oslo: Kunnskapsforlaget. `www.dokpro.uio.no/ordboksoek.html`.

Cannegieter, M. 1997. *Analyse av stilforskyvninger i to nederlandske oversettelser av Knut Hamsuns roman 'Under høststjernen' (1906)*. Universiteit van Amsterdam MA thesis.

Foolen, A. 2003 [1996]. Pragmatic particles. In J. Verschueren, J.-O Östman & Ch. Bulcaen (eds.), *Handbook of pragmatics online*, `www.benjamins.com/online/hop/`.

Freidhof, G. 1991. Umgangssprachliche Eröffnungssignale in Dialogen der Schönen Litteratur (Narrative Prosa). In K. Hartenstein & H. Jachnow (eds.), *Slavistische Linguistik 1990 (= Slavistische Beiträge 274)*, 71–97. München.

Girke, W. 1978. Sätze mit *no*, *i* und *a*. In W. Girke & H. Jachnow (eds.), *Slavistische linguistik 1977*, 7–26. München.

Gvozdanovič, J. 1996. 'and' has four different correlates in western south slavic. *Studies in South Slavic and Balkan Linguistics (= Studies in Slavic and General Linguistics 23)* 24. 101–109.

Hulthén, L. 1947. *Studier i jämförande nynordisk syntax ii*, vol. LIII. Göteborg: Göteborgs Högskolas Årsskrift.

Jasinskaja, K. & H. Zeevat. 2008. Explaining Additive, Adversative and Contrast Marking in Russian and English. *Revue de Sémantique et Pragmatique* (24). 65–91.

Krejdlin, G. E. & E. V. Padučeva. 1974a. Vzaimodejstvie associativnych svjazej i aktual´nogo členenija v predloženijach s sojuzom *a*. *Naučno-techničeskaja informacija* 2(10). 32–37.

Krejdlin, G. E. & E. V. Padučeva. 1974b. Značenie i sintaktičeskie svojstva sojuza *a*. *Naučno-techničeskaja informacija* 2(9). 31–37.

Kručinina, I. N. 1988. *Struktura i funkcii sočinitel´noj svjazi v russkom jazyke*. Moscow.

Mendoza, I. 1996. *Zur Koordination im Russischen: **i**, **a**, und **da** als pragmatische Konnektoren (= Slavistische Beiträge 338)*. München: Otto Sagner.

NNO. 2010. *Nynorskordboka. definisjons- og rettskrivingsordbok*. Oslo: Det norske Samlaget. `www.dokpro.uio.no/ordboksoek.html`.

Padučeva, E. V. 1997. Ėgocentričeskaja semantika sojuzov a i no. In V. S. Efimova T. M. Nikolaeva & I. Fougeron (eds.), *Slavjanskie sočinitel´nye sojuzy*, 36–47. Moscow.

Perridon, H. 1987. 'och' en 'og'. een kontrastieve studie naar het gebruik van de voegwoorden zweeds 'och', deens 'og' en nederlands 'en'. *Tijdschrift voor Skandinavistiek* 8(1). 3–39.

Post, M. 1997. *I, A, No, Og en Men: over zinsinleidende voegwoorden in het Russisch en het Noors*. Universiteit van Amsterdam MA thesis.

Post, M. 1999a. *I*, *A* i *No* v sopostavlenii s sosedstvujuščimi sojuzami v germanskich jazykach. In O. Kovačičova et al. (eds.), *Russkij jazyk, literatura i kul´tura na rubeže vekov*, vol. I, 140–141. Bratislava: IX meždunarodnyj kongres MAPRJaL. Tezisy dokladov i soobščenij.

Post, Margje. 1999b. *A, I* og *No* mot *Og* og *Men*: en kontrastiv undersøkelse av tre russiske og to norske konjunksjoner i setningsinnledende posisjon. *Nordlyd* 27. 24–43.

Prijatkina, A. F. 1970. Konstruktivnye osobennosti sojuza *a* v prostom predloženii russkogo jazyka. In T. P. Lomtev & A.A. Kamynina (eds.), *Issledovanija po sovremennomu russkomu jazyku: sbornik statej, posvjaščennyj pamjati e.m. galkinoj-fedoruk*, 190–205. Moscow.

Rathmayr, R. 1985. *Die russische Partikeln als Pragmalexeme*. (= Slavistische Beiträge 187).

Rejmankova, L. 1975. Ob upotreblenii častic pri osuščestvlenii svjazi meždu replikami. *Československá rusistika* 16–19.

Røsstad, K. 2009. *En korpusundersøkelse av konjunksjonen **a** i russiske originaltekster og i oversatte tekster*. Bergen. University of Bergen MA thesis.

Sannikov, V. Z. 1989. *Russkie sočinitel´nye konstrukcii: Semantika. Pragmatika. Sintaksis.* Moscow: Nauka.

SAOB. 1898–. *Ordbok över svenska språket*. Lund: Utgiven av Svenska Akademien. `http://g3.spraakdata.gu.se/saob/`.

Uryson, E. V. 2002. Sojuz A kak signal «povorota povestvovanija». In N. D. Arutjunova (ed.), *Logičeskij analiz russkogo jazyka: Semantika načala i konca.*, 348–357. Moscow.

Uryson, E. V. 2004. *Sojuzy A i NO i figura govorjaščego*. Voprosy jazykoznanija 2004 (6).

Uryson, E. V. 2008. Sojuzy A TO i A NE TO: Počemu v nekotorych kontekstach oni sinonimičny. In *Trudy meždunarodnoj konferencii «Dialog 2008». Moscow*, 530–538. `http://www.dialog-21.ru/dialog2008/materials/pdf/82.pdf`.

Vasilyeva, A. N. 1972. *Particles in colloquial Russian*. Progress.

Šimchuk, E. & M. Ščur. 1999. *Slovar´ russkikh častic*. Peter Lang AG.

Švedova, N. Ju et al. 1980. *Grammatika russkogo jazyka. vol. i, ii.* Moscow: Nauka.

Yokoyama, O. T. 1981. On sentence coordination in Russian: A functional approach. In R. A. Hendrik, C. S. Masek & M. F. Miller (eds.), *Papers from the seventeenth regional meeting of the Chicago Linguistic Society. Chicago*, 431–438.

Zaliznjak, A. A. 2004. *Drevnenovgorodskij dialekt*. `http://gramoty.ru/?id=dnd`.

Zaliznjak, A. A. & I. Mikaeljan. 2009. Russkij sojuz A kak lingvospecifičeskoe slovo. In *Dialog 2005. Trudy meždunarodnoj konferencii. Retrieved 13*, vol. 12., `http://www.dialog-21.ru/Archive/2005/Zalizniak_Mikaelian/Zalizniak_Mikaelian.pdf`.

Øyslebø, O. 1978. *Stil- og språkbruksanalyse*. Oslo.

AUTHOR CONTACT INFORMATION

Margje Post
Institutt for framandspråk
Universitetet i Bergen
Postboks 7805, 5020 Bergen
Norway
Margje.Post@if.uib.no

OSLa A. Grønn & I. Marijanovic (eds.) *Russian in Contrast*, Oslo Studies in Language 2(2), 2010. 433–466. (ISSN 1890-9639)
http://www.journals.uio.no/osla

CORRECTIVE CONTRAST IN RUSSIAN, IN CONTRAST

KATJA JASINSKAJA
University of Heidelberg

ABSTRACT

In many languages, correction (e.g. *John didn't go to Paris, but to Berlin*) and various kinds of contrastive relations are often expressed by the same marker (e.g. *but* in English). In Russian, correction is marked by the conjunction *a* as part of what is often believed to be a fixed collocation *ne ..., a ...* 'not ..., but...', but conjunction *a* can also be used to encode, for instance, contrastive comparison *Oleg ljubit futbol, a Roma basketbol* 'Oleg likes football, and Roma likes basketball'. This paper addresses two issues. First, it tries to define a single 'meaning' of *a* that covers both its contrastive and corrective uses. Second, concentrating on the differences between Russian and English, it investigates the structure of the contrast-correction semantic space to predict which of a range of contrast markers of a given language are more likely to be used in the correction function than the others.

[1] INTRODUCTION

This paper is concerned with corrective uses of contrastive markers, such as the Russian conjunction *a* and the English conjunction *but*, illustrated in (1) and (2), respectively. It is characteristic for these uses that some contextually salient proposition is explicitly negated in one conjunct (*John didn't go to Paris*), while the other conjunct (*to Berlin*) presents an element that should "replace" the wrong part of the negated proposition (*to Paris*).[1]

(1) a. Oleg ezdil *ne* v Pariž, *a* v Berlin
Oleg went not to Paris but to Berlin
b. Oleg ezdil v Berlin, *a* *ne* v Pariž
Oleg went to Berlin but not to Paris

(2) John did*n't* go to Paris, *but* to Berlin.

[1] In German this function is unambiguously expressed by the conjunction *sondern*, (i). This can be used as a heuristic if it helps to understand our terminology: correction is roughly *the* function that is expressed by *sondern* in German.

(i) Hans ist nicht nach Paris *sondern* nach Berlin gefahren
Hans is not to Paris but to *Berlin* gone

This understanding of the term *correction* is common in descriptive and typological literature (e.g. Malchukov 2004; Mauri 2008), where it figures next to the *additive* and the *adversative* type in various functional classifications of coordinative constructions. To prevent terminological confusion, this notion should be distinguished from the *speech act* of correction in e.g. Asher and Lascarides (2003, pp. 345–350), such as the utterance (a) of speaker B in (3).

(3) A: They gave Peter the new computer.
B: a. No, they gave JOHN the new computer.
b. No, they didn't give it to PETER, but to JOHN.

Of course, correction as a type of coordinative construction in (1)–(2) can be used to perform correction as a speech act, cf. (b) by speaker B in (3). One might even argue that from an evolutionary point of view this is the primary use of corrective coordination. However, both corrective coordination has other uses, and the correcting speech act can be done by other means.[2] The focus of this paper is on corrective coordination.

Apart from correction, the Russian conjunction *a* has other functions which all lie in the domain of contrast taken broadly. Work on *a* in Russian linguistics has mainly concentrated on these other functions of *a* (Kreidlin and Paducheva 1974a,b; Sannikov 1989; Fougeron 1990; Uryson 2002, among others), while the corrective function has usually been attributed to a fixed collocation *ne ... a / a ne* consisting of *a* and the negative particle *ne*, and was excluded from the general analyses of *a*. However, it is a common pattern across languages that the same marker is used for correction and for (one or other type of) contrast—the English *but* is another famous case—so a reduction of correction to a special case of contrast is an obvious thing to try. This is the goal of the present paper. We present an attempt to derive the properties of the corrective uses of *a* from the general characteristics of *a* as a contrastive marker, the semantics and pragmatics of negation, and the properties of the context of use. In doing so we will always keep an eye on the English *but* as another marker that combines corrective and contrastive uses. Although not all of the findings about Russian corrections can be generalised to the English case, many nevertheless can.

This paper is structured as follows. Section [2] takes a closer look at the cross-linguistic regularities in correction marking, particularly at the question which other functions from the contrast semantic space correction markers tend to have. Section [3] briefly recapitulates the theory of contrast from Jasinskaja and Zeevat (2008, 2009), while in section [4] that theory is applied to correction. Finally, section [5] presents the conclusions and discusses further questions raised by this study.

[2] See Kasimir (2006) for detailed discussion of the terminological issue.

[2] CORRECTION MARKING ACROSS LANGUAGES

Some languages do not mark correction at all, i.e. correction is expressed by simple juxtaposition of a negative and a positive sentence, which is also possible in English: *John didn't go to Paris. He went to Berlin.* Other languages have dedicated markers of correction, i.e. markers that unambiguously express correction and nothing else, such as the German *sondern*, the Spanish *sino*, etc. Yet other languages use the same marker for correction and some other functions. Among those languages, correction is frequently coupled with functions that can be characterised as contrastive in one or another sense. Russian and English clearly belong to this group. This section will first present the most important distinctions between various kinds of contrast. This will make it possible to adequately describe the similarities and differences between (the non-corrective uses of) the Russian *a* and the English *but*. Then the most relevant theoretical perspectives upon the emerging picture will be presented.

[2.1] *Non-corrective uses of correction markers*

ADVERSATIVE: The first group of uses includes at least two relevant subgroups. The first one covers the 'prototypical' instances of Lakoff's (1971) *denial of expectation*, i.e. cases where the second conjunct denies some normal consequence of the situation presented in the first conjunct, as in (4), where being short usually implies bad performance in basketball, but this expectation is denied. In English, this function is expressed by *but*, the same marker that is used for correction, while the Russian adversative marker is *no*, a different one from the correction marker *a*.

(4) John is short, but he is good at basketball.

The second subgroup includes the so-called *argumentative* uses of *but* and the Russian *no* (Anscombre and Ducrot 1977). The argumentative function is fulfilled where the conjuncts A and B present an argument and a counterargument for a claim C. E.g. in (5), the fact that the ring is beautiful normally implies that we should buy it, but the fact that it is expensive implies that we shouldn't.

(5) This ring is beautiful, but expensive.

There has been a lot of effort to reduce both types of use either to denial of expectation or to the argumentative function. The theory summarised in section [3] presupposes a reduction of the latter kind. In any case, the distinction is irrelevant for our present purposes, both subgroups together constitute one class of non-corrective uses that we will refer to as *adversative*.

CONTRASTIVE COMPARISON: This term taken from Blakemore (1987) will be used to describe the second group of cases, where the conjoined clauses are presented

in a parallel fashion, so as to highlight the similarities and differences between them. There is no restriction to two conjuncts here, there can be three and more, as in (6). Crucially, the conjuncts must differ in *two* (or more) constituents, e.g. the subject and the object of liking in (6), leading to a contrastive topic-focus structure: *Oleg*, *Roma* and *Vera* are the contrastive topics, *football*, *basketball* and *tennis* are the contrastive foci. Contrastive comparison in the present sense corresponds closely to what is known in Russian linguistics as the *sopostavitel'noe značenie* ('comparative meaning') of the conjunction *a* (Kreidlin and Paducheva 1974b). Thus this function is conveyed in Russian by the same marker as is used for correction, while English uses a simple additive marker *and*.

(6) Oleg ljubit futbol, Roma basketbol, *a* Vera tennis
Oleg likes football Roma basketball and Vera tennis
Oleg likes football, Roma likes basketball, and Vera likes tennis.

Examples very similar to (6) with *but* instead of *and*, e.g. *John is tall, but Bill is small*, also appear in the literature under labels such as *semantic opposition* (Lakoff 1971), or *formal contrast* (Asher and Lascarides 2003). These labels, as well as Blakemore's *contrastive comparison* were introduced originally to distinguish such uses of *but*, which are also characterised primarily by parallel presentation and contrasting of the conjuncts from the proper adversative uses illustrated above. Indeed it seems possible to use *but* in the function we have just defined when the number of conjoined clauses is exactly two (Foolen 1991). However, as will become clear presently, there is a subtle difference between those uses of *but* and *contrastive comparison* in our definition.

As a final terminological remark, it is not clear that the requirement of at least two points of difference between the conjuncts and the contrastive topic-focus structure plays any important role in the original definitions of *contrastive comparison* or *semantic opposition*. It does, however, in our definition, because this is the feature that licenses the use of *a* in Russian. If the conjuncts only differ along one dimension, as in *John did the dishes and went shopping*, where *did the dishes* and *went shopping* present distinct actions, but the actor is the same, a different conjunction is used in Russian—a simple additive marker *i* (see Jasinskaja and Zeevat 2008, for detailed illustration).

FOOLEN'S TESTS: The third relevant type of contrast, typically expressed by *but* in English, does not have any widely accepted label of its own and has rarely been distinguished as a special function, or use, or meaning of contrastive conjunctions. It is very similar to contrastive comparison in that the conjoined propositions also have to differ along two dimensions. However, along one of those dimensions the values should not just be different, but in some sense opposite, e.g. the antonyms in (7), the positive vs. negative polarity in (8).

(7) John is tall, but Bill is small.

(8) John likes football, but Bill doesn't.

The opposition can also be pragmatic in nature, as in (9) where one conjunct confirms and the other denies a contextually salient proposition: *John lives in Amsterdam* confirms Speaker A's expectation that John lives in Amsterdam, while *Peter lives in Rotterdam* denies the expectation that Peter lives in Amsterdam, too, so in this case it is a way of saying that Peter does *not* live in Amsterdam (opposite polarity). Note that in (10) no assumptions of Speaker A are denied, i.e. there is no motivation for polar interpretation of the conjuncts, therefore *and* is the preferred conjunction. The contextual tests in (9) and (10) were introduced by Foolen (1991) to argue that *but* in all its uses involves a denial of expectation, as in (9). Whether or not we want to subscribe to Foolen's reduction of *but* to denial of expectation, his tests do draw the crucial distinction between contrastive comparison (10) and the type of contrast in question (9), for which we will reserve the term *opposition*. In both cases the conjoined propositions differ along two dimensions at least. However, in oppositions the values along one of those dimensions have to be polar.

(9) A: John and Peter both live in Amsterdam, don't they?
B: No. John lives in Amsterdam, but ($^{??}$and) Peter lives in Rotterdam.

(10) A: John and Peter don't live in the same place, do they?
B: No. John lives in Amsterdam, and ($^{??}$but) Peter lives in Rotterdam.

Thus opposition in the present sense is expressed by *but* in English. It should be obvious that the "oppositeness" of the conjuncts implies that there can be only two, which is in accordance with *but*'s restriction to two conjuncts.

In contrast, Russian uses *a* in this function, the same marker as for contrastive comparison, and not the same as for denial of expectation. Apparently, the parallel presentation and the contrastive topic-focus structure turns out to be decisive for the choice of conjunction.

Finally, this section can be summarised as shown in Table 1 on the following page. Apart from correction, the Russian conjunction *a* marks contrastive comparison and opposition, while the English *but* marks opposition and adversative contrast. Thus both the Russian *a* and the English *but* are markers of contrast, but they mark different types of contrast.

[2.2] *Typological theories of correction*

Why is correction often marked in the same way as contrast? And why does Russian use a contrastive comparison marker for correction, while English uses an

	CONTR. COMPARISON	OPPOSITION	ADVERSATIVE
Russian	*a*		*no*
English	*and*	*but*	

TABLE 1: Russian and English contrast markers

adversative? In this section we take a brief look at typological theories that bear on these questions.

A well-established approach to describing multifunctionality patterns of grammatical markers across languages is based on *conceptual*, or *semantic maps*. This approach has also been applied to correction and contrast marking; we will review two recent proposals in this framework: Malchukov (2004) and Mauri (2008). The notion of semantic map assumed in those studies is most closely related to Haspelmath's (2003) proposal. The approach is summarised below in a rather simplified form which might not reflect amply its philosophical motivation, but is consistent with the way it is applied by Malchukov (2004) and Mauri (2008).

A semantic map is a contiguous graph, whose nodes represent the possible functions of grammatical markers (such as CONTRASTIVE COMPARISON, OPPOSITION, ADVERSATIVE from the previous section), and whose arcs connect "most closely related" or "most similar" functions. The standard assumption is that both the set of possible functions and this "closeness" or "similarity" relation are universal. The relation is the basis for predictions concerning which marker-function mappings are possible in natural languages. In its strong form, the claim is that the set of functions expressed by the same marker must be a contiguous subgraph of the semantic map. The arcs also have a diachronic interpretation: a marker can only acquire a new function that is immediately connected to one it already has, and cannot "jump" over functions in between. This development can occasionally create exceptions to the contiguity claim in its strong form: if marker A acquires a new function formerly covered by marker B, it can split B's subgraph into two unconnected parts.[3]

The set of functions of a semantic map should be fine-grained enough to represent relevant differences in the usage of markers within a single language and across languages. If the meanings of two markers (in two languages) are equivalent, they are mapped to the same set of nodes; if the meanings are different, the sets of nodes must be different, too. Thinking of CONTRASTIVE COMPARISON, OPPOSITION and ADVERSATIVE as nodes of a semantic map, it becomes clear that having OPPOSITION separate from both other nodes is important to express the difference between the Russian and English contrastive conjunction systems, cf. Table 1. In cases where more than one function is expressed by the same marker,

[3] The theoretical status of such exceptions is, however, a matter of debate.

the approach is neutral with respect to the question whether those functions constitute different *senses* of that marker (a polysemous, or homonymous marker), or whether those functions are just different *uses* of a marker with a single abstract meaning.

The claim that there is a universal semantic map goes hand in hand with the assumption that the set of functions and the connections between them are somehow cognitively motivated, i.e. there are some fundamental characteristics of human thinking, or language processing, or communication, that determine which functions are likely to be expressed in natural languages and which of them are more likely to be expressed in the same way. However, semantic maps as such only represent claims about the existence of functions and relationships between them, but not about their nature. In some cases the nature of the relationship is well understood. In other cases it is less clear, so the semantic map is just the result of induction from polysemy patterns of markers from a representative sample of languages.[4]

Let's now consider the place of correction in relation to contrast in the semantic maps proposed by Malchukov (2004) and Mauri (2008), shown in figures 1 and 2 on the next page. Malchukov's function ADVERSATIVE is the same in all relevant respects as our notion of the adversative function. The function CONTRASTIVE, however, corresponds roughly to Lakoff's (1971) semantic opposition, and thus conflates our present notions of contrastive comparison and opposition. Mauri's OPPOSITION, in turn, corresponds closely to our contrastive comparison (not to our opposition!), while opposition in our sense and the adversative function are conflated under the label COUNTEREXPECTATIVE.[5] Thus, Malchukov's claim is that whenever a contrast marker is used for correction it should be the same marker as is used to connect sentences with contrastive topic-focus structure, no matter whether the conjuncts are 'opposite' or just distinct along two dimensions (both being part of the CONTRASTIVE function). Mauri's map amounts to (almost) the same claim. Although contrastive comparison is separated from opposition in our sense (the latter being part of COUNTEREXPECTATIVE), CORRECTION is placed between them, and thus can share markers with either of them.

Although both maps are consistent with the Russian and English correction marking patterns, i.e. they do not create non-contiguous marking regions, cf. tables 2 and 3 on the following page, they leave space for improvement and some open questions. First, both maps do not cleanly delineate the functions of different contrast markers *within* Russian and English systems. Since Malchukov lumps together contrastive comparison and opposition, the subtle difference in the us-

[4] See Janda (2009) for critical discussion of the universality claims of the semantic maps approach.

[5] The '...' node in both figures stands for a set of functions including plain additive and temporal conjunction, i.e. functions covered by the non-contrastive uses of *and* in English and the conjunction *i* in Russian.

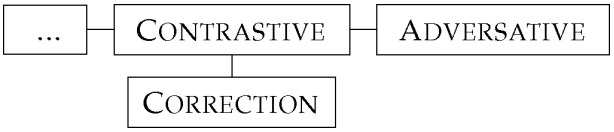


FIGURE 1: Correction in Malchukov's (2004) semantic map

FIGURE 2: Correction in Mauri's (2008) semantic map

<table>
<tr><th></th><th>CORRECTION</th><th>CONTRASTIVE</th><th>ADVERSATIVE</th></tr>
<tr><td>Russian</td><td colspan="2">a</td><td>no</td></tr>
<tr><td rowspan="2">English</td><td colspan="3">but</td></tr>
<tr><td></td><td>and</td><td></td></tr>
</table>

TABLE 2: Russian and English marking patterns in Malchukov's map

<table>
<tr><th></th><th>OPPOSITION</th><th>CORRECTION</th><th>COUNTEREXPECTATIVE</th></tr>
<tr><td>Russian</td><td colspan="2">a</td><td>no</td></tr>
<tr><td>English</td><td>and</td><td colspan="2">but</td></tr>
</table>

TABLE 3: Russian and English marking patterns in Mauri's map

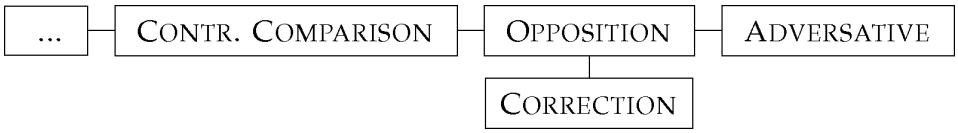


FIGURE 3: Correction and opposition in a semantic map

age of the English *and* and *but* observed by Foolen (1991) is not reflected by the map. Mauri's map, in turn, creates the wrong impression that the only difference between the Russian and the English systems is "on whose side" CORRECTION is. This is because Mauri follows Foolen in regarding opposition as a special case of denial of expectation and does not separate it from her COUNTEREXPECTATIVE function. However, Foolen's reductionist approach, which might be useful in finding a single abstract meaning for the English *but*, is not very helpful in constructing a semantic map. The distinctions that it blurs might be indeed spurious in English, but they are real in other languages, e.g. in Russian. As was pointed out above, another difference between the English and the Russian systems is in marking opposition, cf. table 1: in Russian both correction and opposition are coupled with contrastive comparison in *a*, whereas in English they are both coupled with the adversative in *but*. One might formulate a stronger hypothesis based on these observations, namely that CORRECTION is only related to OPPOSITION in our sense. A semantic map that suggests itself is shown in figure 3.[6] This map represents our (preliminary) answer to the question why Russian uses a contrastive comparison marker for correction, while English uses an adversative. Whenever a contrast marker is recruited for correction, it should be an OPPOSITION marker. Since in Russian OPPOSITION is coupled with CONTRASTIVE COMPARISON in *a*, the same marker is used for CORRECTION. Since in English OPPOSITION is coupled with the ADVERSATIVE function in *but*, CORRECTION is also expressed by *but*.

The second problem is not with the semantic maps as such, but with their motivation. What is the nature of the relationship between different contrast types? What makes CORRECTION and contrast, especially the OPPOSITION type of contrast so closely related? This is the central question to be addressed in this paper. To make this relationship explicit we will make use of the analytic tools of formal semantics. Only if CORRECTION and OPPOSITION can be represented as special cases of a single non-trivial category, or 'meaning', and only if the realisation of one or the other function can be predicted from context, can we talk about corrective uses of a general contrastive *a*, rather than a special corrective 'meaning' of *a*. In

[6] As will be shown in section [5] this semantic map is falsified once we consider a broader selection of languages, but it is consistent with the Russian and English data, so we will stick to it for the time being.

order to answer this question, the next section presents a theory of contrast from our previous work, and section [4] integrates correction into that theory.

[3] A THEORY OF CONTRAST

The central idea of Jasinskaja and Zeevat (2008, 2009) is that additive and contrastive markers like *and* and *but* convey information about the discourse topics addressed by the clauses they connect, where discourse topics are represented as questions under discussion (QUD) along the lines of e.g. Roberts (1996) and Büring (2003). Questions are represented as sets of Hamblin-style alternatives (Hamblin 1973), e.g. the question *Who smokes?* corresponds to the set of mutually compatible possible answers {*John smokes, Mary smokes, Bill smokes, ...*}. Contrast markers can indicate the *type of question* that their conjuncts answer. The question types relevant for the description of the English and Russian conjunction systems differ according to two main parameters: the number and the type of question variables. In terms of the number of variables, the most important distinction is between single and multiple variable questions, which corresponds to the number of dimensions in which the question alternatives differ. The canonical cases are single (*Who snores?*) vs. multiple *wh*-questions, e.g. *Who likes what?*, *Who gave what to whom?*, etc., respectively. In the most general form, the x notation is used to refer to a single variable, $\vec{x}$ for an unspecified number of variables (a tuple of one or more), and $\langle \vec{x}, y \rangle$ for multiple variables (a tuple of two or more). The most important variable types are, informally, *wh* for various types of entities that can answer questions like *who*, *what*, *when*, etc., and the *y/n* type for negative vs. positive polarity instantiated by negation and an identity operator of the same logical type. This is the variable type of *yes/no*-questions like *Does John like football?* and corresponds to the word *whether* in embedded questions. Abstracting away from the meanings of specific markers, let's apply this idea to the definition of the different types of contrast—CONTRASTIVE COMPARISON, OPPOSITION and ADVERSATIVE—which make up the semantic map proposed in the previous section in figure 3 on the previous page.

CONTRASTIVE COMPARISON: Two or more clauses stand in a relation of CONTRASTIVE COMPARISON to one another if (a) they address a discourse topic that can be represented as a double or multiple *wh*-question, i.e. a $\langle \vec{x}_{wh}, y_{wh} \rangle$-question, and (b) they give *distinct* answers to such a question so that the instantiations of each variable in the question are distinct. For example, in (6) repeated below, the QUD can be assumed to be *Who likes what kind of sports?* with two variables *who* and *what kind of sports*. *Oleg*, *Roma* and *Vera* are mutually distinct instantiations of the *who*-variable, while *football*, *basketball*, and *tennis* instantiate the *what kind of sports*-variable and are also mutually distinct:

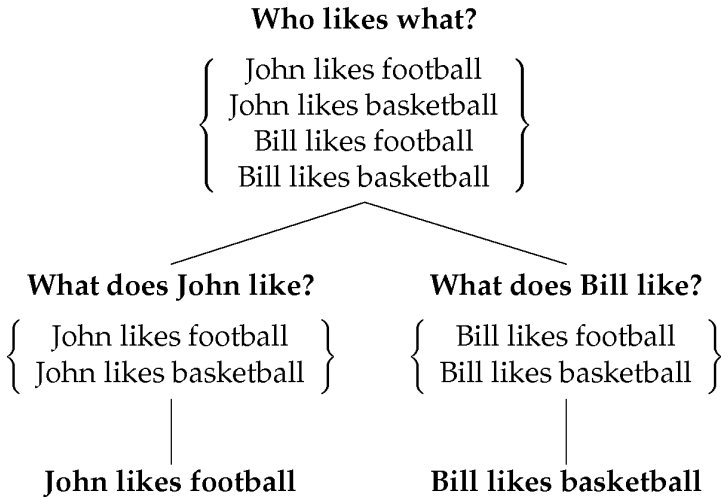


FIGURE 4: A $\langle \vec{x}_{wh}, y_{wh} \rangle$-question

(11) Oleg ljubit futbol, Roma basketbol, *a* Vera tennis
Oleg likes football Roma basketball and Vera tennis
Oleg likes football, Roma likes basketball, and Vera likes tennis.

The alternative set of a similar double *wh*-question is shown in cf. figure 4. Notice that this set can be partitioned into subsets that correspond to single variable subquestions *What does John like?*, *What does Bill like?* This is what we will refer to as *splitting up* a question into subquestions, or a *strategy* in Roberts' (1996) and Büring's (2003) terminology.

OPPOSITION: The OPPOSITION relation also involves giving distinct answers to a double or multiple variable question, however one of those variables has to be of the *yes/no* type: $\langle \vec{x}_{wh}, y_{y/n} \rangle$. We will also refer to this type as *wh-yes/no*-questions.[7] The alternative set of such a question is shown in figure 5 on the following page: the alternatives differ as to *who* likes or doesn't like football, and in the presence vs. absence of negation. Neither English, nor Russian (nor any other natural language we are aware of) can express this type of question by a simple interrogative sentence. A possible gloss one could give to the set of alternatives in figure 5 is *Who "whether" likes football?* In English, one can express this question either by conjoining a number of *y/n*-questions, as in figure 5, or by conjoining two *wh*-questions *Who does and who doesn't like football?* For the rest, the analogy between

[7] This is a necessary but not a sufficient condition for OPPOSITION. As will be shown below, also ADVERSATIVES and CORRECTIONS address *wh-yes/no*-questions, but have to satisfy a number of additional conditions. To keep the functions in a semantic map mutually exclusive we will ultimately define OPPOSITION as a relation between answers to a *wh-yes/no*-question that is *not* adversative or corrective.

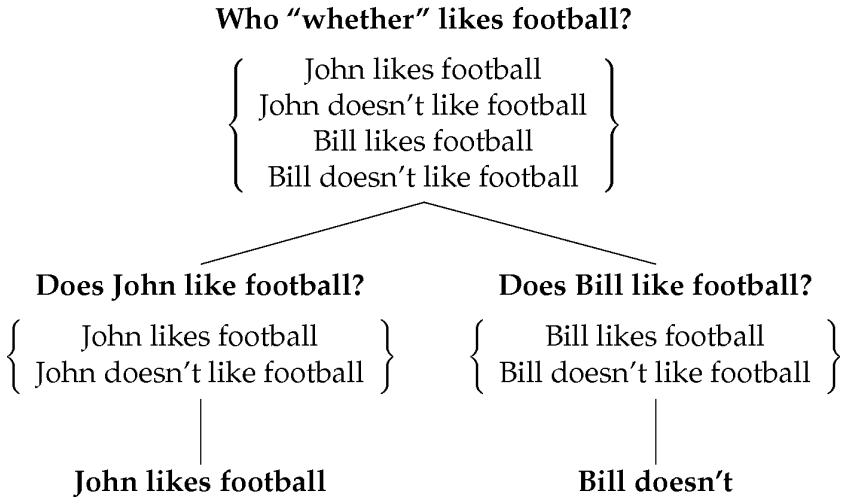


FIGURE 5: A $\langle \vec{x}_{wh}, y_{y/n} \rangle$-question

wh-yes/no-questions and standard multiple *wh*-questions is obvious, cf. figures 4 and 5.[8]

The *y/n* variable in the topic question is what accounts for the "oppositeness" of the clauses that stand in an OPPOSITION relation. It predicts that either one clause should be positive and the other negative as in (8) or in figure 5, or that one clause is used to deny an implicit supposition of the right form. E.g. in (9),

[8] Although the present representation of the discourse topic is inspired by Büring (2003), unlike Büring and more in line with Hamblin (1973) we assume that the alternative set of a *yes/no*-question contains both a positive and a negative alternative. Assuming that the alternative set of a question is partitioned by the alternative sets of its subquestions, this gives us that double *wh-yes/no*-questions also contain both positive and negative alternatives, cf. figure 5. If they only contained positive alternatives, then the *wh-yes/no*-question *Who "whether" likes football?* would be indistinguishable from the single *wh*-question *Who likes football?* There are various semantic reasons for keeping *yes/no*-questions to just the positive alternative and one might even argue that double *wh-whether* interrogative sentences do not exist precisely because the alternative set containing both positive and negative alternatives is not a legitimate semantic object, while its positive subset is indistinguishable from the single *wh*-question. However, *wh-yes/no*-questions as pragmatic objects, i.e. as issues to be interested in, certainly do exist and are distinct from single *wh*-questions. In the first case, both the positive and the negative extension of the question predicate P (for the question *who "whether"* P*?*, e.g. $\lambda x[x$ likes football$]$ in the present example) must be explicitly named. If some object is not named one may conclude that it is not relevant, but not that it is not P. In the second case, only the positive extension is asked for, while for the remaining relevant objects 'not P' is inferred by the process of exhaustive interpretation (e.g. Schulz and van Rooij 2006). Thus, including the negative alternatives gives us a representational handle on this pragmatic distinction (even if it does not *per se* explain it).

repeated below, *Peter lives in Rotterdam* is a way of saying that Peter does not live in Amsterdam, and thus an answer to the question *Does Peter live in Amsterdam?*[9]

(12) A: John and Peter both live in Amsterdam, don't they?
B: No. John lives in Amsterdam, but ($^{??}$and) Peter lives in Rotterdam.

Similarly, the *y/n* variable is responsible for the restriction to exactly two clauses in an OPPOSITION relation, because there can only be two distinct answers to a *yes/no*-question—*yes* and *no*.

ADVERSATIVE: Finally, one of the necessary conditions for an ADVERSATIVE relation is that the clauses give distinct answers to a particular kind of $\langle x_{wh}, y_{y/n} \rangle$-question, which we called *why-whether-* or *why-yes/no*-questions in Jasinskaja and Zeevat (2008, 2009). For example, *Why "whether" we should buy this ring? — [Why should we buy this ring?] It is beautiful, but [why shouldn't we buy this ring?] it is expensive.* Further necessary conditions concern the kind of causal relation involved and the place of the *wh* and the *y/n* variable in the structure of the alternatives. As is made explicit in (13), the alternatives all involve a causal relation between two statements expressed by a two-place operator BECAUSE, whose first argument is the cause, and whose second argument is the effect. In fact, this is not just a general BECAUSE, but its argumentative variety, i.e. an epistemic or a speech act BECAUSE in Sweetser's terminology (1990), so it would be more adequate to say that its second argument is a claim or suggestion, while its first argument gives support to that claim, i.e. the reason to think that it is true or the reason to accept the proposition.

(13) $$\left\{ \begin{array}{l} \textsc{because}(\text{this ring is beautiful}, \text{we buy it}) \\ \textsc{because}(\text{this ring is expensive}, \text{we buy it}) \\ \textsc{because}(\text{this ring is beautiful}, \textsc{neg}(\text{we buy it})) \\ \textsc{because}(\text{this ring is expensive}, \textsc{neg}(\text{we buy it})) \end{array} \right\}$$

The relevant question type must satisfy two more conditions. First, the *wh* variable of the question is the reason, i.e. first argument of BECAUSE. Second, the *y/n*

[9] For our purposes it does not matter that the question in (12)A is biased. After an unbiased question *Does John live in Amsterdam, and does Peter live in Amsterdam, too?* (cf. also (11) on p. 215 in Umbach 2005) we predict the same result, i.e. that *but* will be used. The question explicitly asked by the interlocutor is not always precisely the same as the one taken up by the speaker who gives an answer (though of course they must be closely related, otherwise communication would not be possible). Our claim is that regardless of our assumptions about alternative sets for biased questions, the question answered by speaker B is *Who "whether" lives in Amsterdam?* with an alternative set of the form { *John lives in Amsterdam*, *John doesn't live in Amsterdam*, *Peter lives in Amsterdam*, *Peter doesn't live in Amsterdam* }.

variable is the polarity of the consequent.[10] This specific subtype of a *wh-yes/no*-question defines the ADVERSATIVE relation, or the ADVERSATIVE function in the semantic map in figure 3 on page 441. To keep the nodes of the semantic map disjoint one should assume that OPPOSITION involves all kinds of *wh-yes/no*-questions except this subtype of *why-yes/no*-questions, so when we talk about *wh-yes/no*-questions in the rest of the paper we will usually mean them in this narrow sense, to the exclusion of questions like (13).

As was already mentioned, semantic maps do not specify whether a certain multi-functional marker has a single abstract meaning, or as many different meanings as many functions of the semantic map it covers (or something in between). The representation of the different types of contrast in terms of the type of question under discussion can in fact be used in both ways. For example, the semantics of the English *but* can specify two options for the discourse topic: the *wh-yes/no*-questions in the narrow sense characteristic of OPPOSITION and *why-yes/no*-questions characteristic of the ADVERSATIVE, cf. table 1 on page 438 (multiple meanings). It can also be defined in terms of *wh-yes/no*-questions in the broad sense which covers both OPPOSITION and ADVERSATIVE (a single meaning). The latter approach is developed in Jasinskaja and Zeevat (2008, 2009). The English *but* is just a marker of *wh-yes/no*-topics in the broad sense and the Russian *no* marks *why-yes/no*-topics. The meanings of other markers can be defined in negative terms: e.g. the English *and* receives an abstract meaning as a marker of distinct answers to an unspecified type of question, but since it stands in a kind of paradigmatic relationship to *but*, the topic types for which *but* is more appropriate are excluded

[10] A few remarks are in order here. First, it still needs to be investigated whether negation in the negative alternatives needs to take scope over the whole consequent of BECAUSE or can have narrower scope within it. In any case, however, the consequents of positive and negative alternatives must be mutually exclusive.
Second, BECAUSE expresses a veridical relation, i.e. $\textsc{because}(P, Q)$ entails both P and Q. This means that distinct answers to a *why-yes/no*-question are always mutually exclusive: $\textsc{because}(P_1, Q) \wedge \textsc{because}(P_2, \neg Q)$ entails both Q and $\neg Q$. This is why adversative conjunctions like *but* and *no* always mark one of their conjuncts (usually the second one) as decisive:

(i) a. The ring is expensive, but it is beautiful. (We will buy it)
b. The ring is beautiful, but it is expensive. (We will not buy it)

This is also why BECAUSE in adversatives is the argumentative BECAUSE. Adversatives are used when the issue whether Q is not settled and is a matter of actual or possible dispute. The consequent of P_1 is a concession to the contrary view, while the consequent of P_2 is the proposition the speaker really endorses.
Third, one can think of other possible alternative sets that involve a causal relation, a *wh*-type variable and a *y/n*-type variable, but if they do not meet the specific conditions stated above, then they do not give rise to an ADVERSATIVE relation. An interesting case are corrections of causes, such as:

(ii) John hit Peter not because he was angry, but because he was drunk.

Here what is negated in the negative and asserted in the positive alternatives is not the consequent, but the identity of the cause. In contrast, the consequent *John hit Peter* holds in all the alternatives.

from its marking domain (the *blocking* mechanism in Jasinskaja and Zeevat 2009). As a result, *and* admits all topic types except *wh-yes/no*-topics. It is sometimes difficult to decide which marker in a system should receive a positive definition, and which an abstract function restricted by blocking. The historical development of the system can give an effective clue: a relatively young marker that is expanding its set of functions should be defined positively, whereas an older marker that loses its functions to a newcomer is blocked by it.[11]

For the sake of readability, less technical terminology will be used in the rest of the paper. We will refer to $\langle \vec{x}_{wh}, y_{y/n} \rangle$-questions as *wh-y/n* and use the term 'double *wh*' for double variable questions that do not have a *y/n*-type variable. Double (variable) questions are thus a supertype of double *wh* and double *wh-y/n*. We will mainly talk about double questions assuming that the extension to multiple questions in general is trivial.

[4] CORRECTION AS A TYPE OF CONTRAST

This section will present an argument for the claim that both OPPOSITION (14) and CORRECTION (15) are realisations of a *wh-yes/no* strategy. At first glance these realisations look very different: (14) shows a contrastive topic-focus pattern, with a *wh*-type topic and polarity focus.[12] In contrast, (15) has focal accent on the instantiations of the *wh* variable, while a contrastive topic seems to be missing altogether.

(14) a. Oleg KURIT, a Roma ne KURIT.
Oleg smokes but Roma not smokes
Oleg smokes, but Roma doesn't.
b. Oleg ne KURIT, a Roma KURIT.
Oleg not smokes but Roma smokes
Oleg doesn't smoke, but Roma does.

(15) a. Kurit OLEG, a ne ROMA.
smokes Oleg but not Roma
b. Kurit ne Oleg a ROMA.
smokes not Oleg but Roma

Moreover, Russian corrections obligatorily contain what is traditionally called *constituent negation* (in contrast to *sentential negation*, see Babby 1980, 2001; Brown

[11] A system of markers can also undergo reorganisation though, which can lead to an older marker acquiring a new positive meaning.

[12] Polarity focus both positive and negative is realised in Russian by a focal accent on the finite verb. The negative particle *ne* is a clitic, so it normally remains unaccented and does not function as a negative polarity focus exponent. In contrast, the finite verb is accented in both conjuncts in (14) even though the lexical verb itself is given at least in the second conjunct (normally, given material is deaccented). Thus the morphosyntactic constraints on focal accent placement overrule the considerations of givenness.

1999), i.e. the negative particle *ne* appears immediately before the constituent to be corrected, cf. *ne Roma, ne Oleg* 'not Roma', 'not Oleg' in (15). The standard assumption is that sentences with constituent negation of the form *not* X P presuppose that some object has property P (Borschev et al. 2006), i.e. their meaning is similar to that of the English negated clefts: *It is not John who smokes*. In contrast, *sentential negation* is expressed by the negative particle appearing immediately before the finite verb, e.g. *ne kurit*, lit. 'not smokes' in (14). Sentential negation is possible in opposition sentences, but it cannot introduce the negative conjunct in corrections.

The goal of this section is to show, on the one hand, that all these structural differences fall within the range of options in addressing a *wh-yes/no* discourse topic, and on the other hand, that they correlate with precisely those functional features that make out the difference between the OPPOSITION and the CORRECTION function. We will start with an overview of logical possibilities in how a *wh-yes/no* topic can be addressed in section [4.1]. Section [4.2] singles out one subtype of opposition sentences which bears the closest resemblance to correction in terms of those logical possibilities. The functional differences between the members of such minimal pairs are formulated. The last three sections relate those functional differences to sentential vs. constituent negation (section [4.3]) and differences in information structure (sections [4.4] and [4.5]).

[4.1] *Topic and focus in* wh-yes/no

There are always two ways to address a double question like *Who ate what?* You can go by people, or you can go by food. In the first case, the double question *Who ate what?* is split up into a series of single variable questions like *What did John eat?*, *What did Bill eat?*, etc., where the *who*-variable is instantiated by different persons from the relevant domain. In the second case, the double question is split up into subquestions *Who ate the beans?*, *Who ate the carrots?*, etc. According to Büring (2003), the choice between these two strategies determines which constituent is marked as contrastive topic and which one as focus: contrastive topic is the variable that is instantiated in the subquestion, i.e. people when you go by people, and food when you go by food; the focused constituent corresponds to the *wh*-variable in the subquestion.

Applying the same idea to *wh-yes/no*-questions we also get two possible strategies. Suppose the question is *where "whether" John went*. If we go by the locations instantiating the *where* variable, the question is split up into a series of *yes/no*-questions: *Did John go to Paris?*, *Did John go to Berlin?*, etc., as shown in figure 6 on the next page. In this case *to Paris*, *to Berlin*, etc., are contrastive topics ($[...]_T$), while the polarity is the focus ($[...]_F$), which surfaces as the focal accent on the auxiliary verb *did* or *didn't*. This is the structure underlying the canonical examples of OPPOSITION such as (8).

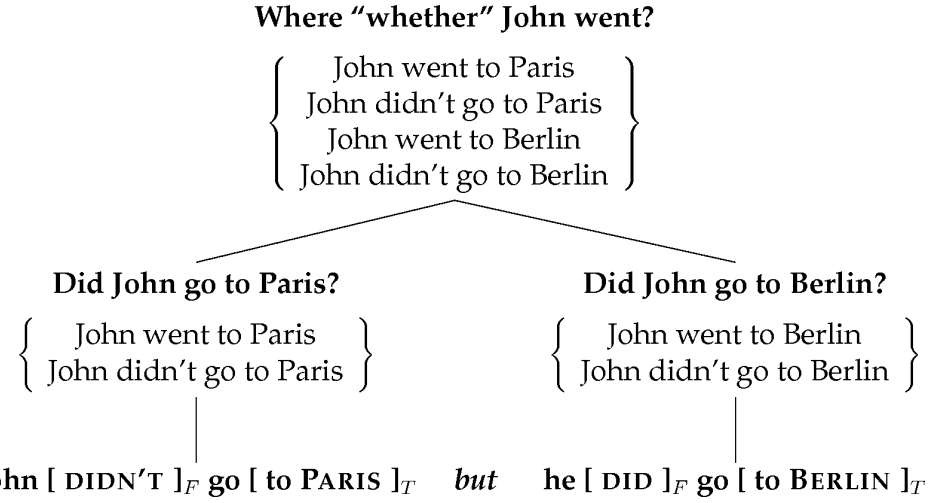


FIGURE 6: A *wh-yes/no* question split up by the *wh* variable

The other possibility is to instantiate the *y/n*-variable first, which splits up the *wh-yes/no*-question into two *wh*-questions, one addressing the positive part of the question and the other addressing the negative part, e.g. *Where did John go?*, *Where didn't John go?*, cf. figure 7 on the following page. In this case, the polarity would be marked as contrastive topic, and the answers to the *where*-question as focus.[13]

[4.2] *Corrections vs. oppositions with* y/n*-topics*

The main claim we would like to put forward is that corrections (16-b)/(17-b) have the same underlying QUD structure as oppositions with *y/n*-topics (16-a)/(17-a), i.e. they both address an overarching *wh-yes/no*-question, which is split up by polarity as in figure 7. The assertive propositional content of the conjuncts in both cases is the same: one conjunct states that it is not the case that John went to Paris, so it provides an answer to the question where John did not go; the other conjunct states that he went to Berlin, which is an answer to where John went.[14]

[13] Apparently, in English contrastive topics and foci can be marked just by intonation: topics receive a type B and foci a type A pitch accent (Büring 2003), which includes topics *in situ* that linearly follow the focus, as in figure 6. In German, there is a constraint that a topic must be followed by at least one focus in the same sentence (Büring 1997). In a sentence like that in figure 6 this can be achieved by topic fronting: *[Nach Paris]$_T$ ist er [nicht]$_F$ gefahren, aber [nach Berlin]$_T$ [schon]$_F$*. Russian is more like German in this respect: accented contrastive topics have to precede foci; the melodic form of the pitch accent in turn is a less reliable cue to the topic/focus distinction than word order. There is a lot of variation in the form of the topic and focus accents (see Mehlhorn and Zybatow 2000, for a convincing illustration), and one and the same accent can mark both topic and focus depending on the context (Kodzasov 1996, p. 198).

[14] Since corrections have no contrastive topics, this contradicts Büring's (2003) claim that the presence of a strategy—a double question split up into single variable questions—is a sufficient condition for contrastive topic marking. The proposal developed in section [4.4] assumes that exceptions from this claim are possible when the pair or sequence of clauses realise two different strategies at the same time and the other strategy requires a structure without a contrastive topic.

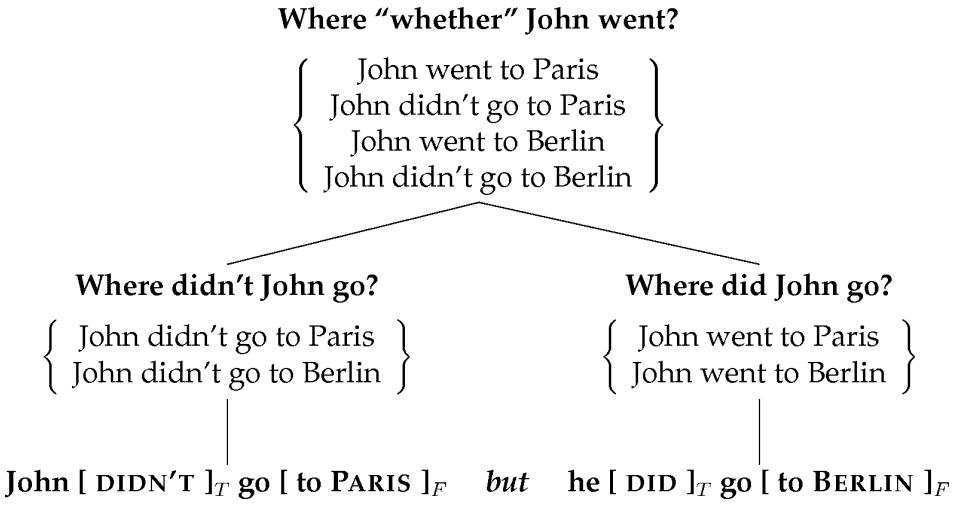


FIGURE 7: A *wh-yes/no* question split up by the *y/n* variable

(16) a. John [DIDN'T]$_T$ go [to PARIS]$_F$, but he [DID]$_T$ go [to BERLIN]$_F$.
b. John didn't go [to PARIS]$_F$, but [to BERLIN]$_F$.

(17) a. Oleg [ne EZDIL]$_T$ [v PARIŽ,]$_F$ a [EZDIL]$_T$ [v BERLIN.]$_F$
Oleg not went to Paris but went to Berlin
b. Oleg ezdil ne [v PARIŽ,]$_F$ a [v BERLIN.]$_F$
Oleg went not to Paris but to Berlin

The functional differences between the two versions (a) and (b) lie in the domain of presuppositions and/or implicatures. The Russian sentence (17-a) is rather marked, presumably because it can only be felicitously used in a context where *going* and *not going* to different places has been at issue. It seems to presuppose that there is a place that Oleg did not go to, and another place that he did go to, and specifies the first one to be Paris and the second to be Berlin. Its English counterpart (16-a) might sound less marked, but with really heavy contrastive topic accentuation on the auxiliaries it seems to have similar presuppositions.

In contrast, (16-b)/(17-b) only presupposes that John/Oleg went somewhere. The first conjunct negates that on a particular occasion John went to Paris, while the second conjunct states that on *that* occasion, in *that* event of going to a place, John went to Berlin rather than Paris. In other words, the wrong element *Paris* is *replaced* by the correct element *Berlin* in the description of a particular '*John went to* X'-event. We will refer to this property as *replacivity*, which is the most important distinctive feature of corrections among other kinds of contrast.[15] Notice that in the (a) versions going to Paris and going to Berlin are treated as distinct

[15] The term is derived from Jacobs' *replacive negation*, i.e. a type of negation that requires a correction according to Jacobs (1982, 1991). The same property has also been referred to as 'denial by substitution' by Umbach (2004).

possibilities, while in the corrections there is only one relevant occasion of going somewhere and it can either be to Paris, or to Berlin.

We have been using the term 'presuppose' in a rather non-technical sense here. In the following two sections we will make more precise assumptions about the nature of the 'presuppositions' involved and the linguistic means that contribute those presuppositions. Our discussion will concern primarily the Russian examples and will only touch upon a possible generalisation to English, which cannot be developed in detail in the present paper.

[4.3] *Negation and its presuppositions*

Our first assumption concerning negation will be that it 'presupposes' in a certain weak sense the proposition it negates. This is not the traditional, strong notion of presupposition which requires the presupposed material to be entailed by the context. It is enough that that material is somehow suggested, a possibility that could be entertained by someone on the basis of the current information state. Horn (1989) calls it 'supposition', others have used the term 'weak presupposition' (Zeevat 2008). It is a general characteristics of the pragmatics of overt negation that reflects the fact that one would never say that *John didn't go to Paris* unless it were somehow possible that John would go to Paris. This is equally true for English and Russian negation.

Of particular interest to us is the distinction between what is traditionally called *sentential* and *constituent* negation. Although we will stick to traditional terminology, one should keep in mind that it is rather misleading. It suggests that sentential negation takes scope over the whole sentence, while constituent negation takes narrower scope, but as was convincingly shown by Jacobs (1982) this is not at all the relevant distinction. From a syntactic point of view, sentential negation is verbal negation, i.e. the negative particle *ne* appears immediately before the finite verb and takes scope over the VP. It has received a lot of attention in the literature on Russian especially because it licenses the genitive of negation, as well as negative polarity (negative concord) items (Babby 1980, 2001; Brown 1999; Borschev et al. 2006). From a semantic point of view, its assertive content is just logical negation. For convenience, we will assume *ne* to denote $\lambda P \lambda Q[Q(\lambda x \neg P(x))]$ where P is a property that stands for the meaning of the VP, and Q a quantifier denoted by the argument (typically, the subject) that still needs to be supplied to make it a full proposition.[16] Accordingly, the weak presupposition it introduces is simply $Q(P)$. For example in (18), Q is $\lambda P[P(Oleg)]$ and P is $smoke$ which gives us $\neg smoke(Oleg)$ for the assertive meaning of the sentence, and $smoke(Oleg)$ for its weak presupposition.

[16] It is immaterial for the present discussion whether the given logical type is basic for the Russian negative particle, or the result of syntactically or semantically motivated abstraction operations on a lower basic type.

(18) Oleg ne kurit
Oleg not smokes
Oleg doesn't smoke.

In contrast, constituent negation is marked by the particle *ne* appearing in front of "the constituent that is negated," cf. (19), which can be (almost) any constituent: quantificational and referential DPs, PPs, etc., and in particular also VPs or whole sentences. Thus from a syntactic point of view, constituent negation is cross-categorial negation (at least superficially).[17] Normally, the negated constituent receives focal accent.

(19) a. ne [OLEG]$_F$ kurit
not Oleg smokes
b. kurit ne [OLEG]$_F$
smokes not Oleg
It is not Oleg that smokes.

Semantically, "the constituent that is negated" does not just mean that negation takes scope over that constituent in the standard sense. Sentences with constituent negation have altogether rather different semantics from the sententially negated ones. Constituent negation is typically assumed to presuppose the positive part of the sentence, e.g. (19) presupposes that someone smokes (Borschev et al. 2006). In fact, a stronger assumption seems justified: Russian sentences with constituent negation have roughly the same semantics as e.g. the English negated specificational (pseudo)cleft sentences, i.e. *It is not Oleg that smokes*, or *Who smokes is not Oleg*.

The first approximation of how this meaning is composed is shown in figure 8 on the next page. Negation applies to the property of being Oleg ($\lambda x[x = Oleg]$) associated with the negated DP, and takes the quantifier *who smokes* ($\lambda P \forall x[[x \in C \wedge smoke(x)] \rightarrow P(x)]$) associated with the fronted verb *kurit* as its second argument. Simplifying again, the positive part of the sentence *kurit* 'who smokes' is represented as a universal quantifier.[18] Its domain restriction C depends on the context of utterance and realises the idea that only relevant individuals that smoke are concerned. Notice that the same semantics is assigned to constituent

[17] This means that when the negative particle appears immediately before the finite verb the sentence is ambiguous between a constituent and a sentential negation reading.

[18] It is more common to treat free relatives, which participate in pseudocleft constructions, as definites, or *maximal* individuals (Jacobson 1995; Rullmann 1995). Notions like maximality, however, implicitly involve universal quantification.

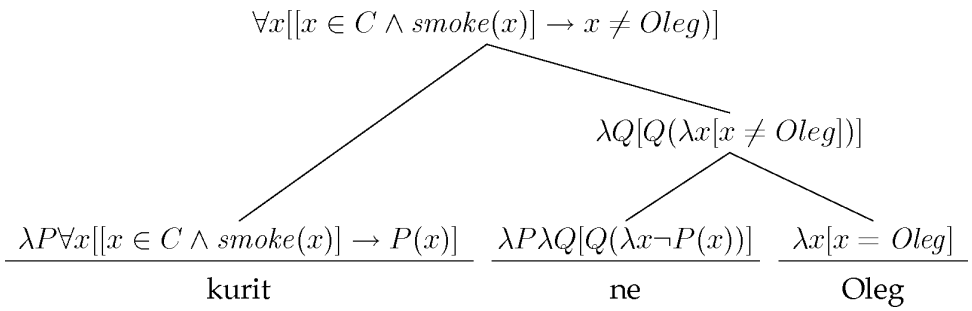


FIGURE 8: Semantic composition for a sentence with constituent negation, (19-b)

and sentential negation,[19] all the difference comes from the meanings associated with the rest of the sentence—the negated and the positive parts. We assume that these differences are accounted for by whatever syntactic operations are responsible for the marked word order and accentuation, and especially for the position of the negative particle in sentences with constituent negation. However, no details of the syntactic analysis will be discussed.[20]

As far as presupposition is concerned, first of all, the fronted verb *kurit*, just like a free relative *who smokes* (and Fregean definites), introduces an existential presupposition that someone relevant ($x \in C$) smokes, i.e. $\exists x[x \in C \wedge smoke(x)]$. This accounts for the intuitions of Borschev et al. (2006). Second, negation weakly presupposes what it negates, i.e. in the present case it is the meaning of a positive (pseudo)cleft *It is Oleg that smokes* or *Who smokes is Oleg*: $\forall x[[x \in C \wedge smoke(x)] \rightarrow x = Oleg]$. Notice that in conjunction with the existential presupposition this is equivalent to saying that *Oleg* is an exhaustive answer to the question *Who smokes?*, which can also be roughly paraphrased as *only Oleg smokes*. In other words, a sentence with constituent negation like (19) presupposes that the question *Who smokes?* has previously been answered exhaustively, and the answer was *Oleg* (or this answer was expected or possible). Thus constituent negation brings in ex-

[19] This is partly due to the wide scope of the quantifier Q over negation, which in turn only takes scope over the predicate P in our definition. In other words, P represents the negated and Q the positive part of the sentence. This might not be general enough to account for all possible readings of sentences with sentential negation. Certainly, a more general and principled analysis can be provided in the future.

[20] If constituent negation applies to the VP, as in (i), the sentence is to be interpreted roughly as *what Oleg does isn't smoke*, i.e. the fronted DP *Oleg* represents in this case a quantifier over eventualities that involve Oleg as agent. Once again, this quantifier must result from the interpretation of syntactic substructures of the sentence which in turn result from presumably the same syntactic operations as those that apply in sentences like (19) with DP constituent negation. The difference is that here the surface word order happens to coincide with the unmarked word order with the negative particle before the finite verb, so the sentence *Oleg ne kurit* is ambiguous between the constituent and the sentential negation reading.

(i) Oleg ne kurit, a p'jot.
Oleg not smokes but drinks

haustivity into the semantics of the sentence. As we will see in the following section exhaustivity is one of the crucial ingredients of replacivity in corrections.

From a pragmatic point of view, the distinction between constituent and sentential negation in Russian is close to the replacive vs. non-replacive distinction introduced by Jacobs (1982, 1991). Replacive negation calls for a correction. The sentences in (19), for example, sound incomplete without a continuation stating who actually smokes. This incompleteness does not lead to ungrammaticality, but there is a clear sense that after (19) the question *Who smokes instead?* is somehow "in the air."

[4.4] *Exhaustivity and replacivity*

Coming back to corrections, let us first show that their replacivity property can be derived from the assumption that one conjunct negates an exhaustive answer to a question, while the other conjunct gives a new exhaustive answer to the same question. As was argued in the previous section, the negative conjunct in (20) presupposes that *Oleg* is an exhaustive answer to the question *Who smokes?*, which can be derived by applying, let's say, Groenendijk and Stokhof's (1984) exhaustivity operator EXH[21] to the quantifier that represents the meaning of the answer $\lambda P[P(\textit{Oleg})]$, and applying the resulting exhaustive quantifier to the predicate of the question $\lambda x[x \in C \wedge \textit{smoke}(x)]$, i.e. $\textsc{exh}(\lambda P[P(\textit{Oleg})])(\lambda x[x \in C \wedge \textit{smoke}(x)])$, which is equivalent to $\forall x[[x \in C \wedge \textit{smoke}(x)] \leftrightarrow x = \textit{Oleg}]$.

(20) Kurit ne OLEG, a ROMA.
smokes not Oleg but Roma
Not Oleg, but Roma smokes.

The proposition $\forall x[[x \in C \wedge \textit{smoke}(x)] \leftrightarrow x = \textit{Oleg}]$ is "removed" from the information state after the processing of the incompatible proposition $\forall x[[x \in C \wedge \textit{smoke}(x)] \rightarrow x \neq \textit{Oleg}]$ that constitutes the content of the negative conjunct.[22] For Roma to *replace* Oleg as the one relevant person who smokes, we have to assume that the positive conjunct provides another exhaustive answer to the question *Who smokes?*: $\forall x[[x \in C \wedge \textit{smoke}(x)] \leftrightarrow x = \textit{Roma}]$. Given that Oleg and Roma are two distinct persons, these two exhaustive answers are mutually incompatible.[23] Moreover, it is essential that the question with respect to which

[21] For reference, this is one of the existing formulations of Groenendijk and Stokhof's exhaustivity: $\textsc{exh} = \lambda\mathcal{Q}\lambda P[\mathcal{Q}(P) \wedge \neg\exists P'[\mathcal{Q}(P') \wedge P \neq P' \wedge \forall x[P'(x) \rightarrow P(x)]]]$.

[22] One would have to assume some kind of non-monotonic notion of update (such as e.g. Asher and Lascarides 2003) to make such "removal" of incompatible propositions possible. Otherwise we reach an absurd information state.

[23] The distinctness of Oleg and Roma would normally be part of general world knowledge, but it also follows from the requirement that the conjuncts of a contrastive conjunction provide distinct answers to the same question. In the present case, Oleg and Roma would be distinct instantiations of the *wh*-variable in a *wh-yes/no*-question.

the answers are exhaustivized is strictly the same, which also means the same instantiation for the implicit contextual restriction C. The question might, and normally would also contain other implicit restrictions that constrain, for instance, the time and location of the relevant smoking events. Keeping all those restrictions the same in the two instances of exhaustivization ensures that Roma comes *in place* of Oleg as the only relevant smoker on that particular relevant occasion.

The idea to derive replacivity and mutual exclusiveness of the conjuncts in corrections from the assumption that the conjuncts represent exhaustive answers to the same question has been previously developed by Kasimir (2006) in her account of the German *sondern*. Our proposal implements the same idea, except that if Kasimir makes exhaustivity of the conjuncts a presupposition conventionally associated with *sondern*, in our case it is not part of the semantics of the Russian *a*, but is contributed by constituent negation, which is obligatory in corrections.[24]

[24] There are some curious exceptions to the claim that corrections always involve constituent negation. In (i), both the negative existential predicate *net* (Borschev et al. 2006) and the negative concord item *ni odnogo* '(not) a single' indicate sentential negation. In (ii) it is the negative concord item *nikakix* 'no'. Nevertheless, both are followed by a correction with *a*.

(i) Na ètoj grjadke net ni odnogo ovošča a tol'ko sornjaki.
on this patch there isn't no single vegetable but only weed

(ii) Oleg ne ugonjal nikakix mašin, a igral ves' večer so mnoj v karty
Oleg not high-jacked no cars but played all evening with me in cards
Oleg didn't high-jack any cars, but was playing cards with me all evening.

Possibly, what is going on here is when the speaker utters the first clause she is not planning yet to produce a correction. The plan to make it into a correction appears after the production of the negative clause, so it is reinterpreted *post hoc* as one with a constituent negation: *There are no vegetables* ⇒ *What there is is not vegetables*; *Oleg didn't high-jack any cars* ⇒ *What Oleg was doing is not high-jacking cars*. This is supported by the fact that, although generally Russian corrections with *a* can be turned around—positive conjunct first, negative second, cf. (15-a) vs. (15-b)—this is not possible in these examples, cf. (iii-b) and (iv-b). This is because by the time the speaker reaches the negative conjunct, she is already committed to the plan of producing a correction, so constituent negation must be used. Notice that the versions without the negative concord items (iii-b) and (iv-b) are felicitous.

(iii) Na ètoj grjadke tol'ko sornjaki,
on this patch only weed
a. # a (net) ni odnogo ovošča.
but there isn't no single vegetable
b. a ne ovošči.
but not vegetables

(iv) Oleg igral ves' večer so mnoj v karty,
Oleg played all evening with me in cards
a. # a ne ugonjal nikakix mašin.
but not high-jacked no cars
b. a ne ugonjal mašiny.
but not high-jacked cars

More precisely, constituent negation provides the exhaustivity of the presupposed proposition to be corrected, while the exhaustivity of the positive conjunct still needs a short comment.

It is common to assume that answers to a question are interpreted exhaustively by default, even if they do not contain linguistic devices that encode exhaustivity, like *only* or cleft constructions. Presumably, default exhaustivization of this sort takes place in the positive conjunct in corrections. However, what guarantees that the question with respect to which the answer is exhaustivized is the same (including all the implicit contextual restrictions) as the one with respect to which the presupposed incorrect answer is exhaustivized? An anonymous reviewer has suggested that it should be adopted as a general characteristic of corrections that the negative and the positive conjunct address the same single *wh*-question like *Who smokes?* in (20), rather than a double *wh-yes/no*-question which falls apart into a negative and a positive subquestion *Who doesn't smoke?* and *Who smokes?* as was proposed in section [4.2]. Indeed, this could be made to guarantee the preservation of the same contextual restrictions. Also, this construction of the discourse topic (in contrast to the double *wh-yes/no* strategy) would explain the characteristic information structure of corrections, i.e. their accentuation and word order, and in particular the absence of contrastive topics, cf. (14) and (15). Notice also that Russian sentences with constituent negation are more appropriate as answers to positive than to negative questions, so (22-a) is more appropriate after (21-a) and (22-b) after (21-b).

(21) a. Kto kurit?
who smokes
'Who smokes?'

b. Kto ne kurit?
who not smokes
'Who doesn't smoke?'

(22) a. Kurit ne Oleg
smokes not Oleg
'It is not Oleg who smokes.'

b. Ne kurit Oleg
not smokes Oleg
'It is Oleg who doesn't smoke.'

Moreover, if the default exhaustivization of a negative answer to a positive *wh*-question provided something to the effect of the Russian sentences with constituent negation or the English negated clefts—EXH(*not John*)(*smokes*) = 'it is not John who smokes' (including the weak presupposition of the corresponding positive sentence 'it is John who smokes')—then the present analysis could be easily extended to English corrections, where neither the positive nor the negative conjunct need to contain any linguistic devices that encode exhaustivity by convention:

(23) John didn't go to PARIS, but to BERLIN.

Unfortunately, this is not the prediction made by most existing theories of default exhaustivization. For example, Groenendijk and Stokhof (1984) predict that the exhaustive interpretation of *not to Paris* with respect to the question *Where did John go?* implies that John didn't go anywhere (which is obviously wrong); von Stechow and Zimmermann (1984) predict in the same case that John went everywhere else except Paris.[25] In general, this is incompatible with the exhaustive interpretation of the positive conjunct which says that John went only to Berlin. Or rather, the two exhaustive interpretations are only compatible on the assumption that Paris and Berlin are the only relevant places under discussion: $C = \{\textit{Paris}, \textit{Berlin}\}$. However, on this assumption the exhaustive interpretation of the negative conjunct already implies that John went (only) to Berlin, so the positive conjunct is made completely redundant. This contradicts the intuition that the negative conjunct in (23) leaves it open where John actually went instead of Paris, whereas this information is provided no sooner than in the positive conjunct. Schulz and van Rooij (2006) argue that the exhaustive interpretation of a negative answer to a positive *wh*-question should be identical with its literal meaning, i.e. John didn't go to Paris, but it is unknown whether he went anywhere else. This would at least be compatible with the intuition that the positive conjunct provides a further specification of the place John went to in relation to 'some place other than Paris'. However, the way Schulz and van Rooij (2006) achieve this result in their theory is hard to reconcile with our assumptions about corrections. In their proposal, the fact that the speaker uses negation in an answer to a positive question should be seen as a signal that the speaker deviates from the standard form for answering positive questions, and ultimately as an indication of the speaker's limited competence on the issue *Where did John go?* The assumption of limited competence blocks inferences about other places John could have gone to. However, the normal exhaustive reading of the positive conjunct, that John went to Berlin and nowhere else, can only be derived on the assumption of the speaker's full competence on the issue where John went. In other words, in order to make Schulz and van Rooij's proposal work for us, one has to assume that the speaker is incompetent on the current question when uttering the negative conjunct, but competent on the very same question when uttering the positive conjunct in the very next instant. This assumption seems rather implausible. Moreover, none of these approaches predicts that exhaustivized negative answers to positive questions (weakly) presuppose an exhaustive positive answer that they negate—an assumption which is essential in our derivation of replacivity. It seems, what we need is an exhaustivity operator that makes roughly $\neg[max(\lambda x[\textit{John went to } x]) = \textit{Paris}]$ out of *Where did John go? John didn't go*

[25] Basically, von Stechow and Zimmermann (1984) exhaustivize with respect to the negated predicate, i.e. the question *Where didn't John go?* If Paris is the only relevant place where John didn't go, then he must have gone to all other relevant places.

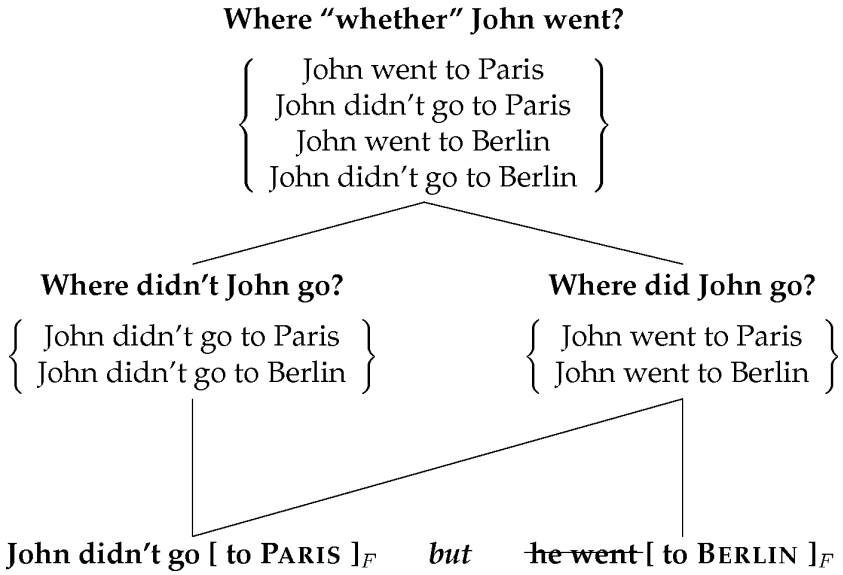


FIGURE 9: Discourse topics in corrections: The negative conjunct addresses both the negative part of the *wh-yes/no* question and shares the positive single *wh*-question with the positive conjunct

to Paris, where max returns the maximal (possibly plural) individual that has the property λx[*John went to* x] (cf. Rullmann 1995). This is close to Zeevat's (1994) notion of default exhaustivity, with the additional assumption that negation is neither part of the answer (or focus), nor of the question (or background), but takes scope over the whole proposition after the application of the exhaustivity operator to the answer and the question.

In sum, the assumption that the negative and the positive conjunct both answer the same single variable *wh*-question establishes a link between the deviant information structure of corrections (only foci, but no contrastive topics) and their replacivity property. However, does this mean that our view of corrections as instances of contrast that address double *wh-yes/no*-questions should be abandoned? In fact, it seems that corrections simply have this hybrid nature: they answer a single *wh*-question and at the same time a double *wh-yes/no*-question split up into a negative and a positive part, as shown in figure 9. The single *wh*-topic construal is responsible for the information structure and replacivity of corrections, while the *wh-yes/no*-construal accounts for the fact that there must be exactly one negative and one positive conjunct, as well as for other properties that corrections share with contrast relations, in particular the use of contrastive markers like the English *but* and the Russian *a*. In other words, this hybrid nature is what distinguishes corrections from oppositions with *y/n*-topics. Oppositions only have overarching *wh-y/n*-topics, cf. figures 6 on page 449 and 7 on page 450, while corrections are defined by a combination of the two topic types.

Of course, a number of hard questions arise once we allow for such non-tree structures as in figure 9 on the facing page, where one and the same utterance may be immediately dominated by two distinct topic questions. What constrains the set of questions addressed by the same utterance? Without strong constraints on this set the whole QUD-based theory of discourse structure loses its predictive power. What does it depend on, for instance, which of the questions addressed by the utterance determines its information structure? Although we cannot address these issues in detail here, it should be said that there is a number of phenomena that motivate structures with sentences dominated by more than one question. One of them is second occurrence focus, where the accentuation pattern suggests one construction of the discourse topic whereas a focus-sensitive particle requires a different construction to set its domain correctly. This group of phenomena still awaits a systematic study.

[4.5] *Why oppositions with* y/n*-topics are rare*

Finally, a few remarks are due on the observation made in section [4.2] that oppositions with contrastive polarity topics are barely acceptable and require very special context to be felicitous (cf. the discussion of (16-a)/(17-a)). This would follow naturally on the assumption that there is a preference for contrastive topics that are also given, contextually activated and talked about. For Büring's example (24) it implies that if Fred and Mary were previously mentioned and are talked about, one would prefer to go by people making Fred and Mary the contrastive topic as in (24-a). In contrast, if the talk is about food, so the beans and the eggplant were mentioned or are accessible via a bridging inference while Fred and Mary are new then it might be better to go by food and choose the structure in (24-b).

(24) a. [Fred]$_T$ ate [the beans]$_F$, [Mary]$_T$ ate [the eggplant]$_F$.
b. [Fred]$_F$ ate [the beans]$_T$, [Mary]$_F$ ate [the eggplant]$_T$.

This is not to suggest that the notions of givenness, referent activation or aboutness topic should be conflated with the notion of contrastive topic. Rather, just like subjecthood, definiteness and animacy are distinct notions which tend to fall together—subjects are definite and animate most of the time (Aissen 2003; Zeevat and Jäger 2002)—different varieties of topic tend to be aligned in a similar way. That is, there is perhaps no categorical requirement that contrastive topics be also aboutness topics or given, but an optimisation process prefers sentences with given contrastive aboutness topics.

In this light it is obvious that polarity values make bad topics. Although we have seen that in Büring's theory they are just as good contrastive topics as they are foci, one has to admit that they make little sense as aboutness topics or as entities subject to activation in memory. Therefore *ceteris paribus*, splitting by the

wh-variable (25-b), which makes a usual term the contrastive topic, is always preferred to splitting by the *yes/no*-variable (25-a). In fact, what seems like a *yes/no*-topic in (25-a) is most probably something bigger—perhaps, an open proposition of *Oleg not going* vs. *Oleg going* somewhere. These are entities that can be activated and talked about. This would explain why (25-a) is only appropriate in a context where Oleg not going and Oleg going somewhere (and not just *yes* and *no*) are somehow activated or salient.[26]

(25) a. Oleg [ne EZDIL]$_T$ [v PARIŽ,]$_F$ a [EZDIL]$_T$ [v BERLIN.]$_F$
Oleg not went to Paris but went to Berlin
b. Oleg [v PARIŽ]$_T$ [ne EZDIL,]$_F$ a [v BERLIN]$_T$ [EZDIL.]$_F$
Oleg to Paris not went but to Berlin went

[5] CONCLUSIONS AND OUTLOOK

We started with the observation that correction is often signalled by the same markers as contrast, especially the opposition type of contrast. This regularity is captured by universal semantic maps of Malchukov (2004) and Mauri (2008), and the one proposed in section [2], figure 3. Now we can say more about the nature of the links between the functions CONTRASTIVE COMPARISON, OPPOSITION, ADVERSATIVE and CORRECTION. All of them are relations between distinct answers to various subsorts of multiple variable questions. Moreover, ADVERSATIVE is more closely related to OPPOSITION than to CONTRASTIVE COMPARISON because both ADVERSATIVE and OPPOSITION involve *wh-yes/no*-questions—questions whose one variable is of the polarity type. In this paper we have shown that CORRECTION can be seen as another special case of a *wh-yes/no*-strategy. Its specific characteristics are: (a) the multiple variable *wh-yes/no*-question *Who "whether"* P*?* is split into single variable subquestions by polarity, i.e. one subquestion addresses the positive (*Who* P*?*) and the other the negative part of the question (*Who not* P*?*) ; (b) one conjunct negates that A is an exhaustive answer to the question *Who* P*?* while the other conjunct asserts that B is. Given an appropriate notion of exhaustivity, the latter condition could also be derived from the assumption that both conjuncts give exhaustive answers to the positive *Who* P*?* question (i.e. the negative con-

[26] Even in languages like German which seem to allow for contrastive topicalisation (movement to the prefield position) of just the negative particle as in (i), the contextual appropriateness conditions of such sentences are similar to those of (25-a). It is the open propositions of her confessing vs. not confessing something that serve as the contextual anchor, rather than just *yes* and *no*.

(i) Sie has zugegeben, den Mann ausgeraubt zu haben. Nicht aber hat sie gestanden, ihn
she hat confessed the man robbed to have not but has she confessed him
ermordet zu haben.
murdered to have
'She confessed that she robbed, but not that she murdered the man.'

junct addresses both the negative and the positive part of the question). Thus what makes CORRECTION and OPPOSITION so closely related is again the fact that they both answer a *wh-yes/no*-question. In fact, in light of the present proposal one would have to define OPPOSITION as a relation between distinct answers to a *wh-yes/no*-question *except* the subtypes characteristic of corrections and adversatives—therefore its position on the crossroads. In other words it is the similarities and differences between the types of discourse topic that determine which functions on the contrast semantics map are more closely, and which are less closely related.

The other question that we asked in the beginning of the paper was whether the combination of negation and *a* in Russian (or *but* in English) is a fixed collocation with correction semantics, or whether the correction semantics results independently from the properties of *a* as a general contrast marker in combination with the properties of negation. We have gone a long way in proving the latter point. Indeed, using the same notion of discourse topic one can define a single general meaning for *a*: a relation between distinct answers to a multiple variable question whose variable types are unspecified (except the *why-yes/no*-questions of the adversative type, because for those conjunction *no* is the preferred marker), as this is done e.g. in Jasinskaja and Zeevat (2008). Since correction is a special case of *wh-yes/no*, which in turn is a special case of a multiple variable question, correction falls within the domain of *a*. The same holds for the English *but*: since *but* marks *wh-yes/no*-strategies of all kinds, it can in particular be used for correction.

In turn, the replacivity property and the presuppositions that we find specifically in corrections could all be put on the account of negation and exhaustivity. The Russian case was relatively easy to handle because corrections in Russian require constituent negation, and exhaustivity is simply built into the conventional semantics of sentences with constituent negation. The same approach could be applied to English if negation is made to interact with pragmatic exhaustivity in the right way.

In sum, this paper offers a theory of correction that explains its marking patterns in Russian and English and its most central semantic and pragmatic properties.

There are still many loose ends, unanswered questions and problems. Let's mention just one of them because we did not get a chance to discuss it in the body of the paper. The semantic map proposed in section [2], figure 3, only connects the CORRECTION function to OPPOSITION. At least, this arrangement of functions is best motivated from the point of view of the theory of contrast based on topic question types. According to the strong contiguity claim of the semantic map approach, this predicts that whenever a contrast marker is used for CORRECTION it should also be able to mark OPPOSITION. Or in other words, if CONTRASTIVE COM-

PARISON and CORRECTION are marked in the same way, then OPPOSITION should be marked in the same way as well. Japanese is a language that falsifies this prediction. The relevant contrastive relations can be conveyed in Japanese by the converb marker *-te* (*-de*), roughly 'and', and the clause final marker *-ga*, roughly 'but' (Mauri 2008). *-Ga* has an ADVERSATIVE function. *-Te* looks like a general additive marker which in particular can be used for CONTRASTIVE COMPARISON. It is also used in corrections, as in the following example from Mauri (2008, p. 134):

(26) tyuumonsi-ta-no-wa kootya-de-naku-**te** koohii-desu
order-PRF-NR-TOP tea-COP-NEG-AND coffee-COP
What I ordered is not tea, but it's coffee.

However, OPPOSITION in examples like *John likes football, but Bill doesn't* is expressed by the marker *-ga*. This makes the marking region of *-te* discontinuous.

(27) John-wa sakka-ga suki da-**ga** Bill-wa suki ja-nai
John-TOP football-NOM likes COP-but Bill-TOP likes COP-NEG
John likes football, but Bill doesn't.

An *ad hoc* solution would be to draw an additional arc between CORRECTION and CONTRASTIVE COMPARISON, though this is not so appealing since it makes the semantic map weaker. Another possibility is to use the weak, diachronic interpretation of semantic maps: if *-te* were an older marker with a general additive/contrastive function (distinct answers to an unspecified type of question), while *-ga* were expanding from a purely adversative marker and took over OPPOSITION as a new function, it would create a 'hole' in the marking region of *-te*. Finally, the single meaning approach would come to terms with this deviant marking pattern if it could be shown that there is some independent reason that prevents *-ga* from being used in corrections. Then *-te* once again receives a general function of marking distinct answers to an unspecified type of question. This is so general that it covers in particular also CORRECTION. *Wh-yes/no*-questions are excluded from the marking domain of *-te* since there is a better marker for them, namely *-ga*. However, the CORRECTION-type *wh-yes/no*-questions are not excluded if there for some independent considerations *-ga* is not the preferred marker for that question type. Which of these or other solutions is right is to be clarified by future research.

REFERENCES

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, 21(3):435–483.

Anscombre, J. C. and Ducrot, O. (1977). Deux mais en français? *Lingua*, 43(1):23–40.

Asher, N. and Lascarides, A. (2003). *Logics of Conversation*. Studies in Natural Language Processing. Cambridge University Press.

Babby, L. H. (1980). *Existential sentences and negation in Russian*. Karoma Publishers.

Babby, L. H. (2001). The genitive of negation: a unified analysis. In *Annual Workshop on Formal Approaches to Slavic Linguistics*, volume 9, pages 39–55. Univ of Michigan/Michigan Slavic.

Blakemore, D. (1987). *Semantic Constraints on Relevance*. Basil Blackwell.

Borschev, V., Paducheva, E., Partee, B., Testelets, Y., and Yanovich, I. (2006). Sentential and constituent negation in Russian BE-sentences revisited. In *Formal Approaches to Slavic Linguistics: The Princeton Meeting 2005 (FASL 14)*, pages 50–65.

Brown, S. (1999). *The syntax of negation in Russian: A minimalist approach*. CSLI Publications.

Büring, D. (1997). *The Meaning of Topic and Focus: The 59th Street Bridge Accent*. Routledge, London.

Büring, D. (2003). On D-trees, beans and B-accents. *Linguistics and Philosophy*, 26:511–545.

Foolen, A. (1991). Polyfunctionality and the semantics of adversative conjunctions. *Multilingua*, 10(1/2):79–92.

Fougeron, I. (1990). *A* et les rapport interphrastiques en russe contemporain. *Revue des études slaves*, LXII(1-2):125–137.

Groenendijk, J. and Stokhof, M. (1984). *Studies in the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, University of Amsterdam.

Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10:41–53.

Haspelmath, M. (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In Tomasello, M., editor, *The new psychology of language: Cognitive and functional approaches to language structure*, volume 2, pages 211–243. New York.

Horn, L. R. (1989). *A Natural History of Negation*. University of Chicago Press.

Jacobs, J. (1982). *Syntax und Semantik der Negation im Deutschen*. Fink, München.

Jacobs, J. (1991). Negation. In von Stechow, A. and Wunderlich, D., editors, *Semantik. Ein Internationales Handbuch der Zeitgenössischen Forschung*. De Gruyter, Berlin.

Jacobson, P. (1995). On the quantificational force of English free relatives. In Bach, E., Jelinek, E., Kratzer, A., and Partee, B. H., editors, *Quantification in natural languages*, pages 451–486. Kluwer Academic Publishers.

Janda, L. A. (2009). What is the role of semantic maps in cognitive linguistics. In Stalmaszczyk, P. and Oleksy, W., editors, *Studies in honor of Barbara Lewandowska-Tomaszczyk*, pages 105–124. Peter Lang Publishers, Frankfurt am Main.

Jasinskaja, K. and Zeevat, H. (2008). Explaining additive, adversative and contrast marking in Russian and English. *Revue de Sémantique et Pragmatique*, 24:65–91.

Jasinskaja, K. and Zeevat, H. (2009). Explaining conjunction systems: Russian, English, German. In Riester, A. and Solstad, T., editors, *Proceedings of Sinn und Bedeutung 13*, volume 5 of *SinSpeC. Working Papers of the SFB 732*. University of Stuttgart.

Kasimir, E. (2006). On *'nicht ... sondern ...'* (contrastive *'not ... but ...'*). In Ishihara, S., Schmitz, M., and Schwarz, A., editors, *Interdisciplinary Studies in Information Structure*, volume 5, pages 107–151.

Kodzasov, S. V. (1996). Zakony frazovoj akcentuacii. In Nikolaeva, T. M., editor, *Prosodičeskij stroj russkoj reči*, pages 181–204. Institut russkogo jazyka Rossijskoj akademii nauk, Moskva.

Kreidlin, G. E. and Paducheva, E. V. (1974a). Vzaimodejstvie associativnyx svjazej i aktual'nogo členenija v predloženijax s sojuzom *a*. *Naučno-texničeskaja informacija*, 1974(10):32–37. Ser. 2, Avtomatizacija obrabotki tekstov.

Kreidlin, G. E. and Paducheva, E. V. (1974b). Značenie i sintaksičeskie svojstva sojuza *a*. *Naučno-texničeskaja informacija*, 1974(9):31–37. Ser. 2, Avtomatizacija obrabotki tekstov.

Lakoff, R. (1971). If's, and's and but's about conjunction. In Fillmore, C. J. and Langendoen, D. T., editors, *Studies in Linguistic Semantics*, pages 114–149. Holt, Rinehart and Winston, New York.

Malchukov, A. L. (2004). Towards a semantic typology of adversative and contrast marking. *Journal of Semantics*, 21:177–198.

Mauri, C. (2008). *Coordination Relations in the Languages of Europe and Beyond*. Mouton de Gruyter.

Mehlhorn, G. and Zybatow, G. (2000). Alte und neue Hüte der russischen Informationsstruktur. In *Linguistische Arbeitsberichte*, volume 74. Institut für Linguistik, Universität Leipzig.

Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics*, 49:91–136.

Rullmann, H. (1995). *Maximality in the Semantics of Wh-constructions*. PhD thesis, University of Massachusetts at Amherst.

Sannikov, V. Z. (1989). *Russkie sočinitel'nye konstrukcii*. Nauka, Moskva.

Schulz, K. and van Rooij, R. (2006). Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy*, 29(2):205–250.

Sweetser, E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge University Press.

Umbach, C. (2004). On the notion of contrast in infromation structure and discourse structure. *Journal of Semantics*, 21:155–175.

Umbach, C. (2005). Contrast and information structure: A focus-based analysis of *but*. *Linguistics*, 43(1):207–232.

Uryson, E. V. (2002). Sojuz *a* kak signal "povorota povestvovanija". In Arutjunova, N. D., editor, *Logičeskij analiz jazyka. Semantika načala i konca*, pages 348–357. Indrik, Moskva.

von Stechow, A. and Zimmermann, T. E. (1984). Term answers and contextual change. *Linguistics*, 22:3–40.

Zeevat, H. (1994). Applying an exhaustivity operator in update semantics. In Kamp, H., editor, *Ellipsis, Tense and Questions*, pages 233–269. ILLC, Amsterdam. Dyana-2 deliverable R2.2.B.

Zeevat, H. (2008). *Only* as a mirative particle. Ms. University of Amsterdam.

Zeevat, H. and Jäger, G. (2002). A reinterpretation of syntactic alignment. In de Jongh, D., Zeevat, H., and Nilsenova, M., editors, *Proceedings of the 3rd and 4th International Symposium on Language, Logic and Computation*, Amsterdam. ILLC.

AUTHOR CONTACT INFORMATION

Katja Jasinskaja
Centre for Advanced Study
Drammensveien 78
0271 Oslo
Norway
katja.jasinskaja@cas.uio.no