UiT The Arctic University of Norway

Faculty of Humanities, Social Sciences and Education

Aspect and Meaning in the Russian Future Tense:

Corpus and Experimental Investigations

Daria Kosheleva

A dissertation for the degree of philosophiae doctor, July 2022

Aspect and Meaning in the Russian Future Tense: Corpus and Experimental Investigations

Table of Contents

1 Introduction

This dissertation presents a comprehensive study of the Russian future tense in the framework of cognitive linguistics. The dissertation is article-based and consists of three peer-reviewed articles. Article I *Why markedness is always local: the case of Russian aspect* is accepted for publication, Article II *Looking into the Russian future* is published, and Article III *Russian future: an inside and an outside perspective* is under resubmission in international journals. According to the Regulations on the degrees philosophiae doctor (PhD) and philosophiae doctor (PhD) in artistic research at the University of Tromsø – the Arctic University of Norway (UiT)[1], a dissertation can be a monograph or can consist of several smaller works – a collection of articles (cf. §18, subsection 2). The current chapter – the discussion chapter[2] – serves as an extensive introduction to the articles.

My dissertation examines the distribution of the perfective and imperfective future forms, their future and non-future meanings, and the use of the future tense verb forms by both native and non-native speakers.

In Russian there are two main ways of referring to the events that will happen in the future. One is a perfective synthetic non-past form like *pročitaju* '(I) will read' and the other is an imperfective analytical future form comprised of an auxiliary verb *byt'* 'be' conjugated in the future tense and an imperfective infinitive — like *budu čitat'* '(I) will read'. In the dissertation, I study both forms and apply the same methods to them to compare their usage in the corpus and in an experiment.

The dissertation is divided into two parts: the discussion chapter, which serves as a general introduction, and a collection of three articles.

The articles in the dissertation are unified by a common set of themes and assumptions. The connecting thread in the articles is the usage-based approach. We explore the future tense via actual examples from the corpus, and by means of forms produced by the speakers in the experiment. We rely on the frequency parameter in order to show which aspectual form is considered "the default", or the unmarked member. We apply the notion of local markedness to the aspect-tense system to demonstrate that in some cases it is more relevant to talk about markedness on the level of tense, and not on the level of the verb. We also show that frequency is not a good indicator of markedness for all types of speakers: non-native speakers are not sensitive to the relative frequency of the verb in the aspectual pair. Finally, we dive deep in the corpus to establish how much of the future tense actually has a future time reference meaning. The ratio of future to non-future usages and the connections

[1] https://lovdata.no/dokument/SF/forskrift/2022-12-01-2087 (in Norwegian)

[2] *kappa* in Norwegian

between them show a complex interweaving, and this dissertation is an attempt to unravel the knots.

The discussion chapter serves to place the articles in the context of previous scholarship, shows the relationships between the articles, and is an opportunity to reflect on the results obtained and their limitations. The discussion chapter consists of sections 1 to 7. Section one is a brief introduction to the dissertation. Section two provides first and furthermost the context for the current research. It covers such broad topics as future time (2.1), future tense (2.2), and Russian future tense (2.3). Future time (2.1) is one of the primary meanings of the future tense. The ways people perceive time can be important for how they express themselves when talking about the future events. The subsection 2.2 on future tense gives a typological overview of the presence or absence of the future tense forms in the languages of the world. The subsection 2.3 about the Russian future tense explores the topic from different perspectives including diachrony, morphology, semantics, and language learning and acquisition. The primary aim of this subsection is to show what research has already been conducted and what we already know about the Russian future tense.

Section 3 shows the specific mechanisms from the framework of cognitive linguistics that my co-author Laura A. Janda and I implement in the articles in connection with the collected data. The section begins with a brief excursion into the history and basic concepts of cognitive linguistics (3.1). Subsequent subsections discuss the concept of prototype (3.2) as a central member of a radial category (3.3), the phenomena of metaphor and metonymy (3.4) as concepts used to describe the relationships between lexical and grammatical meanings, followed by the usage-based approach (3.5). Additionally, subsections 3.6 and 3.7 are dedicated to two complex concepts – markedness and complexity, which find their place in various linguistic theories.

Section 4 presents the methods applied in the articles of the dissertation. It begins with an overview (4.1) of the existing corpora for Russian placing the Russian national corpus in the context and giving a direction for potential future research (same techniques can be used with different corpora). Further I describe the method used in Article III (Kosheleva, under resubmission) concerning conducting a survey (4.2). Finally, in subsection 4.3, I draw attention to the statistical analyses performed in Article I (Kosheleva & Janda, accepted) and Article III (Kosheleva, under resubmission).

Section 5 is a collection of short summaries (5.1-5.3) of the articles in the dissertation. Each summary contains information about the background, hypotheses, methods, and main findings.

Section 6 focuses on the discussion of the articles. Subsection 6.1 lists the main findings and implications of the dissertation. Subsection 6.2 links the topics in the articles while 6.3 explores the common grounds for the findings. Subsection 6.4 shifts to proposals for further research that stem from the research questions posed in the articles. Section 7 is a brief conclusion of the discussion chapter (kappa).

2 Research status

This section provides a description of the current research status and connects it to the questions raised in the articles. All three articles are concerned with future tense, and it is therefore necessary to clarify some foundational issues about time in language, which could not be explored in detail in the articles. The first part of the section is devoted to the notion of time, namely what time is for human beings and how time, in particular future time, is expressed by language. The second part is a short overview of what future tense is and its typological characteristics. Finally, in the third and biggest part I refer to the scholarship dedicated to different facets of the future tense in Russian.

2.1 Future Time

2.1.1 What is time; how people perceive time

The average modern person views events through the prism of time. Life's journey passes through certain phases, including what is happening now, what we remember happening earlier, and what we expect to happen later. Usually, these phases are expressed through the past, present and future tenses, which will be discussed in subsection 2.1.2. Lakoff (1993, 218) describes time through the metaphor of motion in space. A metaphor [3] is defined as a mapping of one domain into the other, i.e., it is "cross-domain mapping" (Lakoff 1993, 203). Thus, time is understood in terms of a metaphorical journey. Humans do not have any special perceptual mechanisms that would help them to comprehend time, so according to Lakoff, the concepts of motion, entities, and locations facilitate the understanding of time.

Metaphorical motion in space can be of two types: TIME IS A MOVING OBJECT[4] and TIME IS STATIONARY AND WE MOVE THROUGH IT (Lakoff and Johnson 1980, 34). These two metaphors are also known as MOVING TIME and MOVING EGO in Clark's terminology (1973, 50). As the names suggest, the time metaphors differ in what is moving (Figure), and what remains static (Ground). In linguistics, the cognitive-semantic categories Figure and Ground are introduced by Talmy (1978, 419). Consider examples (1a) and (1b) from the Russian National corpus below:

(1a)	Približa-l-o-s′	Roždestv-o	Bogorodic-y…	Tix-ij
	approach-PST-N-REFL	Nativity-NOM.SG	Virgin-GEN.SG	quiet-M.NOM.SG
	i	nežn-yj	prazdnik,	kotor-yj
	and	gentle-M.NOM.SG	holiday.NOM.SG	which-M.NOM.SG

[3] The metaphor as theoretical concept is discussed in section 3.4 on metaphor and metonymy.

[4] Lakoff & Johnson (1980) use small capitals to refer to a metaphor. I am following their style.

ja vsegda očen′ žd-u.

I.NOM always very wait-PRS.1SG

‘The Nativity of the Virgin was approaching… A quiet and gentle holiday, which I always look forward to.’ [E. Kučerenko. O Bogorodice, duročke Maške i prostoj čelovečeskoj žizni (2015.09.21)]

(1b) Sejčas my približa-em-sja k

now we.NOM approach-PRS.1PL-REFL to

Roždestv-u Xristov-u k voploščeni-ju

Nativity-DAT.SG Christ-M.DAT.SG to incarnation-DAT.SG

Syn-a Boži-ja naš-ego

Son-GEN.SG God-M.GEN.SG our-M.GEN.SG

radi spaseni-ja.

for.sake salvation-GEN.SG

‘Now we are approaching Christmas, the incarnation of the Son of God for the sake of our salvation.’ [mitropolit Antonij (Blum). O Evangel′skom blagovestii (1974)]

In example (1a), the event of the Nativity is moving towards the speaker. The speaker remains motionless. Example (1a) demonstrates the MOVING TIME metaphor. Example (1b) shows the opposite situation: the speaker is moving in time towards Christmas. In (1b), the MOVING EGO metaphor is used.

Both metaphors are attested in many languages but the distributions of the metaphors within a language are different. Brdar and Brdar-Szabó (2017) researched time metaphors in English, German, Croatian, Hungarian and Romanian. They found an asymmetry in Croatian and Hungarian: the MOVING EGO metaphor is less acceptable than MOVING TIME.

Moore (2014, 15) indicates that both metaphors are structured around the ego. In addition to the ego-centered metaphors (MOVING EGO and EGO-CENTERED MOVING TIME), Moore (2014, 43-48) discusses three other metaphors associated with time: NOW IS A MOVER, A SITUATION IS A MOVER, and the PURPOSEFUL ACTIVITY metaphor. In the metaphor NOW IS A MOVER, the present (“now”) is a moving entity that is moving from the region behind the moving entity (past) into the region ahead of the entity (future). Example (2) represents the NOW IS A MOVER metaphor in English (Moore 2014, 43).

(2) The hour is approaching dawn.

In (2), the hour signifies “now”, and it is moving towards the future (the dawn), which is ahead of the mover. Example (3) presents an analogous Russian sentence, where “now” is represented by the word *vremja* ‘time’ which is moving towards a specific point, i.e., lunch:

(3)	Vremj-a	š-l-o	k	obed-u.
	time-NOM.SG	go-PST-N	to	lunch-DAT.SG
	‘Time was approaching lunch.’			

Similarly, a situation can play the role of the ego: example (4) illustrates the metaphor A SITUATION IS A MOVER (Moore 2014, 45).

(4)	The	candle	burned	from	dusk	to	dawn.

The candle burning is the situation that is moving along the path from an earlier time (“dusk”) to a later time (“dawn”). Example (5) is a straightforward translation equivalent in Russian:

(5)	Sveč-a	gore-l-a	ot	zakat-a	do	rassvet-a.
	candle-NOM.SG	burn-PST-F	from	dusk-GEN.SG	to	dawn-GEN.SG
	‘The candle burned from dusk to dawn.’					

The third additional metaphor described by Moore (2014, 45) is the PURPOSEFUL ACTIVITY metaphor. In this metaphor, the primary focus is on the destination, which is the purpose of an activity. The movement indicates the agent’s progress towards the purpose. In (6), the purpose is getting the work done:

(6)	We’re	halfway	through	the	job.

The agent is moving from earlier to later while progressing in doing the job. A similar phrase in Russian can illustrate this metaphor:

(7)	My	na	polputi	k	namečenn-oj	cel-i.
	we.NOM	on	halfway	to	outlined-F.DAT.SG.	goal-DAT.SG
	‘We’re halfway there (lit. to the outlined goal).’					

In (7), the destination is a goal expressed explicitly, and the agent is moving towards the goal currently marking that they are in the middle of this imaginary road. However, Russian example (7) is ambiguous: the goal and the way can be interpreted either literally or metaphorically, depending on the broader context.

All the metaphors described above help understand the concept of “moving forward in time” (Moore 2014, 50).

Besides various motion metaphors, there also exist different models of time. In the "Western" folk model of time, the future is in front of us, and the past is behind us (Radden 2003, 237). The distribution of Figure and Ground in the MOVING TIME metaphor may be more consistent with the Western folk model because the ego is facing the future coming at them. The Western model is not the only possible folk time model. From the perspective that the future is what we do not know and cannot see, it makes sense that the future is situated behind, and the past is in front of the eyes. Such a model of time is adopted, for example, in Malagasy, where *aoriana*, *any riana* 'after, behind' refers to the future, whereas *taloha, teo aloha* 'before, in the front' denotes the past (Ø. Dahl 1995, 199). Similar back-oriented futures are attested in Toba, Taos, Jaqaru, Kawki and Quechua (South-American Indian languages), Maori, and Classical Greek (Radden 2011, 16).

Two types of models of time, elements of which are reflected in Russian, are proposed by Arutjunova (1999, 687-695). The models called the Human path and the Time flow, are ego-oriented and connected to the spatial metaphors. The Human path combines the metaphors of movement and the traveler who occupies a spot on the path in the ranks of those walking (Arutjunova 1999, 689). The Ancestors walk along the path, and the Descendants are following in their steps. The Human path model is similar to "the future is behind" model: the Path directs the Human from the unknown future into the known past. The Human path is compatible with the MOVING EGO metaphor. In Russian, the Human path model is reflected in the usage of motion verbs such as *približat'sja* 'approach' and *nastupat'* 'advance' and prepositions of place like *pered* 'in front of', *pozadi* 'behind' in the context of time. In the Time flow model, the future is still behind, and the past is ahead: the wind of time blows in the back of the walking people (Arutjunova 1999, 690). The Time flow model is compatible with the MOVING TIME metaphor. In Russian, the Time flow model is represented in such expressions as *sledujuščij den'* 'following day', *uxodjaščij god* 'passing (lit. going away) year'.

In addition to *ahead* and *behind*, there are other spatial dimensions that help people navigate in time. On a timeline, the future may be situated either on the right or on the left. The rightward or leftward future orientation is usually not reflected in the language (Casasanto 2016, 170). However, speech can be accompanied by gestures pointing in one direction or another. In English, gestures pointing to the right typically accompany future-directed speech (Cienki 1998, 197).

Finally, the duration of time may also be experienced in different ways. Casasanto et al. (2004, 576-577) conducted a corpus study and found that English and Indonesian prefer distance metaphors to indicate the duration of time (e.g., *for a long time*), while Greek and Spanish have a tendency towards quantity metaphors (e.g., *mucho tiempo* 'much time'). Arutjunova's Time flow model also focuses on the importance of the duration. When we say *prošlo dva časa* 'two hours have passed', we direct the attention to how long a certain period of time lasted, not to a dot on a clock-face (Arutjunova 1999, 691).

To sum up, humans do not have a special mechanism for experiencing time. Different spatial metaphors are used to help anchor time in everyday reality. Spatial metaphors are reflected in different modalities (speech and gestures). Time is often anthropocentric (cf. MOVING TIME and MOVING EGO metaphors). The quantification of time can build on metaphors of distance and volume (long vs. much).

2.1.2 What linguistic means express time

Language has several categories associated with time, and the three that are most important are tense, aspect and modality (mood). Additionally, Klein (1994, 14) distinguishes among the following temporal expressions in language:

- "inherent temporal features of the verb (punctuality, durativity);
- complex verb clusters (*begin to sleep*, *continue to smoke*);
- temporal adverbials (*now*, *later*, *yesterday*);
- special particles (Chinese perfectivity marker *le*);
- principles of discourse organization."

In the dissertation, I focus primarily on the three main categories: tense, aspect, and modality.

Tenses help us navigate the dimension of time. Comrie (1985, 9) defines tense as the "grammaticalized expression of location in time". Note that tense is subject to metaphorization: the location in time should not be interpreted literally, as it is intended to serve as a metaphor. The Russian grammatical terminology is confusing because both 'time' and 'tense' are denoted by the same word *vremja*. This ambiguity can be disorienting for Russian native speakers when required to distinguish between time and tense.

Tense places situations on the temporal scale with references to other events (Comrie 1976, 3—5). Langacker reformulates the traditional definition of tense in the terms of cognitive linguistics: "tense imposes an immediate temporal scope, positioned with respect to the speech event, within which the profiled process must be manifested" (2008, 157). In other words, tense refers to the range of time when the situation (the profiled process) is happening viewed with respect to the speech event.

There is past, present, and future time and there are past, present, and future tenses. In addition, there exists a distinction between absolute and relative tenses. Russian primarily relies on absolute tenses, i.e., the tenses that "include as part of their meaning the present moment as a deictic center" (Comrie 1985, 36).

To clarify the relationship between time and tense, let us turn to the terms "topic time" and "time of utterance", introduced by Klein (1994, 3-9). Topic time is the time when

the event happened, and time of utterance is the time when the utterance was pronounced. Tense marks the relationship between the topic time and the time of utterance (Klein 1994, 5). The terms are analogous to Reichenbach's (1947, 288) "point of event" (E) and "point of speech" (S), similarly used by Comrie (1985, 122). According to Reichenbach (1947), tense is the relationship between E and S.

In example (8), which is a line from a famous poem *Zimnjaja noč* 'Winter night' by Boris Pasternak, the verb *goret'* 'burn' is used in the past tense. It means that the candle was burning (the topic time) prior to the time the speaker described the room (the time of utterance).

(8)	Sveč-a	**gore-l-a**	na	stol-e
	candle-NOM.SG	burn-PST-F	on	table-LOC.SG

'The candle **burned** on the table.' B. Pasternak. "Zimnjaja noč′".

For the (prototypical) present tense, the point of event and the point of speech overlap (Comrie 1985, 123). Example (9) from the RNC describes the current moment for the speaker when it's spring and the sun is shining:

(9)	Na	dvor-e	vesn-a,	solnc-e	**svet-it.**
	on	yard-LOC.SG	spring-NOM.SG	sun-NOM.SG	shine-PRS.3.SG

'It's spring, the sun is shining.' [Ljubov′ Kuznecova. «...Sobiraju razroznennye brëvnyški naroda svoego...» // «Vestnik SŠA», 2003.09.03]

In the future tense, the event occurs after the moment of speech (E after S). Example (10) illustrates a prototypical case of future tense use. The speaker says that they will go to the theatre to see the musical with friends. The event will happen the day after the point of speech (*zavtra* 'tomorrow').

(10)	A	ja	s	podrug-ami	zavtra
	and	I.NOM	with	girlfriend-INS.PL	tomorrow
	pojd-u		na	*«Notr-Dam de Pari».*	
	go-FUT.1.SG		on	"Notre-Dame de Paris".ACC	

'And I **will go** to "Notre-Dame de Paris" with my girlfriends tomorrow.' [kollektivnyj. Forum: Poxod v cirk (2010)]

Future tense and future time may also have a less straightforward relationship. When someone uses past tense forms, we are very often confident that the events did indeed happen in the past. You cannot have the same luxury in the future tense, particularly in Russian, where the future tense bears the function of future-in-the-past. So, by the time the speaker is telling about the event it might have already happened (see the discussion around example (11) in Comrie 1985, 110).

(11) Kolj-a skaza-l, čto on

Kolja-NOM say-PST.M.SG that he.NOM

prid-ët zavtra.

come-FUT.3.SG tomorrow

'Kolja said that he **would** (lit. **will**) **come** the next day (lit. tomorrow).'

In (11), the speaker does not specify whether Kolja is coming the same day, the next day or Kolja has already come and left. (11) is definitely a future tense example but not necessarily a future time example.

Languages differ in how much information they convey by lexical and grammatical means and to what extent time is grammaticalized as tense. Tense can be expressed via morphological marking or analytically, using auxiliary verbs. Consider examples (12a) and (12b) from the Norwegian Web 2015 (Bokmål) corpus given below.

(12a) Russer-ne **send-te** ham til front-en.

Russian-NOM.PL send-PST he.ACC to front-DET

'The Russians **sent** him to the front.'

(12b) Jeg **skal** **sende** gav-er og julekort.

I.NOM shall send present-PL and Christmas.cards.PL

'I **will send** gifts and Christmas cards.'

In 12a, the verb *sende* 'send' is in the past tense form *sendte* 'sent'. The past tense is expressed with the morphological marker *-te*. Example (12b) contains a future tense form, consisting of the auxiliary verb *skal* 'will/shall' and the infinitive of the verb *sende* 'send'. The morphology of the Russian future tense is discussed in subsection 2.3.2.

In addition to tense, aspect and modality also contribute to localizing events in time. In contrast with tense, Comrie defines aspect as a "way of viewing the internal temporal constituency of a situation" (1976, 3). In the Russian tradition, there is a distinction between two major classes: perfective and imperfective aspect. In general, the perfective aspect views the situation as a whole, while the imperfective aspect is more focused on the internal organization of the situation (Comrie 1976, 16).

Examples (13a) and (13b) from the National Corpus of Polish use imperfective and perfective aspect respectively.

(13a) <…> Dostojewsk-iego **czyta-ł** w oryginal-e.

<…> Dostoevsky-GEN.SG read-PST.M in original-LOC.SG

(He) **read** Dostoevsky in original. [Z. Smektała. Chcica czyli Billie Holiday to kurwa: poemat romantyczny, 2006.]

(13b)	Dawno	już	**przeczyta-ł**	wszystk-ie	książk-i
	long.time.ago	already	read-PST	all-ACC.PL	book- ACC.PL
	jak-ie	mia-ł	w	dom-u	<…>
	that- ACC.PL	have-PST.M	in	house-LOC.SG	

(He) already **read** all the books that he had at home a long time ago. [A. Barczyński. Ślepy los, 1999.]

Example (13a) is stating the fact that the person has read Dostoyevsky's novels in Russian, using the imperfective verb *czytać* 'read' without any indication of whether they have finished reading the novels. This is merely a statement reporting an activity. In (13b) the perfective verb *przeczytać* 'read' combined with the adverbial construction *dawno już* 'a long time ago already' is conveying the message that the subject finished reading all his books.

Aspect for verbs is in many ways analogous to countability for nouns. Metaphorically, the distinction between perfective and imperfective aspect can be viewed as analogous to the difference between rocks and sand. Janda (2003, 251) suggests the following names for the two metaphors: PERFECTIVE IS A DISCRETE SOLID and IMPERFECTIVE IS A FLUID SUBSTANCE. The properties of discrete solids and fluid substances (rocks, sticks vs. sand, water) manifest in different aspectual functions (for a full overview, see Table 1 in Janda 2003, 253-254). Here is how these two metaphors manifest themselves in Polish in example (13) above. In example (13a), the imperfective aspect shows the property of unboundedness, which corresponds to substances having no inherent edges. There is no reference to the beginning or the ending of the reading. In (13b), the number of books in the house that the subject could have read is limited, so the perfective aspect metaphorically represents a solid object with clear boundaries. Metaphor as a concept in the tradition of cognitive linguistics is described in more detail in subsection 3.4.

The morphological marking of tense and aspect varies across languages. WALS map (Chapter 69, Dryer 2013) shows the distribution of the following strategies for marking tense and aspect across the globe: tense-aspect suffixes are the most typical strategy with 667 languages from their sample of 1131 languages, followed by prefixes (153 languages), no morphological tense-aspect marking (152), a combination of strategies (146), and tones (13). Norwegian (see example (12a) above) uses suffixes, whereas Russian – the primary focus of this dissertation – falls into the mixed category (i.e., a combination of strategies).

Modality refers to the speakers' attitude to the situation and often goes beyond this definition (Bybee et al. 1994, 176). Modality is especially important in the future tense because we often deal with uncertain events.

2.1.3 What is future time and how real it is

"What is future?" is a complex question with a multitude of various facets, including physical reality according to Einstein's theory of relativity, the existence and development of humanity on Earth, physiological capabilities of a human brain and the development of philosophical thought throughout the centuries.

Time can be interpreted radically differently, as something objectively existing in the world, or as a product of our consciousness. Therefore, future time and the question of its reality lies in two domains ("objective" and "subjective"). In the "objective" domain, the debate is between Presentism and Eternalism. Hinchliff (1996, 123) says that according to Presentism, "the only things that exist are things that presently exist". So, Presentism denies the existence of the future because the future does not exist here and now. Eternalism considers both the three-dimensional space and time, which together form a four-dimensional category. In the 4D-space, a future-time version of an object is just one of many existing versions. Future things will become real when the future time comes (Putnam 1967, 240).

In the "subjective" domain, the reality of time depends on the observer. If the object of research is an individual as a conscious creature, we need to turn to an understanding of the mind. Future time may not appear to be real because of the asymmetry of time in the consciousness: human beings have past memories, but not future memories. (Riggs 2015, 50). The future tense is often less morphologically complex, and many languages use periphrastic instead of conjugated future forms (Chapter 67 WALS, Dahl & Velupillai 2013). The lack of certainty about the future time makes the future tense prone to be closely associated with modality.

On the other hand, humans can plan, make predictions, and expect the predictions to come true. In addition, we have emotions, such as fear, dread, and hope, that are directed at future events (Dainton, 2018). Thus, the abovementioned capabilities of the human mind can support the idea that both past and future events may seem to be equally real if they exist in the mind of a conscious being with well-developed cognitive capacities (Bosanquet et al. 1897, 235). Faye (1993, 259) leans towards the reality of the future time, if time is understood as the relationship between groups of events and how particular observers experience them.

As we can see, there is no clear unity among philosophers concerning the interpretation of time, from which different interpretations of the reality of future events follow. As mentioned above, the interpretations of reality (and irreality) of the future events are closely related to modality. Their interaction is discussed in Article II (Kosheleva & Janda 2022).

2.2 Future Tense

In the following subsections, I turn to a discussion of the future tense. The future tense and the future time are not synonymous, meaning that the use of the future tense in a sentence does not always imply that the action being described will happen in the future time. Conversely, an action that occurs in the future time can also be described using other verb tenses.

Future time is something that has not happened yet with regard to the present, the "now" moment. It is a part of the physical world, though sometimes we can imagine different future time scenarios and not all of them will necessarily come true. Future tense, on the other hand, exists in language. Future tense is one of the ways to refer to future time, but it is not the only one. Example (14) from Russian illustrates the use of the present tense in combination with the adverb *zavtra* 'tomorrow' to refer to an event that is going to happen at a future time[5].

(14)	**Zavtra**	ja	**id-u**	k	An'k-e
	tomorrow	I.NOM	go.IPFV-PRS.1SG	to	An'ka-DAT.SG
	na	dnjux-u.			
	to	birthday.party-ACC.SG			
	'Tomorrow I'm going to An'ka's for a birthday party.'				
	[Andrej Klepakov. Opekun // «Volga», 2016]				

Another important observation concerns the point of reference. Future time takes the current moment as a starting point, whereas future tense can be calculated from various points of reference, which are not necessarily equal to "now". Finally, future tense can be used for referring to things other than future time. The non-future time references are one of the main topics of Article II (Kosheleva & Janda 2022).

2.2.1 Future tense: typological characteristics

How common is future tense in the languages of the world? How can a language express future tense? Bybee & Dahl (1989, 56) studied a balanced sample of fifty genetically diverse languages to compare how the future tense is expressed cross-linguistically. Bybee and Dahl found that the future gram, i.e., the grammeme used to express the future tense, is expressed as a bound morpheme in roughly half of the languages. Example (15) from the Udmurt national corpus (udmcorpus.udman.ru) illustrates the use of the future tense in Udmurt (the Permic grouping of the Uralic family). In Udmurt, the future tense is formed by means of the suffixes -o- (-*e*-) or -*lo*-

[5] It can be argued that the present tense is used because the truth conditions refer to the plan and not to the event itself: the sentence is true as long as there exists a plan to carry the action out (cf. Bulygina & Šmelev 1992 in Nesset 1998, 179).

depending on the type of conjugation. The future tense suffix is added to the verbal stem followed by the suffixes denoting person and number except for the first person singular (Kiseleva & Efremov 2015, 2).

(15)	Ton	kyrja-lo-d —	mon	kyrja-lo,
	you	sing-FUT-2.SG	I	sing-FUT.1SG
	ton	bör-do-d —	mon	bör-do
	you	cry-FUT-2.SG	I	cry-FUT.1SG

You will sing — I will sing, you will cry — I will cry <…>.

[Lidiya Nyan'kina. Shuzi-Mazi (1996)]

The other half of the sample prefers a periphrastic variant (with an auxiliary verb, such as *will* in English). The Russian future instantiates both possibilities, differentiated according to aspect. The perfective future is expressed with a bound morpheme, whereas the imperfective future is expressed periphrastically.

The World Atlas of Language Structures (Chapter 67 WALS, Dahl & Velupillai 2013) indicates that around a half of the languages in their sample (110 out of 222) have some grammatical means to express future. The languages with inflectional future include Chukchi, Yukaghir, Nivkh, Evenki, Nenets, and Udmurt, spoken in Russia. European languages with inflected future tense forms from the sample are Latvian, French, Basque, and Spanish. There are a few languages in the Caucasus that express future tense morphologically, including Georgian and Lezgian. Other regions with inflected future languages include West Africa, Papua New Guinea and Central America. Example (16) shows a conjugated future tense form of the Latvian verb *satikties* 'meet'.

(16)	Es	tevi	**satik-š-u.**
	I	you	meet-FUT-1.SG

'I will meet you.'

At the same time, in the other half of the languages, future tense inflections are absent. An example of a language lacking future tense verb forms is Finnish:

(17a)	Tänään	on	kylmää.
	today	is	cold.PART

'It is cold today'.

(17b)	Huomenna	on	kylmää.

tomorrow is cold.PART

'It will be cold tomorrow'.

In Finnish (see examples (17a-b)) present tense is the neutral way of expressing future meaning. WALS focuses exclusively on languages that have a morphological marker for the future tense. Thereby Dahl and Velupillai limited their sample to very strict criteria. The WALS chapter 67 does not mention whether the morphological markers responsible for the future tense have any additional meanings or functions, i.e., such as gnomic or directive uses (cf. Russian perfective non-past in Article II, Kosheleva & Janda 2022).

Cross-linguistically, the main lexical sources of grammaticalization for the future tense are verbs that denote desire, obligation, and movement (Bybee & Pagliuca 1987, 109, 111). Example (18) shows a French future tense verb form (*le future proche* 'the near future'), comprised of the auxiliary verb *aller* 'go' and an infinitive.

(18)	Je	**vais**	**aller**	à	la	plage	bientôt.
	I	go.PRS.1SG	go.INF	to	the	beach	soon

'I will be going to the beach soon.'

Verbs of possession ('have') and existence ('be', 'become') are less common but still possible candidates for future tense markers. In example (19), the German *Futurum I* verb form is comprised of the auxiliary *werden* 'become' and an infinitive.

(19)	Ich	**werd-e**	**lernen.**
	I	will-PRS.1SG	learn.INF

'I will learn.'

Bybee and Dahl (1989, 90) specify that usually the candidates for grammaticalization are auxiliaries (or morphemes) with the meaning of desire; constructions meaning "moving towards the goal"; copulas or possession verbs with an infinitive or other non-finite verb forms.

Russian future tense does not make use of the typologically most common auxiliaries: verbs of desire, obligation, and movement. The Russian periphrastic future uses the verb of existence *byt'* 'be' as an auxiliary, and therefore the Russian future marker for imperfective verbs can be characterized as relatively rare. I describe the Russian future tense in more detail in the next subsection.

2.3 Russian future tense

The Russian future tense stands out for its heterogeneity for several reasons. First, as in the past tense, the opposition of aspects (perfective and imperfective) is preserved. Second, the future tense is heterogeneous in terms of structure: the future is formed

either with the help of the non-past inflections or by means of an auxiliary verb. Third, as I show in Article I (Kosheleva & Janda, accepted), the Russian future tense forms are heterogeneous with regard to frequency: the perfective form occurs fourteen times more often than the imperfective. Fourth, in addition to the prototypical future time meaning, the future tense conveys a whole group of other meanings, which will be partially discussed in subsection 2.3.4 with a review of the existing literature, and in more detail in Article II (Kosheleva & Janda 2022).

This subsection presents the state of the art of the field. I show what has already been shown to be true about the future tense in Russian with focus on the status of the future tense, its polysemy, the modal meanings expressed by the future tense forms, and the difficulties faced by non-native speakers of Russian when learning how to use the Russian future tense.

2.3.1 How future tense in Russian came to be the way it is[6]

Before moving on to the topic of the Russian future tense, I would like to start by giving a historical perspective. There are at least a few words to say about what the future tense looked like in the language spoken by the possible ancestors of those for whom Russian is now their native language.

First, let me outline the time frame. Nesset (2015, 11) highlights the following periodization. Modern Russian (or Russian as we now know and understand it) has been in use since the eighteenth century. As we go deeper in time, the language becomes less familiar from the perspective of Modern Russian. Middle Russian existed during the period from the 15th to the 18th centuries. Middle Russian was already separate from Belarusian and Ukrainian (Nesset 2015, 10) but had many features that distinguish it from Modern Russian. For example, Middle Russian had a past tense called the Pluperfect: *jesm' byl postavil* '(he) had put' (Andersen 2006, 236). The common ancestor for Belarusian, Russian, and Ukrainian — Old East Slavic (also referred to as Old Rusian or Old Russian) — was spoken between 1000—1400 AD (Nesset 2015, 10). Old East Slavic had a past tense Aorist (*xvali* 'praise.2SG') which was mostly obsolete in Middle Russian. In addition to Middle Russian and Old East Slavic, there were also Common Slavic (300—10000 AD), and Pre-Slavic (before 300 AD). We will look at the "nearest" future tenses: first in Old East Slavic, and then in Middle Russian.

In Old East Slavic, there were two types of analytical future. The first was formed using the perfective auxiliary verbs *počati* / *načati* (Nesset 2015, 150) with inceptive meaning (Andersen 2006, 235 mentions verbs *počĭnu* / *načĭnu* / *vŭčĭnu* 'begin' for Late Common Slavic) and an imperfective infinitive, e.g., *načnet' platiti* 'will pay'. Ivanov (1983, 351) and Nesset (2015, 150) also mention that the verbs *iměti* 'have' and *xotěti* 'want' were occasionally used as auxiliaries instead of the inceptive verbs.

[6] The name of this section is a modified quotation of the title of the book by Nesset (2015).

The other analytical future is the future perfect, which consists of the future tense forms of the auxiliary verb *byti* 'be' and a resultative participle (Andersen 2006, 234). The participle can be formed with the help of *-l-* suffix from both perfective and imperfective verbs: *budu stavilŭ* 'I shall have put.IPFV' and *budu postavilŭ* 'I shall have put.PFV'. It is not always clear whether the present tense of the perfective verbs had a future meaning. As the opposition between the perfective and the imperfective aspect grew over time, the perfective present gradually got the meaning of the simple future.

In Middle Russian, all future tenses from Old East Slavic are preserved.

The changes start happening during the transition from Middle Russian to Modern Russian. In the period between the 16th and 17th centuries the analytical future with the auxiliaries meaning 'begin' is being replaced with the future forms of the verb *byti* 'be', e.g., *budu stavit'* 'I will put' (Andersen 2006, 245). The future perfect also goes through a series of changes. First, the future perfect auxiliary verb (*byti* 'be' in the future tense) is used only in the third person singular form *bude(t)* and as an indication of a particular style. Then, *bude* begins to be used as a conditioning complementizer. The future perfect disappears entirely from Modern Russian before the beginning of the 20th century (Andersen 2006, 246-237).

Table 1 summarizes the results of the development of the future tenses. The only future tense that could be used with both perfective and imperfective verbs — Future Perfect — disappeared. The imperfective future acquired a new auxiliary verb. And as we show in Article II (Kosheleva & Janda 2022), the perfective present developed into a very peculiar tense with a primary future time meaning and a handful of metaphorical extensions. A more detailed description of the future tenses in the context of the tense-aspect system of Modern Russian is given in the next subsection 2.3.2.

Period / Aspect	Old East Slavic and Middle Russian	Modern Russian
Perfective	Perfective Present	Perfective Non-past
Both (PFV and IPFV)	Future Perfect	—
Imperfective	Imperfective Future (*načnu*-future)	Imperfective Future (*budu*-future)

Table 1. *Old East Slavic and Modern Russian future tenses.*

2.3.2 Future tense in the Russian tense system: morphological characteristics

In this subsection, I briefly characterize the tense system in Russian and define the place of the future tense(s) in the system. Table 2 is a summary of the Russian verb

tense system. It consists of past, present, and future and two aspects: perfective and imperfective. Every verb in the Russian language is either perfective or imperfective[7], and each column is valid only for the designated aspect. Both perfective and imperfective verbs have past tense. The past tense verb forms of both aspects have three singular forms (feminine, masculine and neuter) and one plural form. Only imperfective verbs have a present tense form. The past tense conjugation remains the same regardless of the aspect. The imperfective present tense form is conjugated according to person (first, second, and third) and number (singular and plural).

	Imperfective	Perfective
Past	*pisal* 'he wrote'	*napisal* 'he wrote'
Present	*pišet* 's/he writes' (non-past conjugated form)	–
Future	***budet pisat'* 's/he will write'** (periphrastic)	***napišet* 's/he will write'** (non-past conjugated form)

Table 2. *Tenses in Russian.*

Russian has two standard primary ways of expressing future tense. Depending on the aspect, the future tense can be expressed either synthetically, with the help of inflection (for perfective verbs), or analytically, with the help of an auxiliary verb 'be' in the future tense and the infinitive of an imperfective verb. The bottom line of Table 2 shows the future tense forms. The future form for the perfective verb *napisat'* 'write' is *napišet*. It is conjugated in the same way as the present tense form. The only difference is that the imperfective non-past form is interpreted as present tense, whereas the perfective is usually used to indicate the future time reference, and thus is referred to as the future tense.

The future form for the imperfective verb *pisat'* 'write' is *budet pisat*' 's/he will write'. The imperfective future tense has an unusual status: it is the only form that consists of two elements, between which other words can be inserted. For example, *budet mnogo pisat'* 's/he will write a lot', where *mnogo* 'a lot' is an adverb. The elements can also swap places depending on the logical stress in an utterance, which is usually put on the first element: *pisat' budet mno*go (lit. 'write will a lot'), *pisat' mnogo budet* (lit. 'write a lot will'), or *mnogo budet pisat'* (lit. 'a lot will write') are all possible options.

[7] Russian has also a few hundred verbs that can be interpreted as perfective or imperfective depending on the context (Zaliznjak & Šmelev 2000, 10). The biaspectual verbs do not have overt aspectual marking.

2.3.3 Non-past, present, or future? The status of the future tenses in Russian

There is no generally accepted opinion (consensus) about the status of non-past forms of the perfective verbs and the future imperfective verb forms. The following two reasons are the main sources of disagreement. First, the inflections of the perfective non-past form coincide with the inflections of the imperfective present (cf. *pišet* 's/he writes' and *napišet* 's/he will write' from Table 2). Hence, technically the form in question is the perfective present. Isačenko (1965, 445) viewed the non-past perfective form as the perfective present with the proviso that it is problematic to consider a form as expressing tense if the form can express the meanings [8]of the future, present, and past. Timberlake (2004, 95) recognizes the discrepancy between form (present tense morphology) and meaning (future time reference). Even though the non-past form does not report "events that are actual at the here and now of speech", Timberlake explicitly chooses to rely on the morphological form and terms the non-past perfective the perfective present.

Šatunovskij (2009, 195) takes a radically different position: the non-past perfective form is the future tense form. The present tense designates a situation that is simultaneous with the current moment of speech, which cannot be combined with the perfective aspect. Maslov (1990/2004, 521) supports labeling perfective non-past forms as "future", but does not provide arguments for his position. Švedova (1980, §1496) and Stojnova (2018) also follow the meaning of the forms, and label the two futures as the complex or periphrastic future (imperfective) and the simple or synthetic future (perfective).

An intermediate position is occupied by Bondarko (1971, 61) and Vinogradov (1947, 466): they both use the term "present-future" to refer to the non-past perfective form. Vinogradov (1947, 466) argues that "present-future" is a suitable name because "the action comes from the present", and then reaches its final stages in the future.

The second argument concerns the status of the imperfective future. The imperfective future form is periphrastic consisting of two separate entities: an auxiliary and an infinitive. Thus, the imperfective future can be viewed as non-grammaticalized. WALS (Chapter 67, Dahl and Velupillai 2013) adhere to this point of view, placing Russian under the "no inflectional marking of future and non-future distinction" category.

I respect the reasoning behind all the above-mentioned decisions. In the dissertation, I refer to the non-past perfective form as the perfective future. As shown in subsection 2.3.2, the periphrastic imperfective future has a place in the tense system of Russian.

2.3.4 Future tenses and future meanings in Russian

As discussed above, the primary meaning of future tense is reference to a situation at a time following the present time (Comrie 1985, 43), i.e., future time reference. The

[8] I will get back to the question of meanings expressed by future tense forms in subsection 2.3.4.

Russian future is not an exception: according to Švedova (1980, §1495-1496), the perfective and imperfective future tenses share the so-called categorical meaning of a situation that follows in relation to the grammatical point of reference. A similar definition, common for both forms, is given by Maslov (1990/2004, 515) and Stojnova (2018).

Another strategy regarding the relationship between the future tenses and future meanings is to start with the distinction between the specific (future-related) meanings of the perfective and imperfective future. Isačenko (1965, 444) and Vinogradov (1947, 466) focus on the temporal distance between the action and the moment of speech (the imperfective future) or a lack thereof (the perfective future). Bondarko (1971, 89-94, 102-104) and Timberlake (2004, 423) both rely on the aspectual differences, while Bondarko operates with a set of features, the presence or absence of which describes a particular form. The perfective future form refers to predicted events that can lead to results (Timberlake 2004, 423). It is a specific single action, which sometimes can be repeated (Bondarko 1971, 103-104). The perfective future form is characterized by the presence of such features as simultaneity, sequentiality, and localization of the action in time (Bondarko 1971, 102).

The imperfective future form is used for habits, iterative activities, or actions that may be left uncompleted (Timberlake 2004, 423); Bondarko (1971, 90) also adds processes to this list. The imperfective future form primarily refers to processes that are incomplete and/or repeated (Bondarko 1971, 89).

Russian future tenses and future time references do not always coincide. Russian future tense forms express a handful of meanings that are not always primarily associated with future time. The non-future meanings of the Russian future tenses and their relationships are the focus of Article II (Kosheleva & Janda 2022). In the current subsection, I briefly show the state of the art concerning future tense when used for other purposes. In Švedova's view (1980, §1498), the periphrastic (imperfective) future bears only the categorical meaning (reference to a point in the future time) whereas the simple (perfective) future has additional meanings.

It is not only the case that future forms can express non-future meanings. The other way round is also possible: future meanings can be expressed in various ways, not only by future verb forms. Stojnova (2018 §1.2 –1.5) lists the following possible forms and constructions that can refer to the future events:

- modal constructions with the auxiliary *byt'* 'be' in future tense (*budet dolžen* 'will have to', *objazan* 'will be required'; *budet nužno* 'will need', *možno* 'will be possible');
- infinitive constructions with future reference (*mne zavtra vstavat' rano* 'I need to get up early tomorrow');
- constructions with passive participles (*budut sudimy* 'will be judged');

- "futural" participles: present active participles derived from perfective verbs (*spojuščij* 'will have sung').

Stojnova also describes use of present and past tense forms with future meaning. The present can express planned actions (*zavtra sažus' rabotat'* 'I (will) sit down to work tomorrow'), prospective (*ja sovsem zasypaju* 'I am completely falling asleep') or immediate future (*dveri zakryvajutsja* 'the doors are closing'). Čujkova (2018, 53) notes that these uses of present tense forms to express future meanings are limited to certain semantic classes, namely verbs of motion and change of state.

Past verb forms in controlled situations (*nu, ja pošla* 'well, here I go [lit. went]') may also refer to actions in the future, cf. Stojnova (2018, §1.5.2) citing Bondarko (1971, 132-134). Furthermore, there is the construction that consists of the verb *stat'* 'begin, become' in the future tense and an infinitive of an imperfective verb (*stanu delat'* 'I will (begin to) do'): see Stojnova (2019) for a detailed description of the rivalry between *stanu* and *budu*.

2.3.5 Russian future tense and modality

Future tenses and modality always go hand in hand: both refer to events that might happen with a certain degree of probability. This expression of probability can be achieved in different ways with respect to future tenses and modality. In this subsection, I focus on the types of modality attested in sentences with Russian future verb forms.

To what extent do Russian future tense forms express modality? There is no common opinion on this matter. There are some who strongly believe that future tense is shaped by modality and can only be interpreted as a manifestation of modality (Klimonow 2011, Radbil 2011) and others who view modality as an "independent" element in the system (Petruxina & Li 2015).

Radbil identifies two types of Russian future tense. There is one type of future tense that describes events that we do not control: e.g., tomorrow will happen despite our attitude toward it – this is "future as a fact" in his terms (Radbil 2011, 255). These are predictions based on stable phenomena, such as the change of seasons and rising of the sun. The other future is termed "future as modality" (Radbil 2011, 254) and it corresponds to a strong and absolute certainty (i.e., modal assessment of the reality) that the speaker will perform the action s/he is referring to. In this case, the prediction is a promise that might or might not be fulfilled. So, if I say that I will come to work tomorrow, there is also a chance that something might prevent this from happening.

The system gets more complex if we consider the existence of both perfective and imperfective future forms and the fact that the perfective future is essentially a non-past form rather than being exclusively marked as future. For instance, Klimonow explains the perfective non-past form as if there are two separate homonymous forms. One form is responsible for a meaning that includes the present tense: *napišu* 'I write.PFV', as in *ja dlja tebja každyj den' napišu paru stroček* 'I write a couple of lines

for you every day' (a line from the lyrics of a pop song *Ljublju* by H1idden). The other homonymous form is responsible for the future tense: *napišu* 'I will write.PFV', as in *o tebe ja napišu krasivyj roman* 'I will write a beautiful novel about you' (a line from the lyrics of a pop song *Roman* by NEANGELY). However, it is not clear that we gain much theoretical insight from positing two homonymous forms for every perfective verb. The two so-called homonyms would then share no apparent semantic or etymological connection, and the fact that they have the same form would appear to be a mere coincidence. Such an observation fails to provide a thorough and insightful analysis.

What kind of modality is expressed by the future tense? Petruxina and Li gave an extensive description of this matter using a small set of 100 corpus examples and a short survey of native speakers. They find that the most common modal meaning for the Russian future tense is the epistemic modality followed by the volitive/performative modality (Petruxina & Li 2015, 79). An example of epistemic modality would be *postupit* 'will be accepted.PFV', as in *s takim otnošeniem k učebe on daže v PTU ne postupit* 'with that attitude toward studies, he won't even get into a vocational school', where the speaker is expressing their belief about what might happen. An example of volitive/performative modality is *zakonču* 'I conclude.PFV', as in *Na ètom zakonču moj doklad* 'With that I conclude my lecture', where the speaker uses the future tense form to signal their intent and the completion of their performance. According to Petruxina & Li, the interaction of future tense and modality motivates the non-future meanings of the Russian future tense forms, the topic of Article II (Kosheleva & Janda 2022).

2.3.6 Russian future tense forms: first language acquisition

This subsection provides a brief overview of the current status of research on the acquisition of Russian future tense forms. The acquisition of the future tense is particularly relevant for this dissertation because it helps to contextualize the challenges of foreign language learning. Additionally, comparing the acquisition of the future tense to other tenses highlights the unique nature of the Russian future tense.

In order to discuss future events, a person needs to be able to mentally navigate time and to have the corresponding verb forms in their linguistic repertoire. Future tense forms are a challenge for acquisition. Children whose native language is English typically do not begin to use future tense before they are three years of age according to Clark (1998, 379); or even four to six years old (Paul 2007, 300).

Russian-speaking children begin to convey future meanings during the so-called Optional Infinitive stage, when children use infinitives instead of finite verbs forms, normally between ages 1;5 and 2;4 (Brun et al 1999, 120). Brun et al. (1999) show that 25.6% of all incorrectly used (from the point of view of the adult grammar) infinitives in their dataset refer to future events or intentions. An example of such use is shown in (20), where the child is talking to his mother and pointing at his shirt at the same time (Brun et al 1999, 124):

(20)	rubašku	**snimat'**
	shirt	take-off-INF[9]

'(I will/want) to take off the shirt'.

Future tense forms may first appear in the child's speech at about the same age, preceded by the fully established past and present tense forms. Gvozdev (1961, 182-183) reports sporadic use of the future tense forms at age from 1;10 to 2 years. Most of the future tense examples found in Gvozdev's data express the perfective future: *pl'idu* [*pridu*] 'I will come', *kusu* [*ukušu*] 'I will bite'. Gvozdev (1961, 183) mentions only one example of the imperfective future and suggests that the child acquired it as a fixed expression that he often heard from the adults: *mama l'ugacca [rugat'sja] budet* 'mom will swear'. At the age of 2 – 2;2 years, the future tense forms, both perfective and imperfective (*budu est'* 'I will eat') are used regularly (Gvozdev 1961, 200-201).

Polinsky (2006, 16), analyzing various studies of Russian language acquisition (Kiebzak-Madera et al. 1997, Pupynin 1996, Gvozdev 1961) arrives at the conclusion that the hierarchy of verb form acquisition is as follows: infinitive > present (3sg > 1sg, pl > 2 sg, pl > 3 pl) > future.

Mastery of future tense forms does not happen without mistakes. Cejtlin (2000, 149) shows a few examples of mismatches between the aspect and the type of the future (synthetic vs. analytical), such as in (21). Example (21) shows an incorrect usage of the future tense where the auxiliary budu '(I) will' is combined with the perfective infinitive *narisovat'* 'draw'.

(21)	*ja	**bud-u**	**narisova-t′**	kukl-u
	I	will-1.SG	draw.PFV-INF	doll-ACC.SG

'I will draw a doll.'

2.3.7 Future tense and Russian language learners

Students studying Russian at university or at a language course typically encounter the Russian future tense in the last weeks of the first semester, which may correspond to the last stages of reaching A1 level according to CEFR (Common European Framework of Reference for Languages). At this level, students may have little motivation to talk about the future. Both in the classroom and as a part of the homework, common practice is to ask students to report about the general facts or the past events: what a student likes or does not like, what a student did last weekend. A popular topic that triggers the use of the future tense at the beginner level is talking about plans, i.e., what they want to do in their upcoming vacation.

[9] Brun et al. (1999, 124) glossed this as imperative, but morphologically this is an infinitive form.

What are the challenges that non-native speakers of Russian encounter when they (finally) need to use a future tense verb form? The main source of mistakes is the seeming easiness of the periphrastic form that non-native speakers tend to use on every occasion they need to talk about the future. Non-native speakers tend to overuse the imperfective form (Swan 2017, 825). In the overview of the different errors associated with aspect attested in the Russian Learner Corpus (the RLC), Olshevskaya (2018) mentions two additional challenges. First, since the morphology of present and future tense is the same, non-native speakers tend to use present tense form instead of the future tense form, i.e., they have a problem identifying the correct aspect. Second, if the non-native speaker is able to choose the aspect correctly, they may still encounter problems with conjugation or the abovementioned overuse of the *budu* future. I discuss the differences between native and non-native use of the future tense in Article III (Kosheleva, under resubmission).

How do we generally teach students to use the future tense in Russian? How do the textbooks and other resources that we use for teaching contribute to helping the students acquire the verb forms? In other words, how do textbooks tackle the weak spots, i.e., the mistakes discussed above?

In order to have a more precise understanding of the situation with pedagogical materials on teaching future tense to learners of Russian as a foreign language, I have selected a few textbooks and grammars for an overview. The materials in question are either widely used across the world, or specifically in Scandinavia. In addition to traditional hardcopy books, I have included two fully electronic resources. I show how the grammar is presented, and what types of exercises are used to train the use of the future tense. I focus on the following facets:

- the theoretical explanation about which aspect is presented first and to what extent;
- how much space and what role is given to the conjugation of perfective verbs in the future tense;
- whether any specific functions of the perfective and imperfective future tenses are given to distinguish between them;
- whether any submeanings of the perfective and imperfective future forms other than the reference to an action happening in the future time are mentioned in the resource.

The main points are summarized in Tables 3 and 4.

Table 3 describes the state of affairs in the textbooks and electronic resources.

In Table 3, two series of books that feature several textbooks that discuss the future tense are Russian *Poekhali! 1.2* and *Poekhali!-2 2.2*, and Norwegian *Azbuka 1* and *2*.

Danish *Møde i Petersborg* and Norwegian *Sosedi 1*[10] are individual textbooks. Norwegian *Min russiske reise* and American *Mezhdu nami* are full-fledged electronic resources that can be used instead of a textbook.

The order in which the aspects appear in the future tense chapter differs among the pedagogical materials. The imperfective future tense is placed first in *Azbuka 1*, *Sosedi 1*, and *Mezhdu nami.* As an intermediate option, both futures are given in a table side-by-side in *Poekhali! 1.2* and *2.2*, and *Møde i Petersborg*. The perfective future is featured first in *Azbuka 2* in the discussion of the submeanings of the two futures, and in *Min russiske reise* perfective and imperfective futures are given as two separate sections with perfective future coming first.

The conjugation is discussed in most resources: indeed, it is hard to avoid form inflections while actually teaching how to use these forms. There are several approaches to how to present the conjugation to students. First, simply listing a few types of conjugations including complex cases and exceptions, as in *Poekhali! 1.2*, *Møde i Petersborg*, and *Sosedi 1*. Second, the use of the conjugation patterns may not be in the focus in the theoretical block but still be a part of the exercises. This approach is practiced in *Poekhali!-2 2.2* as a logical continuation to the detailed explanation in the first part of the series, and in *Min russiske reise*. Third, the conjugation of the perfective future is given in comparison with the imperfective present, which has the same inflections.

As for the explanation of the differences in the use of perfective and imperfective futures, all textbooks seem to use essentially unique descriptions. The only function that has been named twice (once in *Poekhali! 1.2* and once in *Mezhdu nami*) is 'result' for the perfective future. Some submeanings that deviate from the prototypical meaning, namely action with future time reference, are mentioned in *Azbuka 2* and *Sosedi 1*. *Azbuka 2* names three submeanings and gives the following examples:

- potentially repeated action (Gnomic in the terminology of Article II, Kosheleva & Janda 2022)

(22)

On	vsegda	**najd-ët**	vyxod
he.NOM	always	find-FUT.3.SG	exit.ACC.SG
iz	trudn-oj	situaci-i.	
from	difficult-F.GEN.SG	situation-GEN.SG	

'He can (lit. will) always find a way out of a difficult situation.'

- possible/impossible action (Implicative in the terminology of Article II, Kosheleva & Janda 2022)

[10] *Sosedi 1* is a part of the series and the only one that contains the materials on teaching the future tense.

(23) Tol′ko on **ob″jasn-it** èt-o

Tol′ko	on	**ob″jasn-it**	èt-o
only	he.NOM	explain-FUT.3.SG	this-ACC.SG
tak,	čtoby	ja	ponja-l.
so	that	I.NOM	understand-PST.M.SG

'Only he can (lit. will) explain it in a way that I understand.'

- polite request with negation (Directive in the terminology of Article II, Kosheleva & Janda 2022)

(24)

Vy	ne	**skaž-ete,**	kotor-yj	čas?
you.PL	not	say-FUT.2.PL	which-M.NOM.SG	hour.NOM.SG

'Will you tell me what time it is?'

Sosedi 1 mentions the use of the hortative marker *davaj* 'let us' together with the perfective future: such uses fall under Directive in Article II (Kosheleva & Janda 2022).

Overall, the authors are understandably focused on the "correct" use of the forms in their prototypical future time reference meaning, which is optimal at the beginners' level.

Name	Presented first	Conjugation	Distinction and submeanings
Poekhali! 1.2 Černyšov & Černyšova 2019	a side-by-side table, then IPFV	explained in detail (incl. irregular *najti—najdu* 'find')	1 PFV: result vs. 3 IPFV: regular actions, process, offer
Poekhali!-2 2.2 Černyšov & Černyšova 2018	a side-by-side table reminder	not mentioned explicitly but used in exercises	repeated from *Poekhali! 1.2;* additional PFV: request, single action
Azbuka 1 Nordenstam 2013	IPFV first, then PFV, then a summary	not in focus	the meanings are not named explicitly; they are given through examples instead

	table for all tenses		
Azbuka 2 Nordenstam 2016	the submeanings for PFV are given before the IPFV	focus is on the comparison between IPFV present and PF future	3 additional submeanings for PFV: potentially repeated action, (im)possible action, polite request with negation; for IPFV: intent to perform an action
Møde i Petersborg Hertz 1999	a side-by-side table	two additional irregular verbs	PFV: which changes will happen; IPFV: what kind of activity will take place in the future
Sosedi 1 Bjerkeng et al. 2011	IPFV first (takes up a lot of space)	much attention to the conjugation of PFV future	PFV: additional paragraph on the use of *davaj* 'let us' in combination with PFV future; IPFV: the result of the action is not in focus
Min russiske reise Sokolova et al. 2021	PFV first with video and text instructions	heavy focus on conjugation and formation in exercises	no specific meanings are given; the emphasis seems to be on the general difference between IPFV and PFV aspects
Mezhdu nami Debenedette et al. 2015	IPFV first	conjugation of PFV future parallel with IPFV present	PFV: result, sequence, expected outcome; IPFV: duration, repeated actions, naming an action

Table 3. *Summary of the presentation of future tense in RFL textbooks.*

Table 4 shows how information on how the grammar of the future tense is covered in the grammars for learners of Russian. The resources in question include three grammars with exercises from Russia and the USA, and three reference grammars from Denmark, the USA, and the UK.

Like the textbooks, the grammars vary in the order they present the two futures. Only Filosofova (2020) totally prioritizes the imperfective future in terms of both order and space. Wade et al. (2020) first present the imperfective future, but further explains the meanings of the two future tenses side-by-side. Laskareva (2010) also puts perfective and imperfective futures on the same level. Levine (2009) describes the conjugation of

verbs in general before turning to the aspectual differences, so the perfective future naturally appears earlier in the book. In the chapter on aspect in the future (Chapter 7), he mentions the imperfective future first. Christensen (2004) has the perfective future and imperfective presented side-by-side, followed by the imperfective future. And finally, Timberlake (2004) gives priority to the perfective future.

Conjugation patterns are shown in all the observed grammars. The extent to which conjugation is presented varies from being targeted only in the exercises (Laskareva 2010), to mentioned together with the conjugation of other tenses (Timberlake 2004), to having a separate section on conjugation (Wade et al. 2020).

The reference grammars are quite often aimed at intermediate and advanced students, so the explanations on when to use the future tense forms are more detailed. Moreover, there is a bigger overlap between the submeanings that are described in the grammars compared to the textbooks in Table 3 above. For the perfective future, the submeaning 'result' (perhaps the prototypical meaning of the perfective aspect in general) is mentioned in four grammars. The submeanings 'completeness' (arguably synonymous with 'result'), 'single action', 'frequentative/potentially repeated action with *vsegda'* (Gnomic in the terminology of Article II (Kosheleva & Janda 2022)), and 'repetitive historic future with *byvalo'* (Habitual chains in Article II (Kosheleva & Janda 2022)) are mentioned in two grammars each. The authors of the grammars are even more in solidarity when it comes to the meanings expressed by the imperfective future. Those include 'progress' described in five grammars, 'repeated/habitual actions' — in four grammars, and 'facts' (can be classified as Gnomic in Article II (Kosheleva & Janda 2022)) — in two. In addition to the aspect-specific meanings, there are comments on the use of both future tenses. These relate primarily to uses expressing hypothetical situations and the use of the future tense in the context of past events.

Name	Presented first	Conjugation	Distinction and submeanings
Filosofova 2020	IPFV first and prioritized in terms of space; aspectual differences in the future are in a separate section (20.5)	special focus on form	PFV: a single complete action (result); IPFV: intention to carry out an action, repeated/habitual action, an action in progress
Laskareva 2010	IPFV and PFV side-by-side	conjugation is targeted in exercises	PFV: result; for advanced students: potentially repeated action when combined with *vsegda*/*nikogda* 'always/never' with PFV future; IPFV: process, regular action, fact

Levine 2009	PFV appears first in Basic verb conjugation (212); when tenses are presented, IPFV is first (239)	conjugation is given earlier than the aspectual opposition	distinction in the use of PFV and IPFV futures is explained in terms of the use of aspect in general; PFV: completion, limited duration; IPFV: action in progress, habitual actions; common for both aspects: after *esli* 'if' and *kogda* 'when'
Christensen 2004	PFV future side-by-side with IPFV present; IPFV future follows after	special focus on conjugation	PFV: simple action, single action that is sequenced with other future actions, repetitive actions with *byvalo* 'it happened'; IPFV: repetitive actions, action in progress
Timberlake 2004	PFV first	conjugation of the future tense verb forms is described together with other tense forms	PFV: "predicted events that lead to results"(423); IPFV: facts, actions that "are not expected to be completed"(424); hypothetical (*esli* …, *to*…) uses and directive uses with hortative markers *pust* 'let, may' and *davaj* 'let us' are mentioned (in sections on conditionals and imperatives respectively) regardless of the aspect
Wade et al. 2020	IPFV first; the meanings/functions with examples are given side-by-side	a separate section on conjugation	PFV: result, sequences of actions, warnings, impossibility of action (with negation), historic future in alternations *to*…, *to*… and with *byvalo* 'it happened', frequentative use with *vsegda* 'always'; IPFV: progress, "the way the action was carried out", repeated actions including frequentative genuine repetition; both PFV and IPFV: future tense in reported speech, after the conjunctions like *after*, *as soon as*; in question of intent the question is in IPFV, the

			aspect in the answer is context-dependent

Table 4. *Summary of the presentation of future tense in RFL grammars.*

As we see from the grammars presented for review, the future tense in the grammars is described in more detail than in the textbooks, and both aspectual variants are given approximately the same amount of space. In addition, Christensen (2004) even notes that the perfective future is more frequent than the imperfective, which we confirm in the corpus study in Article I (Kosheleva & Janda, accepted). Many of the meanings comprising the radial category of the future tense in Article II (Kosheleva & Janda 2022), are also mentioned in the grammars. In short, students who turn to reference grammars can get a better understanding of the Russian language as opposed to those who just rely on the explanations in textbooks.

3 Theoretical framework

In this dissertation, I follow the theoretical framework of Cognitive Linguistics. The aim of the current section is to provide a schematic picture of Cognitive Linguistics as well as of its historical development (3.1). Further on, I illustrate how the framework's concepts can be applied to various linguistic phenomena, and in particular the future tense. Subsection 3.2 covers the concept of a prototype, which appears in the prototype – periphery structure of marked and unmarked members of an opposition in Article I (Kosheleva & Janda, accepted). The concept of the prototype plays a central role in Article II (Kosheleva & Janda 2022) in the radial category of Russian future tense meanings (for radial category cf. subsection 3.3). The members of the radial category in Article II (Kosheleva & Janda 2022) are connected to each other via metaphorical and metonymic extensions: metaphor and metonymy are explained in detail in subsection 3.4. The basics of the usage-based approach are covered in subsection 3.5. We apply the usage-based approach to support the scalar interpretation of markedness in Article I (Kosheleva & Janda, accepted). Our radial category model relies on empirical corpus data utilizing the usage-based approach in Article II (Kosheleva & Janda 2022). And finally, I use statistical methods to determine the factors decisive in users' choice of forms in Article III (Kosheleva, under resubmission) also in accordance with the usage-based approach. The last two subsections of the current section do not rely solely on the theoretical framework of cognitive linguistics since markedness is a concept shared across frameworks. However, as Janda (1995) has shown, markedness is strongly motivated by the structure of radial categories, a theoretical notion that is central to cognitive linguistics. I dive into two topics: markedness (3.6) – in order to give context to Article I, and complexity (3.7) manifested in both Article I (Kosheleva & Janda, accepted) and Article III (Kosheleva, under resubmission).

3.1 Cognitive linguistics: history and basic concepts

Cognitive linguistics is a cluster of approaches relying on a postulate that language emerges from the general cognitive abilities of a human being (Geeraerts & Cuyckens 2010/2012, 2, 6). Cognitive abilities include the abilities "to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (Gottfredson 13, 1997).

Human beings strive to make sense of their world and convey their understanding of the world using language. Thus, studying language involves understanding how people deal with conveying meanings. As pointed out by Geeraerts and Cuyckens (2010/2012, 5), there are three aspects to meaning that are the fundamentals of cognitive semantics: 1) the semantics itself; 2) "the encyclopedic nature of linguistic meaning"; and 3) "the perspectival nature of linguistic meaning". The perspectival nature presupposes that the language does not merely reflect the objective reality as it is, but imposes a structure on the world though categorization. The categorization

phenomena include prototypicality, polysemy, mental imagery, cognitive models etc. (Geeraerts and Cuyckens 2010/2012, 2). I address these issues as applied to Russian future tense in the articles in this dissertation, together with my coauthor Laura Janda, and on my own.

Cognitive Linguistics as a discipline originated in the 1970-1980s. A forerunner of Cognitive Linguistics is Eleanor Rosch (1973a-b, 1978) who researched human cognitive categories and found that they depart from Aristotelian assumptions. The western philosophical tradition operates with Aristotelian categories, which are based on the concept of boundedness. Categories are fixed and structured by boundaries. Rosch developed a different understanding of a human cognitive category with a central member — a prototype and peripheral members, which form a radial structure. Cognitive categories do not have clear boundaries; they have internal structure motivated by prototype. Consider an example with colors. A human being (who is a native speaker of English) can clearly distinguish between prototypical blue and green. However, turquoise, as a more peripheral entity, may be categorized differently by different speakers. Linguistic categories of meaning as human cognitive categories are radial in nature (cf. subsections 3.2 and 3.3).

Further development of the framework of cognitive linguistics continued in the writings of scholars such as George Lakoff, Ronald W. Langacker, and Leonard Talmy. Their most influential works are *Metaphors We Live By* (Lakoff & Johnson 1980), *Foundations of Cognitive Grammar* (Langacker 1987), and *Toward a Cognitive Semantics* (Talmy 2000).

Lakoff and Johnson, a linguist and a philosopher, did research on metaphor. In their understanding (1980), metaphor is not just a figure of speech, but a cognitive mechanism, which serves to structure human perception. Lakoff and Johnson describe how metaphor functions as a tool for understanding the world on a "simple-to-complex" basis. People's knowledge of simple concrete phenomena through physical and social experience serves as the basis for understanding more abstract ideas and concepts, such as feelings or time. This is achieved through mapping from a source domain to a target domain (see subsection 3.4).

Langacker's *Foundations of Cognitive Grammar* (1987, 1991) presents a framework which portrays grammar as symbolic in nature and defines concepts such as noun, verb, modifier etc. according to the symbolic grammar. For example, to define a noun, Langacker departs from the prototype for the noun, which is a physical object. Then he proposes that a noun "designates a region in some domain, where a region is … a set of interconnected entities" (ibid., 15). For the word *moment* the domain is time, and the region is a certain portion of time.

The foundations of cognitive grammar lie in the "inseparability of semantics and syntax". The difference between grammar and lexicon is not categorical, it is scalar: grammar and lexicon form a continuum, in which various linguistic phenomena function.

To illustrate this statement, let us consider the following three examples: one example of a highly syntactic item (a grammatical case), one example of a highly semantic item (the meaning of an adjective), and an item that has both syntactic and semantic properties (a preposition).

To begin with, let us examine the phenomenon of case. Case might seem purely grammatical. Russian has six cases expressed by desinences that convey the syntactic relationships that hold for noun phrases. However, these cases also bear meanings. For instance, the meaning of the genitive case is comprised of four submeanings: a source, a goal, a whole, and a reference (Janda & Clancy 2002, 112). When genitive is used to mark a source, the focus is not on the source itself but on the item that relates to the source. For example, in the phrase *maslo delajut iz moloka* 'butter is made from milk', the word for milk is marked with the genitive case, indicating the source (and here the source meaning of the genitive is further enhanced by the preposition *iz* 'from'). However, this is not a statement about milk, but rather about butter. In the relationship governed by the genitive case between a whole and a part, the attention is on the part. For example, in the phrase *kusok saxara* 'sugar cube [literally: a piece of sugar]', the noun for sugar refers to a whole type of substance and is marked in the genitive case, but the attention is only on a single cube. In all instances, the genitive tends to have a background role, its primary meaning is "yielding the focus of attention to something else … in its proximity" (ibid., 112). The example of the genitive case shows that even a case that is syntactic in function has meaning.

On the other end of the grammar-lexicon continuum we can find, for example, adjectives – words used to convey various properties, i.e., size, color, shape etc. The primary purpose of adjectives is lexical. The adjective *krasnyj* 'red' is a word, its primary meaning is "the color of blood, ripe strawberries, or a bright poppy flower" (Ožegov 1990). However, *krasnyj* 'red' is rarely used as a separate entity; as an adjective it has to modify something, a noun, e.g., *krasnyj zakat* 'red sunset'.

In between the cases and the adjectives, we can find prepositions, which possess both lexical and grammatical properties. For instance, the Russian preposition *v* 'in' governs the locative and accusative cases and bears the meaning of either a location, as in *v gorode* 'in the city' or a path into a container, as in *v gorod* 'into the city'.

To sum up, even phenomena that at first glance are more lexical or grammatical, carry features of both.

Similar to Lakoff and Johnson (1980), Langacker (1987, 2) raises the importance of figurative language, including idiomatic and metaphorical expressions. Metaphor serves as a means of constructing our mental world, when our bodily experiences are projected into another domain, such as time or purpose (Langacker 2008, 36). I discuss metaphor (and metonymy) in more detail in 3.4. In this subsection, however, I would like to mention another important term of cognitive grammar: construal. Construal is a crucial function in the interpretation of conceptual content. Langacker explains construal by means of a metaphor in which the content is the scene (as in the theater)

and the construals are ways of viewing what is happening on the scene (Langacker 2008, 55). An example of construal is verbal aspect. In Russian, the speaker can use perfective or imperfective aspect depending on the way the person views the event in question. The event of reading a book can be reported with the perfective verb *pročitat'* 'read.PFV' if the result is important for the speaker: the book is read. If the speaker just wants to mention the fact that someone read the book and does not care to specify the outcome, the imperfective verb *čitat'* 'read.IPFV' is preferred. The situation remains the same but the aspect changes depending on the construal.

Figurative language is also a part of conceptualization (Langacker 2008, 43), to which Talmy's book (2000) is dedicated.

In *Toward a Cognitive Semantics* (2000), Talmy is interested in defining and describing the "linguistic representation of the conceptual structure" in the terms of cognitive linguistics. The conceptual structure is understood quite broadly: it reflects the "content experienced in consciousness" (Talmy 2000, Vol.1, 4). The concepts in question include space and time, motion and location, force and causation, etc. The organization of patterns and processes in language responsible for the expression of the abovementioned concepts (i.e., how the concept of time is experienced in consciousness and then expressed in the language) constitutes the basis of cognitive semantics.

Talmy (2000, Vol.1, 177) investigates how language structures space by the process of schematization, which is similar to Langacker's construals (and metonymy to some extent, cf. subsection 3.4). Schematization is a process where language chooses "certain aspects of the scene to represent the whole" (ibid., 177) and ignores others. Important notions for schematization are Figure and Ground adapted from Gestalt psychology (ibid., 184; Langacker 2008, 58). Figure is understood as a prominent (metaphorical) object whereas Ground is the static background. In the sentence *the bike stood near the house* (Example (2a) in Talmy 2000, Vol.1, 182), the bike is the Figure that stands in front of the house, which plays the role of the Ground.The bike situation can also be described in terms of trajectory and landmark alignment (Langacker 2008, 66). The attention is focused on the bike, that plays the role of the trajector, "the most prominent participant, … construed as being located" (ibid., 70). The house receives the secondary focus and is considered a landmark.

In *Structures that relate events* (2000, Vol.1,345), Talmy expands the relationship between Figure and Ground into temporal, causal, concessive, and additive domains. A complex sentence with a main and a relative clause represents cross-related figure-ground events with a temporal relationship. Such cross-related events occur in the data in Article II (Kosheleva & Janda 2022), in the examples with the posterior future. The main clause (the Ground) is usually in the past, while the relative clause contains the

future tense form (the Figure). Example (25) from the dataset[11] serves as an illustration of cross-related events.

(25)	Ja	**duma-l-a**,	čto	my	
	I.NOM	think-PST-F	that	we.NOM	
	sjad-em	na	kak-oe-nibud′	povalenn-oe	
	sit.PFV-FUT.1.PL	on	some-N.ACC	knock.down.PFV-PST.PASS.PTCP-N.ACC.SG	
	derev-o	i	**bud-em**	**razgovariva-t'**.	
	tree-NOM.SG	and	will-1.PL	talk.IPFV-INF	

'I **thought** that we **would** (lit. **will**) sit down on some fallen tree and (lit. **will**) **talk**.' [Kejt Uinslet: «Naše prošloe dolžno byt′ s nami» (2004) // «Èkran i scena», 2004.05.06]

The Ground event is thinking expressed by the past tense verb form *dumala* '(I) thought. The Figure events are sitting down and talking expressed by the perfective future tense verb form *sjadem* '(we) will sit down' and the imperfective future tense verb form *budem razgovarivat'* '(we) will talk' in the subordinate clause. In such cases, the past tense in the main clause serves as the Ground, from which the actual time of the event put in the future is calculated, which is also in the past. The events that happened or did not happen are situated in the past. The subordinate clause events happened later, and that is the Figure that is calculated from the perspective of the Ground.

Another important notion for this dissertation in relation to Talmy's work is the semantic category of force dynamics: "how entities interact with respect to force" (2000, Vol.1, 409). The member of the force dynamics model that implements the force is called the Agonist, and the member that resists the force is the Antagonist (ibid., 413). In Article II (Kosheleva & Janda 2022) the examples containing a main verb with Implicative meaning and an infinitive follow the force dynamics model, as in example (26):

(26)	Rebënok	daže	**smož-et**	**sozda-t′**
	child.NOM.SG	even	be.able.PFV-3.SG	create.PFV-INF
	svo-j	sajt	na	vKIDS.ru.
	own-ACC.SG	website.ACC.SG	on	vKIDS.ru

'Even a child **will be able to create** their own website on vKIDS.ru.' [Jurij Zubcov. Čem paxnet janvar′ (2002) // «Domovoj», 2002.01.04]

[11] The complete dataset is available at doi.org/10.18710/MHWRGE.

In example (26), the implicative verb *smoč* 'be able' is expressing the force dynamic of the subject *rebenok* 'child' (Agonist). The Agonist opposes the Antagonist — the difficulties faced by somebody very young and inexperienced while creating a website.

In this subsection, I briefly reviewed some of the concepts that have been developed within cognitive linguistics. In cognitive linguistics, language is not considered to be a separate human ability: language emerges from basic human cognitive abilities, such as, for example, the ability to categorize.

People use language to fulfill their own communicative needs. They understand what is being said to them based on their personal observations and encounters. And that is how linguistic meaning is anchored in human experience. Linguistic meaning is a product of human interaction; it is based on "mutual assessment of interlocutors' knowledge, thoughts, and intentions" (Langacker 2008, 4).

The products of language use — recordings of oral speech, written texts, videos containing people talking — lay the foundation for data, which can be studied by linguists. By doing research on these data, linguists can confirm or reject hypotheses about the structure of particular linguistic phenomena in individual languages and language in general. Thus, cognitive linguistics deals with data obtained as a result of language use by speakers and can rightly be considered a usage-based discipline. For more detail on the usage-based approach, see subsection 3.5.

In this dissertation, I work with actual language data produced by speakers, which includes but is not limited to, using corpora – large collections of texts (cf. Article I and Article II), and experiments with native and non-native speakers (Article III). In this theoretical framework section, I describe the main linguistic concepts and terms that are of use for my research (subsections 3.2-3.7).

3.2 Prototype

In order to understand what a prototype is, we need to take one step away and look at categorization – a process which strongly relies on prototypes. The act of categorization is a fundamental cognitive activity performed by humans. Croft & Cruse (2004, 74) define categorization as an ability to perceive an individual entity or experience as an instantiation of something more abstract.

Let me demonstrate a few examples of categories. Furniture is a category that is comprised of beds, sofas, tables, and chairs. Chairs[12] are also a category. Other objects that belong to the chair category include kitchen chairs, Windsor chairs, rocking chairs, office chairs, pushchairs etc. In this example, furniture can be called a

[12] The "chair" example is inspired by Lakoff (1987, 41) and Laura A. Janda's comparison of Russian and Czech in a lecture given during the course RUS-3030 *Concepts and categories: contemporary Russian cognitive linguistics* at UiT, Tromsø.

superordinate level category, chair is a basic level category, and office chair is a subordinate level category.

According to Cambridge English Dictionary, a chair is defined as “a seat for one person that has a back, usually four legs, and sometimes two arms”. And when I say “chair”, this is probably what you would imagine: an object to sit on with four legs, like a kitchen chair. What about a zaisu, a type of traditional Japanese furniture that has a back, sometimes armrests, but no legs? A kitchen chair will be a prototype of a chair, and a zaisu, at least for an average European, is not the first thing that comes to mind. So, what is a prototype?

Prototype theory was described in the 1970s by Eleanor Rosch. According to the theory, a prototype is the best exemplar of a category. The prototype is usually the first stimulus that is associated with the category (Rosch 1973a, 330). It is “the clearest case of category membership defined operationally by people’s judgments of goodness of membership in the category” (Rosch 1978, 36).

Prototypes can be found not only in the material world but also on an abstract level, e.g., in linguistics. A linguistic category can have a central prototypical member and peripheral members. As shown in Article II (Kosheleva & Janda, 2022), the Russian future tense has many meanings, such as Future, Extended Future, Gnomic, and Directive. The most salient meaning is Future, it is the prototype. Together, Future and other meanings form a radial category.

3.3 Radial category

A group that consists of a prototype and other less salient objects or ideas of the same type can be seen as a radial category. A radial category has a complex network-like structure consisting of the central subcategory (the prototype) and its noncentral extensions (Lakoff 1987, 91). The noncentral extensions are understood through the central subcategory. The radial category has a family resemblance structure, often with fuzzy boundaries (Geeraerts 2006, 146).

In subsection 3.2, I mentioned the prototypical chair and different types of chairs. In English, the armchair, the wheelchair, the pushchair, and the office chair all belong to the chair category. Russian carves up the relevant semantic space differently. In Russian, there are *stul'ja* ‘chairs’, *kresla* ‘armchairs’, and *koljaski* ‘strollers’. *Kreslo* ‘armchair’ and *ofisnoe kreslo* ‘office chair’ belong to the armchair category. *Detskaja koljaska* ‘pushchair’ is a type of a stroller. The wheelchair is tricky: it is called *invalidnoe kreslo* lit. ‘disabled armchair’ and *invalidnaja koljaska* lit. ‘disabled stroller’. Both terms can be used interchangeably; however, in the Russian National Corpus *invalidnaja koljaska* is more frequently used than *invalidnoe kreslo* (220 examples vs. 121 examples in the whole corpus). The example with chairs shows that the boundaries of categories can be fuzzy both within the frame of one language and when describing the same objects of everyday life in two different languages.

The meanings of grammatical categories may also have radial category structure with a prototype as the central member. The Russian prefix *raz-* has a prototypical meaning 'apart' and at least ten other meanings including 'crush', 'soften/dissolve', 'swell', 'spread' and 'metaphorical[13] spread' (Janda & Nesset, 2010, 489). These ten meanings are connected to the prototype and to each other in different ways. As mentioned in 3.2, Russian future tense meanings also form a radial category (Article II, Kosheleva & Janda 2022). The radial category of Russian future is a multilayered structure with meanings as extensions that share submeanings with one another. For example, Gnomic and Extended Future noncentral extensions share Hypothetical submeaning (Article II, Kosheleva & Janda 2022).

3.4 Metaphor and metonymy

What types of relationships can both lexical and grammatical meanings have with each other? Two central concepts that play a major role in the semantic relationships that structure radial categories are metaphor and metonymy. The metaphors of time have already been discussed in connection with the topic of how people perceive time. In this subsection, I turn to metaphor and metonymy as general phenomena.

In the traditional sense, a metaphor is a type of figurative language use based on comparison and the search for common features in seemingly different phenomena.

A well-known example of a metaphor is America as a melting pot. A melting pot literally means a type of container resistant to heating, in which the melting process can be carried out. Zangwill and Nahshon (1908) first compared America to a melting pot where various peoples and their national cultures merged to form a united American nation. A country and a container do not have many features in common; however, the metaphor helps understand the processes taking place in society by offering a specific object-based image.

Metonymy is metaphor's close relative. While metaphor represents mappings across domains, metonymy is about contiguity relationships inside one domain. Metonymy is a mechanism where a part of something or something otherwise associated with something is used to represent the whole object or concept. A good example of metonymy in fiction can be found in the narrative poem The Bronze Horseman: A Petersburg Tale by Pushkin: *vse flagi v gosti budut k nam* ('all flags will visit us'). The flags as symbols of the countries stand for the people from these countries who are coming for a visit.

What is the connection between the terms used in fiction and cognitive linguistics?

Cognitive linguistics recognizes the essential role of metaphorical and metonymic relationships in cognition and language.

---

[13] The role of metaphor and metonymy in cognitive linguistics is discussed in section 3.4.

Metaphor is understood more broadly, as “experiencing one kind of thing in terms of another” (Lakoff & Johnson 1980, 5). These “kinds of things” are called domains. Domains include mental experiences, representational spaces, and concepts (Langacker 1987, 147). The “kind of thing” that people are experiencing belongs to the target domain; and the target domain is understood in the terms of the source domain. (Kövecses 2010, 4). Thus, metaphor is also defined as a mapping (correspondence) from the source to the target domain or “cross-domain mapping in the conceptual system” (Lakoff 1993).

Let us revisit the melting pot metaphor and look at it from the perspective of cognitive linguistics. The source domain is the concept of a melting pot. The concept includes the container, i.e., the melting pot itself, the metals that are put in the pot, the heat that is applied to the pot, the time it takes for the metals to melt, and the resulting alloy. The target domain is the country of the United States. The target domain contains the country, the people who initially lived there and the people who migrated there, their life and their interactions, the time it takes for the people to build strong societal ties, their common children and grandchildren. The mappings (correspondences) from the source domain to the target domain are the following:

- pot ⇒ country;
- metals ⇒ indigenous people and immigrants;
- heat ⇒ interactions in the society;
- melting time ⇒ time;
- alloy ⇒ American people.

The metaphor of America as a melting pot is quite complex. Each of the ingredients in the metaphor is different in both form and function. The pot is a type of container, in which something can be placed. The metals are the material that can be placed in a container. The heat is the energy that is applied to the container and its content. After a period of time, the content of the container changes into an alloy. The result is different from any of the parts. A similar transformation occurs in the metaphorical context. A country as a container allows the material under the influence of the catalyst (the society as a source of interactions) to transform into citizens with a new American identity over time.

Metaphors do not only concern specific concepts from our lives, but also grammar. While discussing linguistic means that express time in subsection 2.1.2, I already mentioned the difference between perfective and imperfective aspect in Russian described by Janda with the help of metaphor (2003). Different properties of rocks and sand correspond to different functions of perfective and imperfective aspect respectively. Let us take a closer look at the metaphor PERFECTIVE IS A DISCRETE SOLID (Janda 2003, 252). Rocks as discrete solid objects have — among others — the

following internal (i.e., inherent) and external (i.e., related to humans and objects in the world) properties mapping onto perfective aspect functions:

- edges ⇒ boundaries (a beginning and an end of an action);
- a variety of shapes ⇒ a variety of durations (duration is relevant);
- countability ⇒ compatibility with measure adverbials (*dva časa* 'two hours');
- inability to share space with other objects ⇒ sequencing with other perfectives.

The use of the perfective non-past form as a default future tense can be also explained with the help of this metaphor: a rock occupies a certain space, and no other rock can be simultaneously put in the same space. The "now" moment of the present tense is already occupied by a solid object, namely the human observer. A rock is unable to share the same space with the observer and is therefore forced to take the next spot. Mapping this property into the perfective aspect domain, the perfective non-past is interpreted as having future meaning (Janda 2003, 254).

Metaphor also plays a role in the relationships between various members of a radial category (cf. subsection 3.3). Metaphor serves to extend the boundaries of meaning in one domain into other domains. Thus, in radial categories the domains of the prototype and the extensions can have metaphorical relationships. In Article II (Kosheleva & Janda 2022), we propose a radial category for the future tense in Russian. In this radial category, there exist identical meanings like Hypothetical in different domains – Extended Future and Gnomic. This is an example of metaphorical mapping from the source domain (Extended Future) to the target domain (Gnomic). Hypotheticals in both domains have the same structure: '*esli* X.PFV.FUT, *to* Y'. Hypothetical meaning in the Extended Future domain is situated in Projected reality close to Non-reality (Langacker 2008, 528). Hypothetical meaning in the Gnomic domain is not grounded in time due to the nature of this domain (thus Gnomic Hypothetical can occupy all spots in Reality). However, since it is connected by mapping to Hypothetical in Extensive Future, Gnomic Hypothetical creates a similar mental space (Fauconnier 1985, Chapter 3).

The two examples of metaphor used in relation to verbal aspect described above show how the concept of metaphor can be applied to the grammatical meanings of tense.

Metonymy in cognitive linguistics is also defined in terms of mapping. However, instead of mapping across domains, as in metaphorical correspondence, metonymy is a mapping inside one domain. Both the vehicle (the source) and the target are situated within the same domain or cognitive model (Radden & Kövecses 2007, 336).

Let us return to the example of Pushkin's flags: *vse flagi v gosti budut k nam* 'all flags will visit us'. It is a complex metonymy consisting of multiple entities. The flags

represent the countries they belong to. The countries, in turn, represent the people who live there. The metonymy is reinforced by anthropomorphizing the flags, which is a special kind of metaphor in itself: the flags are mapped from the domain of objects onto the domain of human beings. In addition, it is rather rare for the representatives of all countries on the planet to gather and visit another country. Here, another metonymy is happening: several countries stand for all countries of the world. The mapping happens inside the domain of the human world.

Like metaphor, metonymy is a type of relationship widespread in language. Metonymic relationships can also occur within a radial category. I use tense and aspect examples again because they are closely related to the topic of this dissertation. Nesset (2009) discusses the role of metonymic relationships for different submeanings of Russian perfective and imperfective verbs. Nesset proposes a radial category for the Russian perfective aspect where the submeanings (*nagljadno-primernoe* 'salient example', *potencialnoe* 'potential', and *summarnoe* 'summarizing') are related to the central *konkretno-faktičeskoe* 'concrete factual' meaning via metonymic extensions (Nesset 2009, 71). To demonstrate a metonymic extension connecting a submeaning to the prototype, let us consider example (27) from (Nesset 2009, 69, cited from Zalizniak & Šmelev 2000).

(27)	On	**reš-it**	ljubuju	zadaču.
	he	solve.PFV.-FUT.3SG	any	problem
	'He will solve any problem.'			

Example (27) represents the potential submeaning, where the pool of potential problems may be represented by one problem. Nesset concludes that this relationship is metonymic because one instance from the time domain represents the whole group of instances within the same domain.

In Article II (Kosheleva & Janda 2022), the submeaning of Stable scenario is also metonymic. Stable scenarios are attested for both perfective and imperfective aspects. Stable scenario belongs to the Gnomic domain. In Stable scenarios, one instance of action represents the whole domain of general truths. Example (28) represents a Gnomic Stable scenario (Article II, Kosheleva & Janda 2022).

(28)	— Dlja	nas,	pčel,	v	skoš-enn-oj	trav-e
	for	us.GEN	bee.GEN.PL	in	cut.PFV-PST.PASS.PTCP-F.LOC.SG	grass-LOC.SG
	prok-u	nikak-ogo.	Nektar	iz	nee	
	use-GEN.SG	none-M.GEN.SG	nectar.ACC.SG	from	she.GEN	
	ne	**voz'm-eš'**,—	prodolža-l-a	star-aja	Pčel-a.	
	not	take.PFV-FUT.2.SG	continue.IPFV-PST-F	old-F-NOM.SG	Bee-NOM.SG	

‘— For us bees, there’s no use in cut grass. You **can't (lit. won’t) take** nectar from it — continued the old Bee.’ [Viktor Kologriv. Medovyj lug // «Murzilka», 2002]

In example (28) the context presupposes that the speaker will be talking in general. Then the speaker uses a non-past perfective verb form to show with one instance ‘you will not do X’ a broader notion ‘you cannot do X’. In other words, one instance stands for a general prohibition.

To sum up, in this subsection I show that metaphor and metonymy play a significant role not only in poetic expressions but also in development of grammatical relationships. Both concepts are highly relevant in explaining the use of Russian tense and aspect.

3.5 Usage-based approach

In subsection 3.1, I briefly touched on one of the pillars of cognitive linguistics: the usage-based approach. In this subsection, I will reflect on what usage-based means in connection with the linguistic phenomena discussed in this dissertation.

As Diessel (2011, 830) points out, usage-based linguists value the importance of “communication, cognition, and processing for the development and organization of the grammar”. In other words, in order to understand how language works, it is necessary to take into account such aspects of human activity as communication, cognition, and processing. The components of grammar change under the influence of the cognitive and communicative load in the process of language use.

Langacker describes an instance of actual language use as a usage event (2008, 220). A usage event includes all the aspects of human interaction between the speaker and the hearer. Linguistic units (i.e., semantic and phonological structures, and the symbolic pairings of those, ibid. 174) arise from usage events via the cognitive mechanisms of schematization and categorization (ibid., 220).

Schematization is a process of abstraction of common patterns extracted from multiple experiences (ibid. 17). An example of schematization is Langacker’s model of tense and potency, which Laura Janda and I adapted for Russian future tense (Article II, Kosheleva & Janda 2022). The original model (Langacker 2008, Fig. 9.17, 306) is a schema consisting of a cylinder and a cone extending forward from left to right. Different areas of the schema represent different realities (conceived, projected, and potential) where humans can place events that they experience. In the Russian future tense adaptation of this schema, we highlight the abstract areas corresponding to one of four major classes of meanings of the future tense (Future, Extended Future, Directive, and Gnomic).

In subsection 3.2, categorization is defined with respect to Croft and Cruse (2004, 74). Langacker’s definition of categorization is “the interpretation of experience with respect to previously existing structures” (2008, 17). Categorization is best illustrated in subsection 3.3 dedicated to radial categories. Here I would like to mention another

less prototypical example of categorization. In Article I about local markedness (Kosheleva & Janda, accepted), we combine the traditional notion of markedness with understanding of the categorical structure from cognitive linguistics to illustrate how markedness should be evaluated on the local level rather than on the level of the category.

In an online lecture on the usage-based approach (2015), Martin Hilpert discusses the principles on which usage-based linguistics is built. These principles are the quintessence of the book by Joan Bybee (2010). In the following paragraphs I discuss the principles, which have proven to be most helpful for this dissertation. Hilpert (2015) calls them "claims", and Diessel (2011) operates with the term "theses".

The ideas behind a few of these claims have already been discussed above. In agreement with Langacker, Bybee believes that "language is grounded in domain-general cognitive processes" (cited from the review by Diessel 2011, 831). Grammatical categories are gradient, consisting of a prototype and extensions. Categories are based on concrete tokens, i.e., categories are exemplar-based (Hilpert 2015). New members of categories are added through the process of analogy rather than rules: if something fits the template, it becomes a member of a category. A grammatical template is a schema (cf. the discussion of the schemas above).

It is also not surprising that linguistic categories are polysemous (cf. on polysemy Langacker 2008, 37): their meanings are determined by their use in different contexts (or in usage events in Langacker's terms). To illustrate the polysemous nature of linguistic categories, Hilpert (2015) uses the example of the English future tense auxiliary *will*. I would like to use an example straight from the dissertation, namely, the future tense markers in Russian, which are also polysemous. Their meanings (e.g., Posterior Future or Gnomic hypothetical) form a radial category (cf. the example in subsection 3.3).

Finally, Bybee (2010) acknowledges the importance of frequency. The frequency effect of prominence can be seen as relevant for the Russian future tense. If a form is produced more frequently, it becomes more entrenched in memory. The frequency ratio of 14 to 1 of the perfective vs. imperfective future tense in Russian that we have discovered in the Russian National Corpus is another argument that allows us to establish the functioning of markedness on the local level (Article I, Kosheleva & Janda, accepted). In addition, Article III (Kosheleva, under resubmission) shows that native speakers rely heavily on frequency when they are facing the choice between the perfective and imperfective future tense forms while non-native speakers are not sensitive to frequency in this matter.

In sum, according to usage-based linguists, understanding language requires considering communication, cognition, and processing. The components of grammar change under the influence of these factors during language use. Schematization and categorization are cognitive mechanisms that produce linguistic units such as semantic and phonological structures. Langacker's model of tense and potency highlights the

abstract areas corresponding to the four major classes of meanings of the Russian future tense. Categories are gradient and exemplar-based, meaning that they consist of a prototype and extensions and are based on concrete tokens. The polysemous nature of linguistic categories is determined by their use in different contexts, as is the case with the future tense markers in Russian. Additionally, frequency plays a crucial role in language use, as seen in the frequency ratio of the perfective vs. imperfective future tense forms in Russian, and we see that native speakers are sensitive to this frequency distribution.

3.6 Markedness

Markedness plays a key role in Article I "Why markedness is always local" (Kosheleva & Janda, accepted) of the dissertation. The article is devoted to the discussion of the phenomenon of markedness in relation to the future tense verb forms in Russian.

Markedness has a long history in linguistics and has been used in a wide variety of ways across many linguistic traditions. The concept of markedness nevertheless lacks a unitary agreed-upon definition. In the following I will go through some of the most prominent approaches to the concept that have been proposed, and then I will provide arguments for why I have chosen to focus on morphological complexity and frequency, rather than expectedness.

The concept of markedness was developed by Trubetzkoy and Jakobson in the 1930s within the framework of the Prague School (Prague Linguistic Circle). Markedness was originally a phonological concept of binary oppositions. A phoneme is considered to be defined by a set of distinctive binary features, where each feature has two values that are not equal: one is marked and the other one is unmarked. For example, consider the phonemes /t/ and /d/. In terms of place of articulation, both /t/ and /d/ are coronal and alveolar: [+coronal, +alveolar]. In terms of manner of articulation, the common features for /t/ and /d/ are [-approximant, -continuant, -nasal, -sonorant, -strident]. The only difference between the phonemes /t/ and /d/ is that /t/ is voiceless and /d/ is voiced. The voiceless feature value in /t/ is unmarked, and the voiced feature value in /d/ is marked. This is motivated by the fact that there is something present for /d/ that is not present for /t/, namely voicing.

Since the term "markedness" appeared, it has undergone changes in interpretation and scope and can now apply to a wide range of phenomena from phonology to semantics and discourse.

Lyons (1977, 305-307) distinguishes three types of relationships between the marked and the unmarked members of an opposition: formal marking, distributional marking, and semantic marking. The formally marked member receives an additional morpheme: 'host' : 'hostess', 'consistent' : 'inconsistent' (Andersen 1989, 13). Hostess and inconsistent are marked members because they have additional morphemes as opposed to unmarked host and consistent. Distributional marking refers

to the reduced size of the pool of contexts the marked member can be used in. In the question “how high are your ceilings?” the word high is unmarked because its meaning is general (the speaker probably expects a number in feet or meters). If one receives a question “how low are your ceilings?”, it probably means that the ceilings are unusually low, and low is marked. In other words, high has a larger pool of contexts in comparison with low. A semantically marked member has a more specific narrow meaning than the unmarked member. As in the famous example with two animal terms in Russian provided by Jakobson (1971[1932]), osel ‘donkey’ is almost deprived of reference to sex, it is the expected member of the opposition and is used as a default when talking about a donkey, be it male or female. Oslica ‘female donkey’ is semantically marked for sex and is used when talking about a female representative of the donkey species (cf. for other criteria in connection with the donkey example Article I, Kosheleva & Janda, accepted). Andersen (1989, 14) calls only the third type of marking markedness, defining it as an asymmetric relationship between contrasting elements.

Battistella (1996, 7) considers that markedness relationships may involve constructions, rules, and features. Markedness is responsible for evaluation and ranking of linguistic forms and may depend on the context in which the form is used and the forms it is being compared to. Battistella (1996, 13-14 and 50-55) also argues that frequency is an important factor for determining which form is perceived to be marked: the more frequent form is perceived to be more natural and simpler, i.e., unmarked. Russian future tense fits well into this model: two aspectual forms (constructions) have very different frequency distributions with the perfective form dominating in the picture (Article I, Kosheleva & Janda, accepted).

Comrie has written extensively on the concept of markedness in linguistic typology. He refers to some linguistic forms as “basic” or “neutral” (Comrie 1989, 85): they are unmarked because they are more expected. Comrie defines expectedness as the degree to which a form is typical or common, while markedness is defined as the degree to which a form deviates from what is typical or expected. Expectedness is an intuitive criterion for markedness but is hard to measure (Article I, Kosheleva & Janda, accepted).

In addition to expectedness, markedness is often associated with complexity (for review of the phenomenon of complexity, see subsection 3.7). Comrie (1983) notes that the marked forms have “more morphological material”. A typical English noun uses an additional morpheme “-s” to indicate the plural form: *cow – cows*, *spoon – spoons*. By the same logic, the English simple future tense is more complex than the English simple present: will see consists of two elements whereas see is just one. Morphological complexity, which may be considered an example of Lyons’ formal marking discussed above, is a measurable criterion that is suitable for empirical investigations of the kind pursued in this dissertation, and markedness as complexity was therefore used in Article I (Kosheleva & Janda accepted).

As can be seen from the examples of the use of the term markedness above, markedness operates on various levels of language: phonological, morphological, semantic. However, it has also been argued that markedness may be "local", i.e., depending on the linguistic or social context where it occurs (Tiersma 1982). The relationship between the singular and plural of Russian nouns illustrate this, as pointed out in Article I (Kosheleva & Janda, accepted). Generally, the plural is marked since it involves the addition of a suffix. Thus, the plural of *stol* 'table' is formed by adding the suffix -y: *stoly*. However, for certain types of nouns, the markedness relationship is reversed. For many names of nationalities, for instance, it is the singular that involves an additional suffix. A case in point is *angličanin* 'Englishman' vs. *angličane* 'Englishmen', where the singular is more complex (and hence marked), since it has the singulative suffix -in, which is absent in the plural. The concept of local markedness applied to the Russian future tense is further explored in Article I (Kosheleva & Janda, accepted). These examples show that a paradigm is not just based on one single contrast. It contains many contrasts at various levels.

In this subsection I reviewed several approaches to markedness applied to various linguistic phenomena. Markedness is indeed a term the understanding of which varies from one linguist (or linguistic school) to another. In the end the only unifying definition is by Andersen, that markedness is an asymmetry between two elements. And this asymmetry proved to be useful when talking about differences between aspectual forms in Russian future tense. Article I (Kosheleva & Janda, accepted) implement two measurable criteria of markedness to Russian verbs on the local level of the future tense: frequency and morphological complexity. The results contradict the conventional understanding that the imperfective aspect in Russian is always marked.

3.7 Complexity

The issue of linguistic complexity is relevant for two of the articles that comprise this dissertation: Article I – "Why markedness is always local" (Kosheleva & Janda, accepted) and Article III – "Russian future: an inside and an outside perspective" (Kosheleva, under resubmission). In the first article, complexity is mentioned as one of the criteria for distinguishing between the marked and the unmarked members of a grammatical opposition. One way to approach this is to note that complexity in terms of form can be directly measured in terms of the number of morphemes involved. The more complex in terms of form the member is, the more likely it is to be marked. The other two criteria suggested by Jakobson (1971[1932]) are expectedness and frequency. The marked member is less expected and less frequent. While frequency is directly measurable, expectedness is not.

In Article III (Kosheleva, under resubmission), I raise the question of complexity indirectly by comparing the performance of native speakers and learners of Russian when they are asked to produce Russian future tense verb forms. Non-native speakers have more difficulties with perfective synthetic single-word forms due to their more complex morphophonemics than with the analytic imperfective two-word forms that

are presumably more complex due to their larger number of morphemes. In this subsection, I describe different approaches to measuring complexity and show how the complexity of future tenses can be considered in accordance with them.

Researchers have various approaches to defining language complexity as a separate parameter for analysis. First, complexity can be absolute and relative. Absolute complexity is an abstract objective measure used to compare (whole) languages. As a rule, this comparison can be performed using calculations of certain metrics (cf. Kolmogorov complexity as a way of measuring the amount of information conveyed in one line). Miestamo (2008) reduces complexity to the number of elements in the system (Berdicevskis 2012). Absolute complexity is an interesting phenomenon in itself, but has probably little to offer in its broad sense for the Russian future tense provided that we do not conduct typological research on the topic. Whatever the absolute complexity of Russian language might be, it remains constant through the three studies conducted in the dissertation and therefore does not provide any further insight into differences in the complexity of the Russian future tense forms.

Relative complexity is connected to cost and difficulty (Dahl 2004, 38-39 in Berdicevskis 2012). It measures how much effort it costs to learn the language. Relative complexity is subjective in nature because it is tied to the experience of a non-native speaker. Kusters (2008) creates a model of such type of a speaker. He proposes a concept of a "generalized outsider": a prospective learner who lacks any cultural background about the relevant language community. The distance between a learner's native language and the target language, the complexity of which is measured, can play a role (Berdicevskis 2012, 104). The target language can be close to the learner's native language, distant, or not related at all.

In Article III, we are dealing with relative complexity. I could speculate that the native language of the participants might have influenced the outcome of the experiment. Since English is native to the largest group of respondents, the Russian imperfective future analytical construction might be used more frequently because it is more similar to the English future tense than the perfective synthetic form. Both the Russian imperfective future form and the English future tense form entail auxiliary verbs. However, our data analysis does not show any differentiation according to the learner's native language that would confirm this explanation.

Both absolute and relative complexity are concerned with whole languages. The concept of complexity can also apply to individual utterances. Dahl (2004, 42-44, in Berdicevskis 2012) distinguishes between systemic complexity (language as a system) and structural complexity (utterance as a structure). In terms of structural complexity, the stimuli in the experiment (Article III) can be considered equally complex: they are simple homogeneous sentences with similar structure.

Furthermore, linguistic complexity can be examined at various levels of language. Phonological complexity refers to structural organization and arrangement of sounds in a language and can be measured in using such metrics as the number of phonemes

or the presence of distinctive features (e.g., tones). An example of a phonological complexity feature is the number of consonants in a language. Maddieson (Chapter 1, WALS, 2013) categorized languages depending on their consonant inventory in a sample of 563 languages. The four categories include: small (6 to 14 consonants), moderately small (15 to 18 consonants), average (19-25), moderately large (26-33), and large (34 or more). The larger the consonant inventory, the more complex the phonological system of the language.

According to Baerman et al. (2015), morphological complexity is the degree to which a language's grammatical and lexical structures exhibit morphological elaboration and/or irregularity. Morphological complexity considers the number and type of morphemes, and the presence and use of inflection, derivation, and compounding. Latin is a prototypical example of a language with high morphological complexity, whereas Chinese can be seen as a morphologically simple language because words in Chinese are normally composed of a single morpheme.

Syntactic complexity is a measure of elaboration and integration of linguistic structures within and across sentences (Givón 2009). Agreement, case marking, flexibility of word order contribute to syntactic complexity. A sentence with a relative clause is more complex than a sentence without one because the presence of a subordinate clause requires the reader to track multiple relationships to understand the meaning of the sentence.

According to Fillmore (1982), semantic complexity refers to the degree of cognitive effort required to understand a linguistic expression in context. It evaluates the intricacies of the conceptual representation of meaning in a language. The average number of words in a sentence, the degree of polysemy and semantic transparency contribute to semantic complexity. Using a word or a concept in its non-prototypical meaning (e.g., 'bank' not as a financial institution but as a piece of land at the edges of the river) rises the level of semantic complexity.

Linguistic complexity is a multifaceted term, which is most often used in the context of the complexity of individual languages. In this subsection, I have reviewed the main types of complexity and shown how complexity can be applied locally to individual linguistic phenomena, in this case, the Russian future tense forms. Within the framework of this dissertation, relative complexity and morphological complexity play an essential role. Relative complexity is relevant for Article III (Kosheleva, under resubmission), where we are talking about non-native speakers and their challenges with the Russian future tense forms. Morphological complexity appears in conjunction with markedness in Article I (Kosheleva & Janda, accepted) as one of the criteria when distinguishing between the marked and the unmarked members of the opposition.

4 Methodology

The usage-based nature of cognitive linguistics encourages the researcher to extensively use many empirical methods available today. In this section, I cover the methods used in the articles written by my co-author and me. These methods include the analysis of the carefully selected corpus data (4.1), conducting experiments (surveys) online (4.2), and implementing statistical models to interpret both corpus and experimental data (4.3).

4.1 Corpus

This subsection discusses the significance of corpora in cognitive linguistics research. The Russian National Corpus (RNC) is introduced as a major corpus used in the analysis of the Russian future tense forms. The overview of other corpora provides context for the reasons behind choosing RNC for the purposes of this dissertation and provides grounds for potential future research.

Corpora are large collections of texts, assembled and organized following certain rules. The texts usually have metadata about their author, time of creation, genre etc. All the words in all sentences in a corpus are also annotated according to their morphological, semantic, and syntactic characteristics. The annotation is often made automatically but certain subcorpora can be annotated or corrected manually to ensure the quality of the annotation often much needed for the linguistic research.

In cognitive linguistics, corpora are widely used as an important source of data produced by speakers, as opposed to examples constructed by linguists themselves. Examples of areas where corpus data can be of help are studies of synonyms and polysemy (Newman 2011, 531-538), and the relationship between form and meaning (Janda 2016). Below I list several studies featuring the use of corpora, which have contributed to moving the research frontier of cognitive linguistics forward.

Studies of linguistic profiles — an approach within cognitive linguistics based on Construction Grammar (Kuznetsova 2012, 21) — would be impossible without the use of corpora. Linguistic profiling includes grammatical, semantic, and constructional profiling (Janda 2016, 131). Grammatical profiling uses corpus data to establish the relationship between the frequency of grammatical forms and grammatical categories (Janda & Lyashevskaya 2011). Semantic profiling aims at figuring out how meanings and forms relate to each other. Janda and Lyashevskaya (2013) compare the distribution of the semantic tags of verbs in corpus data and the distribution of the (supposedly "empty") Russian prefixes *po-*, *s-*, *za-*, *na-*, and *pro-*. Constructional profiling explores all the constructions in the corpus that the lexeme of interest is found in (Kuznetsova 2012, 23-24). Janda & Solovyev (2009) use constructional profiling to establish the actual relationships, i.e., not the ones described in the

dictionaries, between synonyms based on the relative frequencies of the constructions they appear in.

In this dissertation I primarily use examples taken from the Russian National Corpus. The Russian National Corpus (RNC, ruscorpora.ru) is comprised of texts written mainly during the period between the 18th and the 21st century. The RNC was created in 2004 and has been steadily developing since. Currently the RNC consists of several subcorpora including the newspaper subcorpus, several parallel subcorpora such as Russian-English, Russian-French, Russian-Polish and others, a subcorpus with dialectal texts, a poetic subcorpus and so on. Per May 2022, the total size of the main subcorpus of the RNC (in the new version) is 337,025,184 words in 126,901 documents. The search function in the corpus allows the user to get search results for specified parameters, as well as to download a portion of results in tabular format (.csv). The parameters for the search include grammatical features such as the part of speech, case, number, gender, tense, aspect etc. Semantic classification is also available: the user can choose the semantic class of a word, e.g., movement, placement of an object, physical phenomena, location, mental sphere, etc. for a verb.

Altogether, the RNC is a good instrument, which provides enough data for the analysis of perfective future forms. The imperfective future consists of two separate words and thus is hard to count in the corpus. The issues connected to the imperfective future are described in Article I (Kosheleva & Janda, accepted).

The RNC is not the only corpus available for the Russian language: there are several other corpora, and I would like to mention a few. The Helsinki Annotated Corpus (HANCO, h248.it.helsinki.fi/hanco/index_e.html) was developed between 1999–2016 at the University of Helsinki. In HANCO, the priority is the maximum coverage of grammatical information, not the number of words and documents (Reznikova & Kopotev 2003, 34). A different strategy is used in Araneum Russicum Maximum (Benko & Zakharov, 2016). The creators prioritize size, making the largest web corpus for Russian (Kutuzov & Kunilovskaya 2017, 48). As of May 2022, the version of the corpus called Araneum Russicum Maius contained 1,200,001,911 tokens. Another corpus is the General Internet-Corpus of Russian (GICR, webcorpora.ru/en/), which first appeared in 2012. GICR is also quite large (more than 20 billion words) but genre-specific: it consists of texts collected from Russian social media such as Vkontakte and LiveJournal. Finally, a fundamentally different approach is used in the Collocations Colligations Corpora (CoCoCo, cococo.cosyco.ru). As the name suggests, in CoCoCo the focus is on the co-occurrence of words in collocations. It is designed as a learning project helping students to deal with idiomaticity in the Russian language. CoCoCo uses data from other corpora: the RNC and I-Ru (an automatically annotated corpus of 140 million words). Different corpora are suitable for different purposes; but by far the most standard and, therefore, (one can hope) representative remains the RNC.

4.2 Data elicitation methods

Subsection 4.2 focuses on conducting research through experiments. It explains why experiments are necessary and outlines the steps involved in their design and implementation, such as formulating a hypothesis, designing the experiment, selecting materials, and choosing variables.

Working with corpus data can provide a lot of useful information about a particular phenomenon for a linguist. However, a corpus is not always an ideal source of data. As a part of this dissertation, I am studying how Russian future tense forms are used by non-native speakers in Article III (Kosheleva, under resubmission). At the current stage of development, learner corpora permit us to make certain observations.

For instance, Olshevskaya (2018) studied the types of deviations in the use of verb forms, and verbal aspect in particular. Her research is based on the largest corpus of data collected from students studying Russian — the Russian Learner Corpus (web-corpora.net/RLC). Olshevskaya (2018) uses tags for errors provided by automatic annotation of the RLC, which does not always annotate the examples correctly. The analysis generated by the automatic annotator of the RLC is a good place to start investigating learner errors: the RLC provides data from learners of all levels with various mother tongues. However, the amount of noise (erroneous annotation) in this corpus is excessive and the data that can be extracted from this corpus is insufficient to determine what factors motivate errors in L2 production of future tense verb forms.

In addition, it is a challenge to find comparable data produced by native and non-native speakers. With the limited number of texts written by non-native speakers, it can be hard to find similar contexts used by both native speakers and learners. In Article III (Kosheleva, under resubmission), I conducted an online experiment in the form of a survey where I asked the respondents to change the tense of the verb form in a short sentence from past to future. It is important to remember that experiments have their own limitations, and my survey is not an exception. The respondents were presented with single sentences without additional context sentences since otherwise the survey would take too much time to finish, incurring errors due to fatigue. Such decontextualization means that some factors could not be taken into consideration (Talmy 2007, xix).

In this subsection, I would like to mention what is usually taken into account when designing and conducting an experiment. There are several steps included in the process. Formulating a hypothesis is a first step (Arunachalam 2013, 221). In Article III (Kosheleva, under resubmission) I investigated factors that may influence the choice of aspect in the future tense. I suspected that one or several factors would play a role. For example, that both native and non-native speakers would be sensitive to frequency. This turned out not to be true for the non-native speakers.

The second step is the design of the experiment itself. The experiment should be designed before the participants are recruited (Quené 2010, 270). The creators of the experiment should not only decide which research questions they want to be answered by means of an experiment, but also which parameters they will work with. In the established set of parameters, a researcher also needs to understand which of them are independent variables. The values of the independent variables affect the outcome, or the dependent variables (Arunachalam 2013, 222). In the experiment I ran for Article III (Kosheleva, under resubmission) the independent variables are the original aspect in the stimulus, the relative frequency of the verb in the aspectual pair, and the type of the speaker (native and non-native). The dependent variable in this experiment is called "Match". Match is a binary response dependent variable, where "yes" means that the verb form produced by the participant in the response has the same aspect as the verb form in the stimulus, and "no" means that the aspect in the response is different (== does not match) from the aspect of the verb form in the stimulus. Furthermore, the participants and the stimuli can have their own properties that can affect the study; these variables are called random effects.

In addition to establishing the variables, it is important to choose the materials for the study. This has to be done strategically, since if the number of stimuli is too big, the participants will get tired, and even if they finish all the tasks, the results may be less reliable.

After establishing the variables, the researcher has to develop a set of stimuli to test the participants on. Ideally one would extract stimulus examples from a corpus, however for a study that would include both native speakers and L2 learners, such examples would be too complex. Furthermore, for the purpose of this study, I needed parallel examples that differed only in verbal aspect, and no corpus contains such examples. To this end, I constructed 32 pairs of identical examples for perfective and imperfective verbs. Each example was accompanied by an illustration in order to support comprehension and provide some context. The verbs in the stimuli represented a range from highly frequent to relatively infrequent, and the various types of Russian aspectual morphology, namely marking by prefixes, suffixes, and suppletion.

In order to get an independent reading of responses for both perfective and imperfective verbs, it was necessary to split the stimuli into two sets so that in each set each verb pair was represented by only one aspect. Thus, a given participant saw only one aspect and their response with one aspect for a given verb would not affect a response for the opposite aspect. Each stimuli set targeted sixteen unique perfective verbs and sixteen unique imperfective verbs. Each participant was randomly assigned to one of the two stimuli sets; those with birthdays from January to June got one set, and those with birthdays from July to December got the other set. A pseudorandomized order of stimuli was produced for every participant, thus reducing possible ordering effects.

Psycholinguistic experiments often elicit grammaticality judgements and attempt to mask the target phenomenon by including filler stimuli. However, my experiment did not collect grammaticality judgements, and furthermore required the participants to perform a task (change a verb form from past to future), which therefore could not be masked. The task was already rather long, requiring 32 responses, and filler stimuli would have made the experiment too demanding on participants' time. For these reasons, filler stimuli were not included.

All experiments must conform with ethics guidelines, specifically the Declaration of Helsinki. Experimental design must either a) avoid collecting personal identifier information and avoid potentially traumatic impact on participants, or b) go through a review process prior to carrying out the experiment. In this case, no personal identifiers were collected, and the stimuli were neutral sentences that would not trigger trauma. It is also essential that all participants are given information about the experiment, that they can make informed decisions, that they can quit at any time, and that they receive contact details in case they wish to receive more information.

The third step is to recruit participants. In order to recruit native speakers of Russian, I used announcements on social media, which put a limit on the social strata and age of participants. In order to recruit learners of Russian, I used professional networks connected to the Russian Language Center at the Higher School of Economics in Moscow and the Slavic and East European Languages (SEELANGS) listserv, both of which connected us to teachers of Russian as a foreign language and their students. This is a limitation because although Russian is learned all around the world, our participants were largely drawn from populations in Russia, the US, and Europe. Since participation was voluntary, participants were self-selected, and this is also a limitation. The L2 learners necessarily had some linguistic training (language learning). Both the native speaker and L2 participants represent a large range of ages, but overall young women 18-30 are overrepresented in both groups.

Finally, the data that has been gathered should be analyzed and interpreted in an appropriate way. In my study I used a mixed effects model that I describe in the next subsection.

4.3 Statistical methods

The focus of subsection 4.3 on two statistical analyses, the chi-squared test and mixed effects logistic regression, used to investigate the aspectual difference of the Russian future tense forms. The chi-squared test measures the independence of variables in a matrix of values, while mixed effects logistic regression is used to analyze binary outcomes influenced by multiple factors.

The development of information technology in the past thirty years has not passed by the humanities, and linguistics in particular. On the one hand, digital progress has contributed to the accumulation of language data. Recording, processing, and storing of linguistic data has become easier: there is no need to maintain huge file cabinets

when all information is placed on a virtual cloud. Balanced text corpora with convenient markup have appeared (cf. subsection 4.1). In addition, data and methods that were not previously used in linguistics, such as the results of fMRI and eye-tracking studies, have become available.

On the other hand, new tools for analyzing the obtained data have appeared. Using scripts, from short to sophisticated, written in programming languages such as Python or R, one can get statistical analyses and visualize their results.

All these innovations fit well with the usage-based nature of cognitive linguistics and its focus on analysis of language production (cf. subsection 3.5). This can be observed in the sharp rise and after 2008 strong predomination of quantitative studies published in the journal *Cognitive Linguistics* (Janda 2013, 4-5).

In this dissertation, I use two types of statistical analysis: the chi-squared test (Aritlce I, Kosheleva & Janda, accepted) and mixed effects logistic regression (Article III, Kosheleva, under resubmission).

The chi-squared test can also be called “a test of independence” (Levshina 2015, 210). One can measure whether the variables that define a matrix of values are independent of each other. The test evaluates whether the distribution of the observations is uneven, and whether “this unevenness cannot be attributed to chance” (Janda 2013, 10). In Article I (Kosheleva & Janda, accepted), we investigate whether the number of the imperfective verb forms is significantly bigger than the number of the perfective verb forms in the Russian National Corpus, i.e., the imperfective aspect is more frequent than the perfective. We found out that out of all verb forms, 55% were attributed as “perfective” whereas 45% of the forms were annotated as “imperfective”. We compared the numbers of perfective and imperfective against the total number of verbs using the chi-squared test and found out that the difference is statistically significant while the effect size is quite small. This result is consistent with the traditional assumption that imperfective is unmarked, while perfective is marked. We further applied the chi-squared test to show that differences in the distributions of perfective vs. imperfective verbs forms indicate more nuanced markedness differences at the level of tense.

Another type of statistical analysis that I used is called mixed effects logistic regression. As follows from the name, it is a mixed effect model based on logistic regression. Logistic regression is a method for analyzing data where you expect a binary outcome, and this outcome is influenced by more than one factor. In Article III (Kosheleva, under resubmission) the binary outcome is “yes” or “no” for the variable “match” (whether the aspect in the response is the same as the aspect in the stimulus). The fixed effects factors include the aspect of the stimulus, relative frequency of the aspectual forms in a corpus, and whether the respondent was a native speaker or a learner of Russian. In logistic regression, it is possible to measure the effect of all factors at once, both as main effects and taking into consideration their interaction (Janda 2013, 21). Mixed effects models are used for working with datapoints that are

not independent of each other (Gries 2021, 424), making it possible to combine fixed effects with random effects in a single model. In the data for Article III (Kosheleva, under resubmission) there are two random effects which can affect the datapoints making them connected with each other: the participants (coded as “ID_individual_c”) and the stimuli (coded as “ID_stimulus_c”). These are random effects because a given participant will have their own tendencies in responding, and because a given stimulus will also evoke a pattern of responses. A mixed effects model makes it possible to detect the influence of the fixed factors – in this case aspect, relative frequency, and status as native speaker vs. learner – as separate from the random effects. All in all, mixed effects logistic regression is a solid model for the analysis of the data in my experiment, showing that all three fixed effects do indeed influence outcomes.

5 Overview of the articles

This section presents the brief summaries of the procedures that were carried out and the resulting findings in the articles in this dissertation.

5.1 Why markedness is always local: the case of Russian aspect

The first article tackles the notion of markedness in relation to aspectual opposition between perfective and imperfective aspect in Russian. My co-author Laura Janda and I show that aspect is marked not merely on the level of the verb but also more locally on the level of tense. This article also opens the journey into the landscape of corpus data attesting the use of the Russian future tense.

We understand markedness as a scalar term which serves for "encoding asymmetries" (Diessel 2019, Chapter 11), such as that between the usually unmarked singular and usually marked plural. Markedness fits well into the framework of cognitive linguistics, especially in the model of the radial category (cf. subsection 3.3), with the central prototype (the unmarked member) and the peripheral extensions (the marked member).

In terms of expectedness, complexity, and frequency, the prototypical unmarked member is more expected, less complex, and more frequent (cf. Table 1, Article I). Respectively, the marked members of the radial category are less expected, more complex, and less frequent. To "measure" the markedness of the aspect, we operate with morphological complexity and corpus frequency, and we take expectedness into account where possible.

Dahl (1985) states that in Russian, markedness is "working" on a categorical level, i.e., in the aspectual category, where perfective aspect is marked, and imperfective aspect is unmarked. In the article, we show that Russian aspect can be more optimally described in terms of local markedness. Tiersma (1982) points out that different lexical items can behave differently in terms of markedness. For example, in Russian, for most nouns, the singular is unmarked, and the plural is marked, and this is reflected in the formal marking, as we see in the forms *arab* 'Arab' and *araby* 'Arabs', where the plural has the ending *-y* that is absent in the singular. However, some nouns are marked instead in the singular, with an unmarked form in the plural. We see this in the forms *angličanin* 'Englishman' and *angličane* 'Englishmen', where the singular is marked with a suffix *-in* absent in the plural form.

In our article, we study the phenomenon of markedness on different levels (tense and aspectual morphology).

If the perfective aspect in Russian is more marked, we would expect it to be less frequent than the imperfective aspect. We found that in the Russian national corpus (the RNC) approx. 45% of the verb forms are perfective and 55% of the verb forms are

imperfective. The 10% difference is statistically significant, although the effect size is quite low (cf. footnote Article I, 9).

Morphologically, there are two ways of forming aspectual pairs of verbs in Russian which are used for most verbs. We call them pattern "A" and pattern "B". In pattern "A", the perfective is a prefixed version of the imperfective. For example, the pair consisting of the imperfective *vjazat'* 'knit.IPFV' and the prefixed perfective *s-vjazat'* 'knit.PFV' belongs to pattern "A". The imperfective verb in pattern "A" is called simplex because it does not bear any additional aspectual morphology. On the other hand, the perfective verb in pattern "A" has a prefix and thus is more morphologically complex than the imperfective.

In pattern "B", it's the imperfective verb that gets additional morphology. An example of pattern "B" is the aspectual pair *pri-vjazat'* 'tie.PFV'/ *pri-vjaz-yva-t'* 'tie.IPFV', where the secondary imperfective verb 'tie.IPFV' is formed by adding a suffix *-yva-* to the perfective verb *pri-vjazat'* 'tie.PFV'. In pattern "B", the imperfective verb is more morphologically complex than the perfective.

We reaggregated the data regarding the total frequency and the frequency of the past tense forms (where the aspectual distinction is easily spotted) of verbs belonging to patterns "A" and "B" from (Janda & Lyashevskaya 2011). We found out that in pattern "A", the imperfective verbs are more than twice as frequent as the perfective verbs. The situation for pattern "B" is the opposite: the perfective verbs are more frequent than the imperfective. So, the aspectual markedness for patterns "A" and "B" is local. In pattern "A", the marked member is the perfective, with higher morphological complexity (a prefix) and lower frequency. In pattern "B", the marked member is the imperfective, with higher morphological complexity (a suffix) and lower frequency.

In addition to patterns "A" and "B", there are few verbs that belong to pattern "C". In pattern "C", the simplex verbs are perfectives denoting accomplishments, and imperfectives are derived therefrom by means of suffixes. In the article, we consider the data for four such verbs: *dat'* 'give.PFV' / *da-va-t'* 'give.IPFV', *vstat'* 'stand.up.PFV' / *vsta-va-t'* 'stand.up.IPFV', *rešit'* 'decide.PFV' / *reš-a-t'* decide.IPFV', and *polučit'* 'receive.PFV' / *poluč-a-t'* 'receive.IPFV'. For the verbs in pattern "C", the perfective may be perceived as more expected since the situations described by the perfective verbs from the pattern "C" are "momentary, completed, and unique". Together with morphological complexity and frequency, expectedness points to the perfective verb being the unmarked member of the aspectual pair in pattern "C".

Taking the idea of local markedness one step further, we examined a more challenging part of the verbal paradigm, namely the future tense. For this purpose, we needed to accurately calculate the frequencies of the future tense forms (the synthetic non-past perfective and the analytical imperfective) in the RNC. The major challenges for calculations are homonymy and non-contiguity of periphrastic forms.

There are three sources of homonymy that present challenges for measuring frequency. First, for many perfective verbs, the future second person plural form is homonymous with the second person plural imperative form, e.g., *pogovor-ite* can have two analyses: speak.PFV-FUT.2.PL and speak.PFV-IMP.2.PL. The second source of homonymy is biaspectual verbs, which express different aspect depending on the context. An example from the article is the verb *operirovat'* 'operate', the non-past forms of which, e.g., *operiru-et* can be interpreted as perfective future operate.PFV-FUT.3.SG, or imperfective present operate.IPFV-PRS.3.SG depending on the context. The third source of homonymy is a group of verbs of motion, where homonymous forms can have both different aspect and meaning. For instance, *s-xož-u* is either roundtrip-walk.PFV-FUT.1.SG 'I will go somewhere and come back', or down-walk.IPFV-PRS.1SG 'I am descending'.

In addition to the homonymy issue of the perfective future, the imperfective periphrastic future has its own problem, which originates from its form. The imperfective future consists of an auxiliary verb and an infinitive (e.g., *bud-et spraš̆iva-t'* 'be.FUT-3.SG ask.IPFV-INF' 'will ask'), which can be separated by other words, and/or these two words can appear in different order. Modal constructions with *možno* can also contain the future auxiliary and the imperfective infinitive and give a false impression of an imperfective future tense. Such issues present challenges for counting the number of periphrastic future forms in a corpus.

Due to the issues described above, we decided to count examples manually, and chose to focus on a sample of ten high-frequency aspectual verb pairs. The pairs belong to different morphological patterns ("A", "B", "C", and a suppletive pair) and have a certain degree of semantic variety. The frequencies from the RNC are presented in Table 5 (Article I, Kosheleva & Janda, accepted). The biggest challenge was to calculate the estimated numbers for the imperfective future forms, which do not have their own "label" in the corpus. For this purpose, we developed a special procedure. We extracted samples of infinitives of 100 attestations each. Then we counted how many of these infinitives were a part of the periphrastic future forms. After that, we extrapolated the received numbers to the total numbers of these infinitives in the RNC.

The ratio of perfective to imperfective future forms for the ten verb pairs was calculated to be 14:1. In other words, perfective future forms vastly outnumber imperfective forms.

The obtained results align well with the morphological complexity of the future tense forms. The perfective future is less morphologically complex (lacking an auxiliary verb), and much more frequent than the imperfective, and thus is unmarked. The imperfective future is the marked member, both in terms of high morphological complexity and low frequency. This conclusion goes in opposition with the traditional interpretation of the imperfective aspect as unmarked in Russian and gives evidence for broader use of the notion of "local markedness".

The dominance of the perfective over the imperfective future forms can also have consequences for both native and non-native speakers. This issue is investigated in Article III (Kosheleva, under resubmission).

5.2 Looking into the Russian future

In the second article, together with Laura Janda, we explore the meanings of the future tense forms. Having analyzed corpus data for both perfective and imperfective future forms, we propose radial categories of future and non-future meanings of the future tense.

Both perfective and imperfective futures have been reported to express meanings that are not completely or not always related to future time (Forsyth 1970, Isačenko 1965/2003, Maslov 1990/2004, Stojnova 2016a, 2016b, Vinogradov 1947). Perfective future is in an even more special position, since morphologically it is the perfective present: perfective future and imperfective present have the same conjugation (see Table 2 in subsection 2.3.2).

In order to discover the quantitative distribution of different meanings, we extracted the data from the RNC and prepared two datasets for perfective and imperfective futures respectively with 1,000 examples each. Then we analyzed the data and discovered certain groups of meanings, which form a radial category. We reproduce the diagrams from the article for a better understanding of the description that follows below.

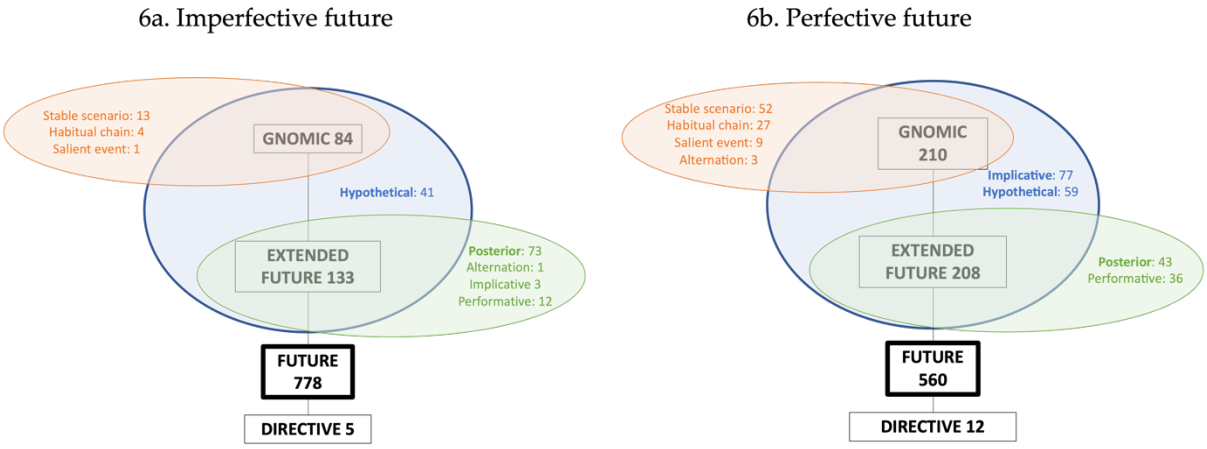

Figure 1. *Radial category of meanings expressed by imperfective Russian future forms* (Figures 6a and 6b reproduced from Article II, Kosheleva & Janda 2022).

The biggest group of examples in both perfective and imperfective future datasets is the future time meaning. There are 560 examples in the perfective dataset and 778 examples in the imperfective dataset with the "pure" Future meaning. Future is the prototypical meaning in the radial category of the future tense. The meaning that is

most closely related to the Future is the Extended Future with 208 examples in the perfective dataset and 133 examples in the imperfective dataset. Extended Future either indicates greater lack of certainty about an action's completion or ties the event to the present or past tense (extending the future). Further away are Gnomic uses: there are 210 examples of Gnomics in the perfective dataset and 84 examples in the imperfective. Gnomic examples constitute a substantial proportion of both datasets; they also differ qualitatively. The future tense forms in Gnomic examples refer to situations that are not grounded in time. Finally, a small group of examples are gathered under the name Directive, with 12 perfective and five imperfective examples. Directives are instructions about actions that should be performed either simultaneously with the instruction or shortly after. Directives are very close to the Future meaning and to imperatives.

In addition to the major meanings, there are minor additional extensions relevant for Extended Future and Gnomic. The extensions specify the general meaning in the group. Several extensions are common for both Extended Future and Gnomic, and both perfective and imperfective futures, others are more unique. In the Extended Future, the common extensions include Posterior (43 perfective and 73 imperfective examples) and Performative (36 perfective and 12 imperfective examples). Posterior Future describes events that begin prior to the moment of speaking; as a rule, the future tense form is used in the subordinate clause, whereas the main clause has a past tense form. Performatives include both traditionally understood illocutionary acts, performed by uttering a phrase, and near-performative actions executed a moment before or after another action.

In the imperfective future dataset, the Extended future also includes one Alternation (*čem*…, *tem*…) and three Implicatives. Perfective Implicatives (77 examples) can be either Extended Future (72 examples) or Gnomic (5 examples). Implicatives are verbs like *smoč'* 'manage' that contribute an "additional layer of meaning" helping to facilitate the action expressed by an infinitive (Karttunen 1971). Another extension common for Gnomic and Extended Future is Hypothetical (59 perfective and 41 imperfective examples). Hypothetical sentences usually contain a space-builder *esli* 'if', which creates an even "less real" space.

The rest of the extensions belong to the Gnomic domain. Stable scenarios (52 perfectives and 13 imperfectives), Habitual chains (27 perfectives and four imperfectives), and Salient events (nine perfectives and one imperfective) are common for both perfective and imperfective datasets. Stable scenarios are utterances based on the experience of the speaker ("general truths" valid for an individual). Habitual chains are sequences of events potentially happening repeatedly (Dickey 2000, Bondarko 1971). A Salient event stands out as a sudden action in contrast with the daily routine (Dickey 2000, Bondarko 1971). In the perfective dataset, there are also three examples with an Alternation extension.

The main groups of meanings of the Russian future tense and their extensions constitute a radial category which has common parallels with the model of tense and

potency proposed by Langacker (2008). Figure 2a-d reproduced from Article II (Kosheleva & Janda 2022) represents the relationship between the various realms of Reality and Non-Realities and the Russian future tense.

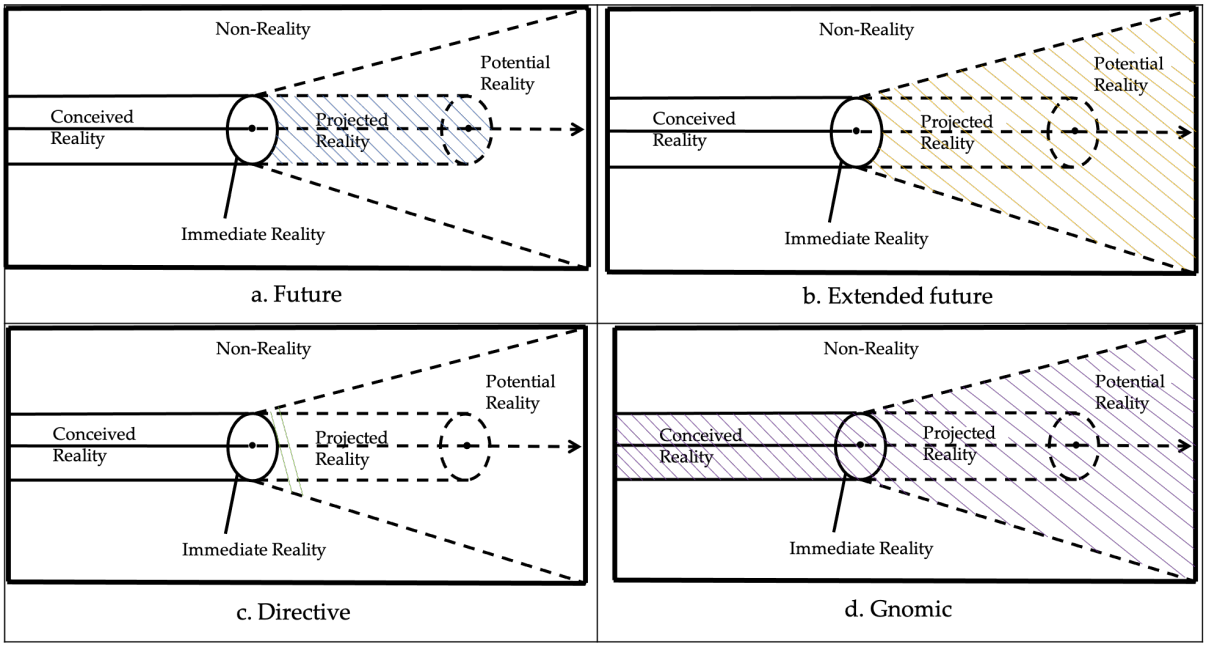

Figure 2a-d. *Adaptation of Langacker's model for Russian future tense. Source: Adapted from (Langacker 2008).*

(Figures 2a-d reproduced from Article II, Kosheleva & Janda 2022).

In Figure 2a, Future corresponds to the area shaded in blue labeled Projected Reality. Extended Future (Figure 2b) is shown with yellow shading and corresponds to the union of Projected and Potential Reality. Directive (Figure 2c) is a small portion of Potential and Projected Realities shaded in green. Finally, shaded in violet in Figure 2d, Gnomic does not have any time limit and so takes up the whole space of reality: Conceived, Projected and Potential Realities.

In addition, we explored the relationship between the future tense meanings and types of modality. To do so, we extracted examples presented in articles with different approaches to modality: Klimonow (2011), Petruxina & Li (2015), Radbil (2011), and Wiemer et al. (2020). Then we analyzed these examples according to our classification of the future tense meanings. We compared the obtained modality-tense correspondences and obtained the following results. There are some examples that we classify as having Future time reference (the prototypical meaning), but are classified primarily as modal uses by Wiemer et. al. (2020) and Klimonow (2011) who view their meaning as possibly potential, and Radbil (2011) who marks it as "undifferentiated modality". Examples that we classify as Habitual chains, Salient events, and Stable scenarios are associated with potential and circumstantial modalities attested by Klimonow (2011), Petruxina & Li (2015), and Wiemer et al. (2020). Examples that we classify as Extended Future are scattered across various types of

modality in these works. No distinct pattern is observed: Extended Future examples range from volitive (Future Performative) to having no modality assigned by the authors of the abovementioned articles (Posterior, Hypothetical, and Performative).

5.3 Russian future: an inside and an outside perspective

The third article in the dissertation aims to broaden our understanding of the acquisition and usage of the Russian future tense by including non-native data. I am particularly interested in comparing the preferences of the native and non-native speakers of Russian when it comes to choosing between the synthetic perfective non-past form and the analytical imperfective future form. Further discussion includes an overview of the particular choices of the speaker and the analysis of the factors that can influence their decision.

In the Article I about markedness (Kosheleva & Janda, accepted), we have already shown that the distribution of the future forms in the Russian National Corpus (RNC) is skewed towards the perfective aspect. The ratio of the perfective to imperfective future forms in the corpus is fourteen to one. Non-native speakers have been reported to rely heavily on the imperfective future form (Swan 2017, 825).

In order to elicit comparable data on the future verb forms in Russian, I conducted a survey experiment targeted at both native and non-native audiences. The participants were charged with one task: replace the past tense forms in the stimuli with the future tense forms of the corresponding verbs. Every participant got a set of 32 sentences, which were shown to them in random order. Every sentence was illustrated by a picture depicting the action presented in the sentence. The set of sentences was accompanied by simple instructions with an example. In addition, I also collected general demographic information from the participants.

I received responses from 160 people: 78 are native speakers of Russian and 82 have been learning Russian as a foreign language. Heritage speakers were excluded. The biggest group of participants are young women (37 natives and 38 non-natives, age range 18—30). The native languages of the learners of Russian range from English (30 individuals) to Swahili (1 individual). They received various amounts of training in Russian, with the most common length of time being four years.

Having collected and manually prepared the data, I undertook a statistical analysis. The aim of the analysis is to find out which factors are associated with the choice of future tense forms. Janda et al. (2019, 270) showed that native speakers of Russian depend on the relative frequency of the aspectual verb forms. I decided to find out whether my data confirms this discovery for native speakers and how non-native speakers behave in relation to the frequency.

I built a mixed effects logistic regression model according to the procedure described in Gries (2021, Chapter 6.4). The response variable Match compares the original aspect of the stimulus with the aspect of the verb form in the response. Match receives

a “yes” if the aspect is the same, and “no” if the aspect of the form given by the respondent is different from the aspect in the stimulus. There are three fixed effect predictors: the original aspect of the stimulus (categorical predictor), the relative frequency of the verb (a numerical predictor; for more on the exact calculations see Article III, Kosheleva under resubmission), and the type of speaker (native vs. non-native, a categorical predictor). In addition, there are two random effects: the identification of the speaker and the identification of the stimulus.

The model shows that all of the fixed effect predictors and the interaction between the relative frequency and the type of speaker significantly contribute to the choice of aspect in the future tense. When the aspect in the stimulus is perfective and the relative frequency of the original aspect is higher than the frequency of the opposite aspect, the prediction of “yes” for the variable Match increases for the native speakers. Thus, my study confirms the observations of Janda et al. (2019). However, non-native speakers are less likely to be successful in establishing the “right” aspect in the answer and the relative frequency does not help them.

In addition to statistical analysis, I carried out qualitative analysis of various deviations that were not directly connected to the choice of aspect. There are several deviations common for both native and non-native speakers and an additional number of deviations specific to the non-native respondents. The most common deviations include the use of verbs synonymous with the verbs in the stimuli (e.g., *žit’* instead of *prožit’* both meaning ‘live’), copying the past tense form from the stimulus into the response, and spelling errors and typos of various kinds. The non-native speakers exhibit the following types of the deviations: the use of the present tense instead of the future (*proživaet* ‘lives’ instead of *prožívet* ‘will live’), combining the future tense auxiliary with the past tense form (**budet snjala* ‘will rented’) or the present tense form (**budet otgrebaet* ‘will shovel away’), and one of the most popular — the use of the future auxiliary together with the perfective infinitive (**budet umyt’* ‘will wash.PFV’).

Summing up, the experiment shows that native and non-native speakers behave differently when it comes to the use of the future tense. When challenged with the specific task of changing the tense of one verb in a short sentence, native speakers favor the perfective future. Even though the perfective future can express additional non-future meanings (Article II, Kosheleva & Janda 2022), it is still the default way of expressing the future time actions and events. The perfective future is also more frequent than the imperfective future (Article I, Kosheleva & Janda, accepted), and as we see in the experiment, when it comes to choosing the aspectual form in the future tense, the frequency is an important factor for the native speakers. Non-native speakers are in a less comfortable position: the morphological form that is simpler from the perspective of a native speaker — the perfective future — remains very challenging for them.

6 Discussion

6.1 Findings and implications

In this dissertation, I have investigated various properties of the future tense verb forms in Russian from the perspective of cognitive linguistics. In this section, I describe the major findings of the research I have conducted and its limitations.

Article I (Kosheleva & Janda, accepted) deals with the phenomenon of markedness and its connection to Russian aspect and future tense. Traditionally, it is believed that the Russian imperfective aspect is unmarked, and the perfective aspect is marked. Based on the criteria of expectedness, complexity, and frequency (Article I, Kosheleva & Janda, accepted, Table 1), we show that markedness does not always operate on the level of the verb. Markedness is also a local phenomenon that operates on the level of the subparadigm, i.e., tense.

My co-author and I developed an algorithm in order to measure the number of uses of perfective and imperfective forms in the Russian National Corpus. The ratio of the attestation of perfective to imperfective future forms in the RNC is 14:1, and this is an important number as it shows that the default way of expressing the future meaning in Russian is via perfective non-past. However, we should also keep in mind that data from other corpora and genres can vary and would need to be investigated separately.

The future meaning is not the only one for the perfective non-past form, as indeed, for the analytical imperfective form of the future tense. Article II (Kosheleva & Janda 2022) is devoted to the qualitative and quantitative analysis of perfective and imperfective forms of the future tense used in context. In some cases, we had to look at the extended context (of the size of a paragraph) to determine the meaning of the forms. This shows that the constituents of a form's meaning can be relatively "distant" from the form itself in the text.

Both perfective and imperfective future forms are polysemous. The future time reference meaning is the most common default meaning for the future tense, also defined in terms of cognitive linguistics as the prototype. In our pseudorandomized samples from the RNC, the examples with future time reference meaning constitute 56% of perfective and 78% of imperfective future tense forms. When we take into account the 14:1 ratio of perfective vs. forms, this means that we can expect 784 (= 56 * 14) perfective forms that express future time reference, as opposed to 78 (= 78 * 1) imperfective forms that express future time reference. In other words, there remains a ten-to-one ratio between the expression of future time reference by means of perfective vs. imperfective verb forms.

Other meanings of the future tense forms include Extended future, Directive, and Gnomic with several additions, such as Hypothetical or Alternation. These meanings have been described in the literature before (cf. Isačenko 1965/2003, Maslov 1990/2004, Vinogradov 1947). We show that these meanings form a radial category

with the future time reference as a prototype and other meanings as extensions gathered around it. In the article, my co-author and I use the model of tense developed by Langacker (2008, 301) and we demonstrate how it can be applied to the Russian future tense.

Future tense verb forms can express modal meanings. Scholars possess a variety of opinions (cf. Klimonow (2011), Petruxina & Li (2015), Radbil (2011), and Wiemer et al. (2020)) about what types of modality can be expressed using the future tense. Having annotated the examples from the abovementioned articles according to our classification, we compared the types of (non-)future meanings with the types of modalities exhibited in these examples. We found no clear pattern of correlation between the two. The non-future meanings cannot be entirely explained via modal attenuations. One type of modality can correspond to several types of future tense meanings and vice versa, one type of future tense meaning can be combined with several types of modality. The question of the interaction of aspectual, temporal and modal meanings in contexts with the use of future tense forms requires further extensive research. Since the types of modality and types of (non-)future meanings combine with each other relatively freely, we cannot assume that the modal component of the future is the sole driver of non-future meanings. We cannot simply blame non-future meanings of future tense forms on modality. Habitual experiences of reality and the sequencing of events also play a role in this complex picture.

In Article III (Kosheleva, under resubmission), I expanded the pool of speakers under consideration and compared the behavior of native and non-native Russian speakers in their use of the future tense. In the experiment, both native and non-native speakers struggled to preserve the original aspect while fulfilling the task of converting past tense forms into future tense forms.

Relative frequency refers to the frequency of the perfective vs. imperfective forms of the verb. One might expect that the aspectual form that is of higher frequency might have a greater chance to be used in the response, and this is an important factor for the native speakers. Non-native speakers are not clearly motivated by frequency in their responses. Sensitivity to the relative frequency brings us back to the discussion of markedness since frequency is one of the three criteria for establishing the marked and the unmarked members of the opposition. The perfective verb is more frequent in twenty-six out of the thirty-two aspectual verb pairs in the stimuli. Considering that a more frequent member is more likely to be found in the responses of the native speakers, one can at least speculate that the experiment is a confirmation that in the future tense, the perfective verb is the unmarked member. However, more thorough research on a different set of verbs is required to confirm or reject this hypothesis.

Non-native speakers may have a different picture of what is marked in the future tense. One of the most notorious mistakes confirmed by the experiment is the overuse of the *budu* auxiliary with both perfective and imperfective aspect. The non-native speakers prefer a form that is more complex from the native speaker's point of view, but is easier to use for learners because it allows them to avoid managing multiple

conjugation paradigms. The overuse of *budu* causes the imperfective form to become the unmarked member for L2 learners. As mentioned above, the non-native speakers are not sensitive to the relative frequency due to the lack of lifelong input, which is impossible for learners to obtain. Another factor might be the traditional teaching strategy, where the imperfective future is taught first or the imperfective and the perfective future are given simultaneously, giving the false impression that the imperfective future is equal to or greater in prominence than the perfective future. To confirm or reject this assumption, one could conduct an in-classroom experiment with two groups of students. In one group, the students could continue to learn the language using the traditional pattern, while in the other group the focus could be predominantly or exclusively on the use of future perfective forms. If the hypothesis is confirmed, this may lead to better quality of teaching the future tense to the future students of Russian.

6.2 Connections between the articles

This subsection is devoted to interrelations among the three articles in the dissertation. The main unifying theme is the Russian future tense, and all further questions arise from particular characteristics of this grammatical category. The three subthemes in the articles revolve around the Russian future tense. Figure 3 demonstrates the structure of the dissertation: the three articles and their topics form the corners of a triangle. The sides of the triangle represent the connections between the articles in the corners of the triangle. The overall theoretical frame is supported by corpus evidence and tested on language users.

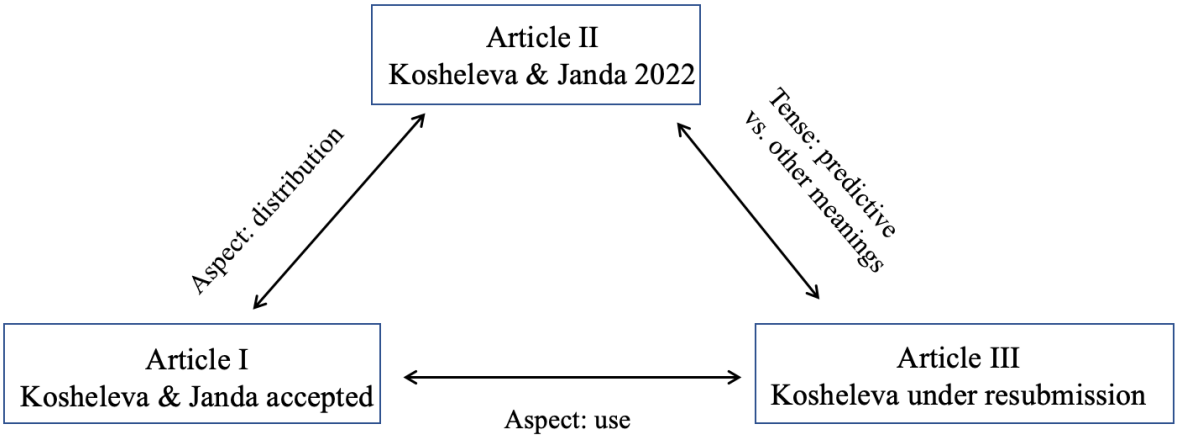

Figure 3. *The cornerstones of this dissertation about the Russian future tense.*

The central theoretical Article II (Kosheleva & Janda 2022) revolves around the semantics of the future tense: what does the Russian future tense mean? And what happens when we disaggregate according to aspect for both perfective and imperfective forms? Of course, various meanings of the future tense forms have been previously discussed in the scholarly literature. In Article II (Kosheleva & Janda 2022) we synthesize the accumulated knowledge about semantic nuances with our own

corpus-based study, and we propose a radial category which comports well with Langacker's (2008, 300) model of tense. Extensive work with corpus data enables us to not only present the radial category but also to visualize the relative scale of the non-future uses of perfective and imperfective future tenses.

We further extend our corpus analysis by looking at the aspectual distribution of future forms in Russian in Article I (Kosheleva & Janda, accepted). A number of factors make the measurement of this distribution far from straightforward. We develop a method that allows us to calculate the ratio of perfective to imperfective future forms in Russian, which according to our calculations turns out to be 14 perfective future forms to 1 imperfective future form. This ratio facilitates a more precise of the interaction of tense and aspect in Russian. Now we know not only what kinds of meanings can be expressed by both perfective and imperfective future forms, but we can also estimate, for example, the ratio of gnomic perfectives to gnomic imperfectives.

In this dissertation I have also considered the experiences of language users, both native and non-native, and the extent to which they use perfective and imperfective future forms (Article III, Kosheleva, under resubmission). Ideally, I would have liked to know how native and non-native speakers use Russian future tense forms with different meanings. In practice, this is a problematic task because semantic nuances are not available for learners until they reach an extremely high level of language proficiency. Therefore, it was not possible to address the full palette of future tense meanings explored in Article II (Kosheleva & Janda 2022). Instead, the experiment was limited to the predictive meaning of future tense; in other words, future time reference in the strict sense. The experiment presented in Article III (Kosheleva, under resubmission) is also connected to the left corner of the triangle, since it shows that native speakers use frequency as a factor when they use future tense. Non-native speakers struggle with the morphophonemics and confirm the well-known rule of Russian language pedagogy: practice makes perfect(ive).

6.3 Findings: common grounds

While subsection 6.2 explains the connections between the articles, this subsection highlights the shared contributions in terms of the theoretical framework, new knowledge regarding the Russian language, quantitative data, and practical implications.

The dissertation is grounded in the paradigm of cognitive linguistics, with Langacker's model of tense (2008) serving as the basis for the theoretical extensions of different types of meanings in a radial category of the future tense. These definitions can be employed in future research on tense in other languages. Tiersma's concept of local markedness (1982) suits the explanation of the discrepancies in the behavior of verbs with different morphological patterns. Our comprehension of markedness based on the data and analysis in Article I (Kosheleva & Janda, accepted) contributes to the development of the usage-based framework of cognitive linguistics (Diessel 2019).

Further on, we have confirmation of the usage-based approach described in (Janda 2019): native speakers tend to use words and structures (in our case, aspectual forms) that they have encountered frequently, rather than relying on a set of strict rules.

The language in focus of this dissertation is Russian. The topic addressed is the grammar of Russian language, specifically the future tense. The findings concerning the topic range from the untangling of the local markedness of "A", "B" and "C" types of aspectually paired verbs to showing the links between the various uses of the future tense. Despite quite a substantial amount of accumulated knowledge about Russian, it seems there are still areas where, especially with the use of the up-to-date methods, it is possible to uncover tendencies and patterns.

Quantitative methods, the description and analysis of data run like a red thread through all three articles. Thanks to the use of corpus data, we have learned a little more about the Russian future, how much and what kinds of future tense expressions there are. In addition, the analysis of corpus data made it possible for us to discover that perfective future tense expresses a greater variety of meanings. Corpus information about frequency helped establish the role of markedness for the perfective and imperfective aspect in the future tense. Furthermore, relative frequency proves to be a significant indicator for the native speakers in choosing between perfective and imperfective aspect.

The results acquired possess practical implications. The relative ratio between perfective and imperfective future tense has been previously underestimated. The subsequent subsection expounds on the potential utilization of the findings obtained.

6.4 Persisting issues and future research

The first article contains a quantitative assessment of the ratio between perfective and imperfective future forms as attested in the Russian National Corpus. Our calculation is 1:14, i.e., there is one imperfective future form for every fourteen perfective forms. The question that remains to be answered is how representative the RNC is for the forms in question. In other words, how often do people write about the future vs. how often do people talk about future? It would be great to work with spoken corpora, to find out the answer for spoken data and compare it to the classical written RNC.

In the second article we sampled corpus data for the perfective and the imperfective future tenses and then analyzed the meanings of the verb forms in the given contexts. We found out that 44 % of the perfective future forms and 22% of the imperfective future forms express other meanings than future. Some of these meanings are atemporal, or gnomic as we call it in Article II (Kosheleva & Janda 2022). It would also be useful to know how many gnomic uses can be found in other tenses: present (as a traditional source for the gnomic meaning), and especially past tense in both perfective and imperfective aspect.

The question raised in the third article that has not been answered yet is simultaneously both simple and hard: how to help non-native speakers to achieve native-like proficiency in their use of the future tense in Russian. There are several ways to tackle the issue.

As shown in subsection 2.3.7, textbooks and other resources often dedicate significant space to the imperfective future. The imperfective future is traditionally introduced first, though in some cases, the imperfective and perfective futures are presented together in a table at the beginning of a chapter or a lesson. Future research on classroom instruction can investigate deliberate shifting of the focus from the seemingly easier (for the learners) imperfective form to the more frequent perfective form.

The aversion of non-native learners to conjugation is justified. Russian is notoriously famous for its rich paradigms, especially when it comes to verbs. Mastering the conjugation takes time, and time is a valuable resource that is always limited. A solution that can be implemented is to create a version of SMARTool that is focused exclusively on the use of perfective future tense forms. The existing SMARTool is a resource for learners of Russian that provides three most frequent forms for 3,000 lexemes divided into groups according to their CEFR level (Janda 2019, 183). Each form is illustrated with an (adapted) example from the RNC. A SMARTool for perfective future forms supplied with a pool of exercises can help learners shift focus from large paradigms to practical solutions, which can boost their ability to use the "correct" future tense.

7 Conclusion

In writing this dissertation, I have come to appreciate the fact that there are two semantic sides to the Russian future tense forms. One side expresses future meaning, while the other expresses a range of other meanings. Especially for the perfective aspect, these non-future meanings are prominent and important. These meanings are connected also to the present tense and to modality in complex ways, and these comport well with Langacker's model of tense. The Russian future tense forms are a great resource for expressing a wide range of meanings.

The imperfective future tense forms are a weaker reflection of what we observe with the perfective future. These forms are of much lower frequency than their perfective counterparts. The tendency to express non-future meanings observed for perfective verbs is also reflected in the imperfective future tense forms, though to a lesser degree.

The perfective future is formally unmarked with respect to the imperfective future, a relationship that is supported by the frequency distribution of forms. However, this pattern seems to be reversed for L2 learners, who cling to the *budu* imperfective future pattern for various reasons: in order to shield themselves from the complex morphology of Russian verbal conjugation, because *budu* as an overt analytical marker of future tense is easy to spot and remember, and because traditional teaching materials tend to overemphasize the imperfective future.

In this dissertation, we have learned more about the frequency distribution of Russian future tense forms across aspect, about the meanings expressed by these forms and their relative prominence, and about the differences between native and non-native speakers in negotiating use of these forms. We have thus expanded our linguistic description of Russian and our theoretical understanding of tense and aspect. Our findings indicate what measures might be appropriate to help L2 learners to achieve expression of future tense in a native-like manner.

Addenda

A References

Andersen, H. 1989. *Markedness theory: The history of a concept.* Berlin, New York: Mouton de Gruyter.

Andersen, H. 2006. Grammation, regrammation, and degrammation: Tense loss in Russian. *Diachronica* 23.2: 231-258.

Arunachalam, S. 2013. Experimental methods for linguists. *Language and Linguistics Compass* 7.4: 221-232.

Arutjunova, N. D. 1999. *Jazyk i mir čeloveka*. Moscow: Jazyki russkoj kul′tury.

Baerman, M., D. Brown, and G. G. Corbett. 2015. Understanding and measuring morphological complexity: An introduction. In Baerman, M., D. Brown, and G. G. Corbett (Eds.). *Understanding and measuring morphological complexity.* Oxford: Oxford University Press, 3-10.

Battistella, E. L. 1996. *The logic of markedness.* New York/Oxford: Oxford University Press.

Benko, V. and V. Zakharov. 2016. Very large Russian corpora: new opportunities and new challenges. *Komp'juternaja lingvistika i intellektuanyje texnologii: po materialam meždunarodnoj konferencii "Dialog"*, vol. 15(22), 79–93.

Berdicevskis, A. 2012. Jazykovaja složnost′ (language complexity). *Voprosy jazykoznanija*, (5), 101-124.

Bjerkeng, M., T. Bräger, and T. G. Nordenstam. 2011. *Sosedi 1.* Bergen: Fagbokforlag.

Bondarko, A. V. 1971. *Vid i vremja russkogo glagola.* Moskva: Prosveščenie.

Bosanquet, B., S. H. Hodgson, and G. E. Moore. 1897. In What Sense, If Any, Do Past and Future Time Exist? *Mind* 6.22: 228-240.

Brdar, M., and R. Brdar-Szabó. 2017. Moving-Time and Moving-Ego metaphors from a translational and a contrastive-linguistic perspective. *Research in Language* 15.2: 191-212.

Brun, D., S. Avrutin, and M. Babyonyshev. 1999. Aspect and its temporal interpretation during the Optional Infinitive stage in Russian. *Proceedings of the Annual Boston University Conference on Language Development* 23: 120-131.

Bulygina, T. V., and A.D. Šmelev. 1992. Kommunikativnaja modal'nost': konstatacija vozmožnosti, gipotezy i kvazi-soobšcenija. Kategorija skazuemogo v slavjanskich jazykach: modal'most'i aktualizacja. *Akty meždunarodnoj konferencii Certosa di Pontignano*, 26-29.

Bybee, J. 2010. *Language, Usage and Cognition.* Cambridge: Cambridge University Press.

Bybee, J. L., and Ö. Dahl. 1989. *The Creation of Tense and Aspect Systems in the Languages of the World.* Amsterdam: John Benjamins.

Bybee, J. L., R.D. Perkins, and W. Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World.* Vol. 196. Chicago: University of Chicago Press.

Casasanto, D., L. Boroditsky, W. Phillips, J. Greene, S. Goswami, S. Bocanegra-Thiel, et al. 2004. How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26. Retrieved from https://escholarship.org/uc/item/9m92g1s5

Casasanto, D. 2016. Temporal language and temporal thinking may not go hand in hand. In B. Lewandowska-Tomaszczyk (ed.), *Conceptualizations of time*. Amsterdam: John Benjamins, 169-186.

Cejtlin, S. N. 2000. *Jazyk i rebenok: Lingvistika detskoj reči.* Moscow: Gumanitarnyj izdatel'stvo centr VLADOS.

Černyšov, S. and A. Černyšova. 2018. *Poexali!-2 2.2. Russkij jazyk dlja vzroslyx*. Sankt-Peterburg: Zlatoust.

Černyšov, S. and A. Černyšova. 2019. *Poexali! 1.2. Russkij jazyk dlja vzroslyx*. Sankt-Peterburg: Zlatoust.

Christensen, A. 2004. *Russisk grammatik.* Copenhagen: Akademisk forlag.

Cienki, A. 1998. Metaphoric Gestures and Some of Their Relations to Verbal Metaphorical Expressions. In J.-P. Koenig (ed.), *Discourse and Cognition: Bridging the Gap*. Stanford, CA: CSLI Publications, 189–204.

Clark, E. 1998. Morphology in language acquisition. In A. Spencer and A. Zwicky (eds.), *Handbook of Morphology.* Oxford: Blackwell, 351-73.

Clark, H. H. 1973. Space, time, semantics, and the child. In T. Moore (ed.), *Cognitive Development and Acquisition of Language.* New York: Academic Press, 27-63.

Comrie, B. 1976. *Aspect*. Cambridge: Cambridge University Press.

Comrie, B. 1983. Markedness, grammar, people, and the world. In F. R. Eckman, E. A. Moravcsik, J. R. Wirth (eds.) *Markedness.* New York: Plenum Press, 85–106.

Comrie, B. 1985. *Tense*. Cambrige: Cambrige University Press.

Comrie, B. 1989. *Language universals and linguistic typology: syntax and morphology*. Chicago: University of Chicago Press.

Croft, W., and D. A. Cruse. 2004. *Cognitive Linguistics*. Cambrige: Cambrige University Press.

Čujkova, O. Ju. 2018. *Semantika buduščego vremeni v russkom, anglijskom i ispanskom jazykax: vzaimodejstvie temporal'nosti, aspektual'nosti i modal'nosti.* Dissertacija. Na pravax rukopisi. Sankt-Peterburg.

Dahl, Ø. 1995. When the future comes from behind: Malagasy and other time concepts and some consequences for communication. *International Journal of Intercultural Relations* 19.2: 197-209.

Dahl, Ö. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.

Dahl, Ö. 2004. *The growth and maintenance of linguistic complexity.* Amsterdam: John Benjamins B.V.

Dahl, Ö., V. Velupillai. 2013. The Future Tense. In M. S. Dryer and M. Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/67, Accessed on 2021-11-04).

Dainton, Barry. 2018. Temporal Consciousness. *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2018/entries/consciousness-temporal/>.

Debenedette, L., W. J. Comer, A. Smyslova, and J. Perkins. 2015. *Mezhdu nami*. https://mezhdunami.org

Dickey, S. M. 2000. *Parameters of Slavic Aspect*. A Cognitive Approach. Stanford: CSLI Publications.

Diessel, H. 2011. Review of *Language, Usage and Cognition*. *Language* 87.4: 830-844.

Diessel, H. 2019. *The Grammar Network: How Linguistic Structure is Shaped by Language Use*. Cambridge: Cambridge University Press.

Dryer, M. S. 2013. Position of Tense-Aspect Affixes. In M. S. Dryer and M. Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute

for Evolutionary Anthropology. (Available online at http://wals.info/chapter/69, Accessed on 2021-08-20.)

Fauconnier, G. 1985. *Mental Spaces.* Cambridge/London: MIT Press.

Faye, J. 1993. Is the future really real? *American Philosophical Quarterly* 30.3: 259-269.

Fillmore, C. J. 1982. Frame semantics. In *Linguistics in the Morning Calm Selected Papers from SICOL-1981*. Seoul: Hanshin Publishing Co, 111-137.

Filosofova, T. 2020. *Da!: a practical guide to Russian grammar.* Milton Park: Routledge.

Forsyth, J. 1970. *A grammar of aspect: Usage and meaning in the Russian verb*. Cambridge : Cambridge University Press.

Geeraerts, D. (ed.). 2006. *Cognitive linguistics: Basic readings* (Vol. 34). Berlin: Walter de Gruyter.

Givón, T. 2009. Introduction. In Givón, T., & Shibatani, M. (Eds.) *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution* (Vol. 85). Amsterdam: John Benjamins Publishing, 1-22.

Gottfredson, L. S. 1997. Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence* 24.1: 13-23. https://doi.org/10.1016/S0160-2896(97)90011-8

Gries, S. Th. 2021. *Statistics for Linguistics with R.* Berlin: De Gruyter Mouton.

Gvozdev, A. 1961. *Voprosy izučenija detskoj reči*. Moscow: APN RSFSR.

Hertz, B. 1999. *Møde i Petersborg. Svidanie v Peterburge.* København: Gylendahl Uddannelse.

Hilpert, M. 2015. *A course in Cognitive Linguistics: Usage-based linguistics*. [Video]. Youtube. https://www.youtube.com/watch?v=NAKqxQVpvFg

Hinchliff, M. 1996. The puzzle of change. *Philosophical perspectives* 10: 119-136.

Isačenko, A. 1965/2003. *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim.* Bratislava: Izdatel'stvo akademii nauk.

Ivanov, V. V. 1983. *Istoričeskaja gramatika russkogo jazyka*. Moscow: Prosveščenie.

Jakobson, R. O. 1971[1932]. Zur Struktur des russischen Verbums. In *Selected Writings II*. The Hague: Mouton, 3–15.

Janda, L. A. 2003. Russian aspect at your fingertips: Why what you know about matter matters, or a user-friendly conceptualization of verbal aspect in Russian. *The Slavic and East European Journal* 47.2: 251-281.

Janda, L. A. 2013. Quantitative Methods in Cognitive Linguistics. In L. A. Janda (ed.), *Cognitive Linguistics: The Quantitative Turn. The Essential Reader*. Berlin: De Gruyter Mouton, 1-32.

Janda, L. A. 2019. Businessmen and Ballerinas Take Different Forms. *Russian. Language Journal* 69: 175-194.

Janda, L. A., and S. J. Clancy 2002. *The Case Book for Russian*. Bloomington, IN: Slavica Publishers.

Janda, L. A., and O. Lyashevskaya. 2011. Grammatical profiles and the interaction of the lexicon with aspect, tense, and mood in Russian. *Cognitive Linguistics* 22.4: 719-763.

Janda, L. A., and O. Lyashevskaya. 2013. Semantic Profiles of Five Russian Prefixes: po-, s-, za-, na-, pro. *Journal of Slavic linguistics* 21.2: 211-258.

Janda, L.A. and T. Nesset. 2010. Taking apart Russian raz. *The Slavic and East European Journal* 54.3: 477-502.

Janda, L. A., V. Solovyev. 2009. What Constructional Profiles Reveal About Synonymy: A Case Study of Russian Words for sadness and happiness. *Cognitive Linguistics* 20.2: 367-393.

Janda, L.A., A. Endresen, A. and R. J. Reynolds. 2019. Aspectual opposition and rivalry in Russian are not discrete: new evidence from experimental data. *Russian Linguistics* 43.3: 249-271.

Kiebzak-Mandera, D., M. Smoczyńska, and E. Protassova. 1997. Acquisition of Russian verb morphology: The early stages. In W. Dressler (ed.), *Studies In Pre- and Protomorphology*. Vienna: Verlag Osterreichischen Akademie Wissenschaften, 101-114.

Efremov, D. A., and L. V. Kiseleva. 2015. O sposobax vyraženija buduščego vremeni v udmurtskom i vengerskom jazykax (sopostavitel′nyj analiz). Ogarëv -Online 2.43: 10.

Klein, W. 1994. *Time in language*. Milton Park: Routledge.

Klimonow, W. D. 2011. Vzaimodejstvie modal′nosti i aspektual′nosti v sovremennom russkom jazyke. *Vestnik Nižegorodskogo universiteta im. N.I. Lobačevskogo* 6.2: 262-265.

Kosheleva, D. and L. A. Janda. 2022. Looking into the Russian future. *CogniTextes. Revue de l'Association française de linguistique cognitive*. https://doi.org/10.4000/cognitextes.2087

Kosheleva, D. Under resubmission. Russian future: an inside and an outside perspective.

Kosheleva, D. and L. A. Janda. Accepted. Why markedness is always local.

Kovecses, Z. 2010. *Metaphor: A practical introduction*. Oxford: Oxford University Press.

Kusters, W. 2008. Complexity in linguistic theory, language learning and language change. In M. Miestamo, K. Sinnemäki, and F. Karlsson (eds.), *Language Complexity: Typology, contact, change.* Amsterdam: John Benjamins, 3-22.

Kutuzov, A., and M. Kunilovskaya. 2017. Size vs. structure in training corpora for word embedding models: Araneum russicum maximum and Russian national corpus. *International Conference on Analysis of Images, Social Networks and Texts*. Springer, Cham. https://doi.org/10.1007/978-3-319-73013-4_5

Kuznetsova, J. 2012. *Linguistic profiles: Correlations between form and meaning*. Tromsø: Universitetet i Tromsø PhD dissertation.

Lakoff, G., and M. Johnson. 1980. *Metaphors We Live By.* Chicago: University of Chicago Press.

Lakoff, G. 1993. The contemporary theory of metaphor. In A. Ortony (ed.), *Metaphor and Thought*, 2nd edition. Cambridge: Cambridge University Press, 202-251.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Langacker, R. W. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites.* Vol. 1. Stanford, CA: Stanford University Press.

Langacker, R. W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.

Laskareva, E. R. 2010. *Čistaja grammatika*. Sankt-Peterburg: Zlatoust.

Levine, J. S. 2009. *Russian Grammar*. New York: McGraw-Hill.

Levshina, N. 2015. How to do Linguistics with R: Data Exploration and Statistical Analysis. Amsterdam/Philadelphia: John Benjamins.

Lyons, J. 1977. Semantics (Vol. 1). Cambridge: Cambridge University Press.

Maddieson, I. 2013. Consonant Inventories. In M. S. Dryer and M. Haspelmath (eds.), The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://doi.org/10.5281/zenodo.7385533 (Available online at http://wals.info/chapter/1, Accessed on 2023-03-12.)

Maslov, Ju. 1990/2004. *Izbrannye trudy. Aspektologija. Obščee jazykoznanie*. Moskva: Jazyki slavjanskoj kul′tury.

Miestamo, M. 2008. Grammatical complexity in cross-linguistic perspective. In M. Miestamo, K. Sinnemäki, and F. Karlsson (eds.), *Language Complexity: Typology, contact, change.* Amsterdam: John Benjamins, 3-22.

Moore, K. E. 2014. *The Spatial Language of Time: Metaphor, Metonymy, and Frames of Reference.* Amsterdam/Philadelphia: John Benjamins.

Nesset, T. 1998. *Russian conjugation revisited: A cognitive approach to aspects of Russian verb inflection*. Oslo: Novus Press.

Nesset, T. 2009. Metonymy of aspect/aspects of metonymy. *Scando-Slavica* 55.1: 65-77.

Nesset, T. 2015. *How Russian Came to be the Way it is: A Student's Guide to the History of the Russian Language.* Bloomington, IN: Slavica Publishers.

Newman, J. 2011. Corpora and cognitive linguistics. *Revista Brasileira de Linguística Aplicada* 11.2: 521-559.

Nordenstam, T. G. 2013. *Azbuka 1: lærebok i russisk*. Oslo: Akademika.

Nordenstam, T. G. 2016. *Azbuka 2: lærebok i russisk*. Oslo: Akademika.

Olshevskaya, M. 2018. Ispol′zovanie russkogo učebnogo korpusa v prepodavanii rki: vid glagola. *Meždunarodnyj aspirantskij vestnik. Russkij jazyk za rubežom* 1: 13-18.

Ožegov, S. I. 1990. *Slovar' russkogo jazyka/SI Ožegov*. Moscow: Russkij jazyk.

Paul, R. 2007. *Language Disorders from Infancy through Adolescence: Assessment and Intervention.* (3rd Ed.). St. Louis: Mosby.

Petruxina, E. V., and Č. Li. 2015. Modal'nost' glagol'nyx form buduščego vremeni v russkom jazyke. *Vestnik Moskovskogo universiteta, Serija 9, Filologija* 4: 72-86.

Polinsky, M. 2006. Acquisition of Russian: Uninterrupted and incomplete scenarios. *Glossos* 8: 1-64.

Putnam, H. 1967. Time and physical geometry. *The Journal of Philosophy* 64.8: 240-247.

Pupynin, Ju. A. 1996. Usvoenie sistemy russkix glagol'nyx form rebenkom (rannie etapy). *Voprosy jazykoznanija* 3: 84-94.

Quené, H. 2010. How to design and analyze language acquisition studies. In E. Blom and S. Unsworth (eds.), *Experimental methods in language acquisition research*. Amsterdam: John Benjamins, 269-284.

Radbil', T. B. 2011. Buduščee kak fakt i buduščee kak modal′nost′ v paradigme «real′nost′-tekst ». In N. D. Arutjunova (ed.), *Logičeskij analiz jazyka. Lingvofuturizm. Vzgljad jazyka v buduščee.* Moscow: Indrik: 254-262.

Radden, G. 2003. The Metaphor TIME AS SPACE Across Languages. *Zeitschrift für interkulturellen Fremdsprachenunterricht* 8: 226–239.

Radden, G. 2011. Spatial time in the West and the East. *Space and Time in Language*, 1.

Radden, G., and Z. Kövecses. 2007.Towards a Theory of Metonymy. In V. Evans, B. Bergen, and J. Zinken (eds.), *The Cognitive Linguistics Reader*. London: Equinox, 335-359.

Reichenbach, H. 1947. *Elements of Symbolic Logic*. New York: Macmillan.

Reznikova, T. I., and M. V. Kopotev. 2005. Lingvističeski annotirovannye korpusa russkogo jazyka. In *Nacional'nyj korpus russkogo jazyka 2003-2005.* Moscow: Indrik, 31-61.

Riggs, P. J. 2015. Contemporary concepts of time in western science and philosophy. In *Long History, Deep Time: Deepening Histories of Place*. Canberra, Australia: ANU Press, 47-66.

Rosch, E. H. 1973a. Natural categories. *Cognitive psychology* 4.3: 328-350.

Rosch, E. H. 1973b. On the internal structure of perceptual and semantic categories. In T. E. Moore (ed.), *Cognitive development and acquisition of language*. Cambridge, MA: Academic Press, 111-144.

Rosch, E. H. 1978. Principles of categorization. In E. Rosch and B. B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale, NJ: Lawrence Erlbaum, 27-48.

Šatunovskij I. B. 2009. *Problemy russkogo vida*. Moskva: Jazyki slavjankskix kul'tur.

Sokolova, S., Bjørgve E., Janda L., Nesset T., Kosheleva D., and Zhamaletdinova E. *Min russiske reise*. https://mooc.uit.no/courses/course-v1:UiT+C001+2020/about

Švedova, N., et al. 1980. *Russkaja grammatika.* Moskva: Nauka.

Stojnova, N. M. 2016a. Nefutural'nye upotreblenija form buduščego vremeni. Materialy dlja proekta korpusnogo opisanija russkoj grammatiki. http://rusgram.ru (2 October, 2018).

Stojnova, N. M. 2016b. Analitičeskaja konstrukcija buduščego vremeni. Materialy dlja proekta korpusnogo opisanija russkoj grammatiki. http://rusgram.ru (2 October, 2018).

Stojnova, N. M. 2018. Buduščee vremja. Materialy dlja proekta korpusnogo opisanija russkoj grammatiki. http://rusgram.ru Na pravax rukopisi. Versija 06.2018

Stojnova, N. M. 2019. Buduščee nesoveršennogo vida i infinitivnoe sočetanie s glagolom stat′ kak konkurirujuščie konstrukcii v sovremennom russkom jazyke. *Russkij jazyk v naučnom osveščenii* 1: 58-83.

Swan, O. E. 2017. Sketching the Russian tense-aspect system for verbs that form simple aspect pairs. *Slavic and East European Journal* 61.4: 825-843.

Talmy, L. 1978. Figure and ground in complex sentences. *Universals of Human Language* 4: 631-649.

Talmy, L. 2000. *Toward a Cognitive Semantics.* Cambridge, MA: MIT Press.

Talmy, L. 2007. Foreword. In M. Gonzalez-Marquez et al. (eds.), *Methods in Cognitive Linguistics.* Amsterdam: John Benjamins, xi-xxi.

Tiersma, P. M. 1982. Local and general markedness. *Language* 59: 832–849.

Timberlake, A. 2004. *A Reference Grammar of Russian*. Cambridge: Cambridge University Press.

Vinogradov, V. V. 1947. *Russkij jazyk*. 2nd ed. Moscow: Vysšaja škola.

Wade, T., D. Gillespie, S. Gural, and M. Korneeva. 2020. *Comprehensive Russian Grammar*. West Sussex: Wiley-Blackwell.

Wiemer, B., A. Rostovtsev-Popel, J. Kwiatkowska, A. Socka. 2020. Četyrexugol′nik snjatoj utverditel′nosti v russko-pol′skom sopostavlenii. *VII Meždunarodnaja*

konferencija Komissii po aspektologii Meždunarodnogo komiteta slavistov. Moskva: Izdatel′stvo RGPU im. AI Gercena, 501-511.

Zaliznjak, A. A. and A. D. Šmelev. 2000. *Vvedenie v russkuju aspektologiju*. Moscow: Jazyki russkoj kul’tury.

Zangwill, I., and E. Nahshon. 1908. The melting pot. *From the ghetto to the melting pot: Israel Zangwills Jewish plays.* Detroit, MI: Wayne State University Press.

B Research data management

As part of the work on this dissertation, two datasets were annotated and analyzed for the second and the third articles. These datasets are publicly available in the TROLLing — The Tromsø Repository of Language and Linguistics. Replication data for the article *Looking into the Russian future* consisting of a readme file, two spreadsheets with examples from the Russian National Corpus (RNC), and terms of use for the RNC are available at: https://doi.org/10.18710/MHWRGE.

For the article *Russian Future: an inside and an outside perspective*, which is currently under submission, an anonymized dataset was created (available at: https://dataverse.no/dataset.xhtml?persistentId=doi%3A10.18710%2FR2EZN8&version=DRAFT). The dataset consists of a readme file, three files with the data received from the participants (same data in three different formats), and two files with the analysis of the data, including the annotation of different procedures and the code in R in two formats. No personal identifying information of the participants in the experiment was collected.

C Author contributions

Article I: Why markedness is always local

This article has two co-authors: Daria Kosheleva and Laura Janda. Kosheleva extracted and analyzed the corpus data, contributed to writing and theoretical background. Janda developed the theoretical background, and did some of the writing. The authors follow the Vancouver rules of authorship.

Article II: Looking into the Russian future

Like Article I, this article is co-authored by Daria Kosheleva and Laura Janda. The authors developed the theoretical background together, and wrote most of the text together. The corpus data was extracted and prepared for analysis by Kosheleva. The annotation and interpretation of the corpus data was performed by Kosheleva. The authors follow the Vancouver rules of authorship.

Article III: Russian future inside and out

The authorship of this article belongs to Daria Kosheleva. The experimental design was developed by Kosheleva. The data obtained in the experiment was preprocessed by Kosheleva. The statistical analysis of the data was prepared with significant help from Laura Janda. Kosheleva was the sole author of the text.

The authors thank the members of the research group CLEAR — Cognitive Linguistics: Empirical Approaches to Russian — for their valuable feedback and suggestions regarding all three articles.

D Acknowledgements

To my grandmother *baba Maxa*

Maus, Maus komm heraus — *myška, myška vyxodi*.

It takes a village to bring up a child. It literally takes the whole world to help build up the confidence to dare write a dissertation. There are extremely many people whom I would like to thank for their support and care; there is simply not enough virtual space on a virtual cloud. And well, my memory also has a limited capacity.

I want to start with the people who were directly involved in writing, reading, editing and other both pleasant and challenging moments of my PhD journey. I am very grateful to Laura Janda, Tore Nesset and the research group CLEAR, which I had the honor to be a part of. You inspired me to get on this road when I was still a master student on exchange at UiT, the Arctic University of Norway. I am proud to have learned from such competent and smart people as you all are. I am also thankful for the support and encouragement from the advisors and colleagues at the faculty and the institute.

Elmira Zhamaletdinova, you have been my friend and colleague for better and for worse. I would not have been able to go through the jungle parts of the PhD-forest if you hadn't been there for me. Valentina Zhukova, I am glad that we got to know each other better. And thank you for showing me places in Tromsø I have never been to.

It is always a pleasure to communicate with alike minds; but talking to people who come from a different framework can also be very enriching. I am glad that PhD-students from another research group within linguistics, CASTL — Anna Katharina Pilsbacher, Myrte Vos, Tor Håvard Solhaug, Maud Westendorp — moved to our hallway. Anna and Myrte deserve a special thanks for their support and fantastic sense of humor. Charlotte Sant, thank you for organizing the cake club:) I got the chance to become a little bit closer to such interesting people as Francesca Clémentine Festi and Paulina Lyskawa.

I would like to thank the “theory of science” gang of PhD-students in humanities who were there for me in good and bad times. Kat Hodgson, Anatolijs Venovcevs, Theodor Sandal Rolfsen, Vegard Stensen, I hope we will meet again, and I wish all of you to achieve the goals you have set in front of you.

The dissertation was not the only project I have worked on during these four years. I am happy that I had a chance to participate in planning, designing, and writing the texts and exercises for *Min russiske reise*. Svetlana Sokolova and Elena Bjørgve, you have been incredible to work with.

I would like to express my gratitude to the students whom I had the privilege to teach. Each and every one of you has made me a better instructor.

My PhD life would have been very different if I had not been a member of TAKk — Tromsø Akademiske Kvinnekor. I have never experienced such a feeling of sisterhood before I joined the choir. My fellow choir sisters, I already miss you. My special gratefulness goes to Christel Misund Domaas, Tone Hugstmyr, Ane Alund, Ellen Marie Bråthen Steen, Inga Bårdsen Tøllefsen, Paula Ryggvik Mikalsen, Emma Margret Skåden. Bjarne Isaksen, you left the choir before I did, making me miss you for one year more than I had anticipated.

To my recently acquired colleagues at The Norwegian Academy of Music: thank you for showing me the world of research beyond linguistics. Thank you for your support and encouragement in the days when everything seemed meaningless. Thank you for the music and the joy it is constantly bringing.

To all my friends in all parts of the world: thank you for being there for me. Space and time apart, I am happy that we still manage to say a word or two to each other.

To all my professors, instructors, and teachers: I am very lucky to have had you along the way of my life. Learning, discovering new things, and teaching are on the list of the most rewarding things I ever get to do.

Dear family, mom Irina Fedorovna and dad Leonid Nikolaevich, thank you for believing in me and letting me discover what I am capable of. I feel incredibly strong because I know how much support I have always had.

Tromsø, July 2022

Articles

I Why markedness is always local: the case of Russian aspect

Daria Kosheleva, Laura Alexis Janda

UiT, The Arctic University of Norway

Markedness is the observation of an encoding asymmetry in which higher complexity (both in terms of form and meaning) tends to pattern with lower frequency. Given that markedness focuses on the relationships between form-meaning patterns and usage patterns, markedness is of inherent theoretical interest for cognitive linguistics. Traditionally it is assumed that the markedness values of Russian aspect are perfective as marked vs. imperfective as unmarked. However, this assumption is not without controversy and conflicts with patterns observed in other languages. Furthermore, neither complexity of form nor corpus frequency support the traditional markedness assumption at the level of the category. We drill down to the levels of both the lexeme (groups of verbs defined by the major patterns of aspectual morphology) and the subparadigm (tense) and find better support for observation of markedness at these levels. While one group of verbs supports the traditional assignment of perfective as marked, two other groups of verbs support the opposite, with imperfective as marked. The subparadigm level of tense presents a special challenge since, due to confounding factors of homonymy and non-contiguous forms, no previous study has presented an accurate measurement of the incidence of future tense forms in Russian. We overcome this obstacle by examining a stratified set of verbs, sampling and manually tagging forms, and then using the sample data to extrapolate a reasonable estimate of future tense forms. We find that perfective future tense forms are approximately fourteen times more frequent than imperfective future tense forms. Russian future tense forms give strong support to the recognition of imperfective aspect as marked due to higher morphological complexity and much lower frequency. We conclude that it makes more sense to evaluate markedness patterns at local levels rather than at the category level.

Keywords: markedness, Russian, corpus, morphology, aspect, encoding asymmetry, future tense.

La notion de marquage décrit une asymétrie codée dans laquelle un plus haut degré de complexité (en termes de forme et de sens) tend à être corrélée avec une fréquence plus basse. Étant donné que la notion de marquage s'intéresse aux relations forme-sens et aux tendances liées à l'usage de la langue, le marquage est intrinsèquement lié aux théories développées dans le cadre de la linguistique cognitive. On considère traditionnellement que le marquage des valeurs aspectuelles en Russe opère ainsi : le perfectif serait marqué, à l'inverse de l'imperfectif qui ne le serait pas. Cependant, cette hypothèse est sujet à controverse et contredit des schémas remarqués dans d'autres langues. En outre, ni la complexité de forme, ni la fréquence d'usage dans un corpus ne soutient l'hypothèse traditionnelle au niveau de la catégorie. Nous irons au cœur des niveaux du lexème (des groupes de verbes définis par des schémas récurrents majeurs de la morphologie aspectuelle) et du sous-paradigme (temps) et nous trouverons plus de soutien pour la présence de marquage à ces niveaux. Alors qu'un groupe de verbes soutient l'hypothèse traditionnelle du perfectif marqué, deux autres groupes de verbes soutiennent l'inverse, où c'est l'imperfectif qui est marqué. Au niveau du sous-paradigme du temps, nous rencontrons un obstacle particulier en raison de facteurs confusionnels d'homonymie et de formes non-contigües. Aucune étude n'avait jusqu'à lors présenté une mesure précise de l'incidence des formes du futur en Russe. Nous proposons de résoudre ce problème en

examinant un groupe stratifié de verbes, en échantillonnant et marquant manuellement des formes, et en utilisant cet échantillon de donnée pour extrapoler une estimation raisonnable des formes futures. Nous trouvons que les formes perfectives du futur sont approximativement quatorze fois plus fréquentes que les formes imperfectives du futur. Les formes du futur en russe démontrent avec force que l'aspect imperfectif est marqué en raison du haut degré de complexité morphologique et d'une fréquence bien plus basse. Nous concluons qu'il fait plus sens d'évaluer les schémas de marquage au niveau local qu'au niveau catégoriel.

Mots clés : marquage, Russe, corpus, morphologie, aspect, encodage de l'asymétrie, temps du futur

1. Introduction

This article tackles the question of the markedness of perfective vs. imperfective aspect in Russian, an issue that has attracted considerable attention in the scholarly literature (see discussion and citations in Section 3). The prevailing assumption is that Russian is a language in which the markedness values for aspect are relevant at the category level and that perfective is marked whereas imperfective is unmarked, the reverse of what is observed for most other languages (Dahl 1985). However, we argue that looking for markedness at the category level of aspect is supported neither by the overt marking of morphology nor by corpus frequencies. Instead, we examine various local levels: lexemes grouped according to their type of aspectual marking, as well as tense, and find more convincing markedness patterns at these local levels. We also tackle the previously unsolved problem of accurately measuring the incidence of future tense in Russian, and present compelling evidence that within future tense, we must acknowledge perfective as unmarked and imperfective as marked. This finding stands in stark contrast to the traditional assumption.

Section 2 presents the theoretical connection of this research to markedness and to cognitive linguistics. We define both markedness and local markedness with respect to scholarly traditions. We identify three parameters that are relevant to the observation of markedness relations, and two of these parameters can be operationalized: morphological complexity and corpus frequency. In Section 3 we present Russian aspect, how it is signaled morphologically, and how this verbal category is traditionally interpreted with respect to markedness. Given the parameters of markedness, we find a lack of convincing evidence for the traditional interpretation. We proceed to investigate the markedness of aspect at the local level of three different groups of lexemes, defined by the three major morphological patterns of aspectually paired verbs: "A", "B", and "C". We furthermore examine markedness at the

subparadigm level of past tense. We show that at all of these local levels, we find better evidence of markedness relations, though they point in different directions: "A" suggests that perfective is marked, while "B" and "C" suggest that imperfective is marked. Section 4 undertakes the task of measuring the frequency of perfective vs. imperfective future forms, and we find that at this local level, there is strong evidence that imperfective is marked. We summarize our findings in Section 5.

2. Markedness and cognitive linguistics

Markedness is a theoretically neutral descriptive concept that focuses on relationships between elements in categories (Battistella 1990: 5), a concept with a long history spanning linguistic traditions (Andersen 1989 and 2001, Prince & Smolensky 2008). When understood in scalar terms (Janda 1995), markedness is entirely compatible with cognitive linguistics (Lakoff 1987: 59-61, van Langendonck 1989: 180, Diessel 2019: Chapter 11) and can serve as a bridge facilitating the integration of linguistic analyses across theoretical frameworks. Haspelmath (2006) argues against the use of the term "markedness" in favor of descriptions of linguistic forms and their relative frequencies, yet the term persists (even in his own subsequent work, cf. Haspelmath & Karjus 2017), thanks to its usefulness in capturing relationships between meaning, form, and frequency.

Comrie (1983: 95) urged linguists to "try to account for markedness in terms of other, independently verifiable properties of people, the world, or people's conception of the world". Comrie's grounding of markedness in these terms resonates well with Langacker's (2008: 85, 39–57) description of cognitive linguistics as a framework that employs only descriptive constructs based on well-known cognitive phenomena and takes into account an encyclopedic view of meaning. The association of markedness with frequency (Haspelmath 2006; see Table 1 below) comports with the usage-based approach of cognitive linguistics and supports a scalar interpretation of markedness.

Markedness describes an asymmetric relationship between two or more elements that are both contrasted and related to each other (Andersen 1989: 37–39), termed "encoding asymmetries" by Diessel (2019: Chapter 11) and Haspelmath & Karjus (2017). The term "encoding asymmetry" refers to a situation in which one item is overtly marked, while the other item is not marked. The prototype – periphery structure of radial categories that is a persistent feature of cognitive linguistics (Rosch 1973a and b, Lakoff 1987, Lewandowska-Tomaszczyk 2007) is a satisfactory model for such an asymmetric relationship. The prototype – periphery structure models an asymmetric relationship, where the prototype is both contrasted to the other members

of a category, while at the same time related to those other members (Mayerthaler 1980: 26). The prototype of a radial category is the unmarked member, while the other peripheral members are marked, and their markedness can be measured in terms of distance from the prototype.

Markedness and radial category structure tend to align along three parameters outlined in Table 1: expectedness, complexity, and frequency. In describing semantic markedness, Jakobson (1971[1932]) states that the unmarked member lacks a semantic mark, as opposed to a semantically marked member that has more restricted distribution. Jakobson's example is Russian *osel* 'donkey' which is unmarked for sex and can refer to any donkey, as opposed to the marked *oslica* which can only refer to a female donkey. In this example, *osel* 'donkey' is the **most expected** item because in most situations when we speak about donkeys, we are not speaking only about female donkeys and therefore do not need to specify the sex of the animal. In terms of **complexity**, *osel* 'donkey' is a monomorphemic lexeme with a semantically simple meaning, whereas *osl-ica* 'female donkey' is more semantically complex, referencing both 'donkey' and 'female', and more morphologically complex, since it is comprised of two morphemes, *os/l* 'donkey' + *-ica* 'female'. In terms of **frequency**, unmarked items are typically of higher frequency than marked items, and this is borne out by corpus data. In the Russian National Corpus (ruscorpora.ru; RNC), the unmarked *osel* 'donkey' has 5774 attestations, over twenty times more than the marked *oslica* 'female donkey', with only 281 attestations. In terms of cognitive linguistics, *osel* 'donkey' is the prototypical member of a radial category in which *oslica* 'female donkey' is a more peripheral member.

Comrie (1989: 85) describes the unmarked member of an opposition as the one that is more expected. Both of these descriptions correspond to the default nature and expectedness of a prototypical instance of a category in contrast to a peripheral instance. Diessel (2019: 224) likewise highlights the function of markedness as a "strategy to indicate constructions that deviate from listeners' linguistic expectations". The less expected marked member motivates the use of overt means to distinguish the marked member from the unmarked prototype, yielding the common observation that marked members tend to have overt morphological marks and thus higher formal complexity. The alignment of semantic complexity (unexpectedness) with formal complexity is termed by Haspelmath & Karjus (2017: 1218) "iconicity of complexity".

The relationship between markedness and frequency has a long tradition, going back at least to Greenberg (see overviews in Andersen 1989: 28–30; Battistella 1996: 13–14, 50–55; Andersen 2001: 50–51). The more expected and less complex unmarked

prototype is likely to be more frequent than the less expected more complex peripheral marked member. Table 1 lays out the typical relationship among the three parameters of expectedness, complexity, and frequency, illustrated with Jakobson's example as described above.[14]

	unmarked ≈ prototype *osel* 'donkey'	marked ≈ periphery *oslica* 'female donkey'
expectedness	more expected used when talking about donkeys in general	less expected used only when talking specifically about female donkeys
complexity	less complex simple meaning, one morpheme	more complex complex meaning, two morphemes
frequency	more frequent 5774 attestations in RNC	less frequent 281 attestations in RNC

Table 1. Comparison of tendencies for marked vs. unmarked to align with prototype vs. periphery along the parameters of expectedness, complexity, and frequency.

For the purposes of this article, while the parameters outlined in Table 1 are certainly linked to each other, we do not presume any causal or necessary relationships among them (for relevant discussion, see Tiersma 1982, Haspelmath & Karjus 2017, Diessel 2019: Chapter 11). The challenge for our analysis is operationalizing the distinction between items that are relatively marked in relation to items that are relatively unmarked. The most accessible means for operationalizing the markedness distinction are via observation of formal (morphological) complexity and corpus frequency. Morphological complexity and frequency are therefore central to the analysis in Sections 3 and 4, although expectedness is also taken into account where appropriate. We recognize phenomena that comport with the tendencies in Table 1 along two or all three parameters as supporting the recognition of a contrast between a relatively marked vs. a relatively unmarked member of a relationship.

[14] The rows in Table 1 correspond to Haspelmath's (2006) "senses of markedness" as follows:

- expectedness = senses 4, 5, 6, 11 (markedness as difficulty, deviation from default)
- complexity = senses 1, 2, 3 (markedness as specification for a distinction and overt coding)
- frequency = senses 7, 8, 9, 10 (markedness as rarity and restricted distribution)

"Local markedness" (also known as "markedness reversal") was first identified by Tiersma (1982) to describe the fact that markedness values are not always uniform for a given grammatical category across all lexical items. Tiersma examined singular vs. plural markedness patterns, showing that for most nouns, plural is marked (and less frequent) with respect to singular. However, some nouns, particularly those referring to objects often encountered in groups or pairs, show the opposite, with the singular as marked (and less frequent). Furthermore, multiple markedness values may converge: "a lexical item may be generally unmarked in one category and locally unmarked in another" (Tiersma 1982: 857). We will expand Tiersma's concepts of general vs. local markedness in Section 4 to show that different markedness values may also compete within a paradigm.

In the next section we explore how Russian verbal aspect has been interpreted in terms of markedness and how these interpretations correspond to the three parameters of expectedness, complexity, and frequency.

3. Russian aspect with respect to markedness

The markedness relationship that characterizes Russian aspect is generally acknowledged to be the opposite of that observed in other languages. However, the received wisdom about Russian aspect and markedness is not without controversy and the relationship is more complex than perceived at first glance.

With regard to most languages that have a perfective vs. imperfective aspectual distinction, scholars observe that perfective usually behaves as the unmarked member of the opposition, while imperfective is the marked member. Slavic languages stand out as an idiosyncratic deviation from this norm (Dahl 1985: 71–85, Binnick 1991: 136–139), with most scholars (see citations below) labeling perfective as the marked member in opposition to the unmarked imperfective. There are two other ways in which Slavic aspect deviates from the typological norm: a) extent – the grammaticalized perfective vs. imperfective distinction holds for all verb forms of all tenses and moods rather than being restricted to the past tense, and b) balance – the imperfective aspect appears in a larger range of contexts than in other languages.

In terms of extent, the Slavic aspectual distinction is realized at the lexical (derivational) level rather than the inflectional level, meaning that an entire verb is either perfective or imperfective. In Russian it is traditional to refer to verb "pairs", consisting of a perfective verb and an imperfective verb that bear the same meaning and are differentiated by aspectual morphology (prefixes and suffixes). The lexical

level of aspect in Russian is illustrated in Table 2. All forms of the perfective verbs *s-vjazat'* 'knit, tie' and *pri-vjazat'* 'tie one thing to another' are perfective, whereas all forms of the imperfective verbs *vjazat'* 'knit, tie' and *pri-vjaz-yva-t'* 'tie one thing to another' are imperfective; this generalization holds also for other verb forms not represented in Table 2, such as imperatives, gerunds, and participles.

		example verb: infinitive	past (M.SG)	inflectional non-past (3SG)	periphrastic future (3SG)
"A" pattern	perfective prefixed with *s-*	*s-vjazat'* 'knit, tie'	*s-vjaza-l* 'he knitted, tied'	*s-vjaž-et* 's/he will knit, tie' [FUT]	-
	simplex imperfective	*vjazat'* 'knit, tie'	*vjaza-l* 'he knitted, tied'	*vjaž-et* 's/he knits, ties'	*budet vjazat'* 's/he will knit, tie' [FUT]
"B" pattern	perfective prefixed with *pri-*	*pri-vjazat'* 'tie one thing to another'	*pri-vjaza-l* 'he tied one thing to another'	*pri-vjaž-et* 's/he will tie one thing to another' [FUT]	-
	imperfective suffixed with *-yva-*	*pri-vjaz-yva-t'* 'tie one thing to another'	*pri-vjaz-yva-l* 'he tied one thing to another'	*pri-vjaz-yva-et* 's/he ties one thing to another'	*budet* *pri-vjaz-yva-t'* 's/he will tie one thing to another' [FUT]

Table 2. Illustration of Russian tense and aspect morphology using verbs and forms related to vjazat' 'knit, tie'. Hyphens indicate morpheme boundaries in order to clarify the presence of aspectual prefixes (s- and pri-) and suffixes (-yva). Forms associated with future tense are shaded and marked "[FUT]".

In the majority of instances, the aspectual difference in Russian is overtly marked by morphology, following one of two patterns, either: A) a prefix marks perfective as

opposed to an imperfective verb without the prefix, or B) a perfective verb (often with a prefix) is opposed to an imperfective verb that is marked with a suffix.[15] In the top half of Table 2, the "A" pattern is illustrated with the prefixed perfective verb *s-vjazat'* and its corresponding imperfective simplex verb *vjazat'*, both of which mean 'knit, tie' (but differ in aspect). The "B" pattern is illustrated in the bottom half of Table 2 with the perfective *pri-vjazat'* 'tie one thing to another' and the suffixed secondary imperfective *pri-vjaz-yva-t'* that shares the same meaning (with the only difference being in aspect).

In terms of balance, Slavic languages use (or allow) an imperfective in situations where most other languages would require a perfective. The skew toward imperfective is particularly strongly documented for Russian. For example, a detailed comparison of contexts with perfective verb forms in Spanish but imperfective verb forms (Russian in Janda & Fábregas 2019) shows that Russian uses imperfective in many contexts where Spanish uses perfective (but note that the reverse can also occur, cf. Fábregas & Janda 2019).

Scholarship on Russian aspect in the twentieth century was dominated by the invocation of features, where the positive value is associated with the perfective aspect as the marked member of the opposition, while the imperfective lacks the feature. The features that characterize perfective aspect fall in three (somewhat overlapping) groups: boundedness, totality, and definiteness.

Boundedness or telicity refers to the reaching of a limit (Jakobson 1971[1957], Avilova 1976). For the perfective verbs in Table 2, this means that the act of knitting or tying has come to a close. Other names for (approximately) the same feature include delimitation (Bondarko 1971), demarcatedness/dimensionality (van Schooneveld 1978), and closure (Timberlake 1982).

Totality captures the idea that a perfective situation is viewed as a complete whole (Isačenko 1960, Maslov 1965, Bondarko 1971, Comrie 1976, Smith 1991, Durst-Andersen 1992). This comports also with Wierzbicka's (1967/2018) observation that the Slavic imperfective (based on Polish examples) has a more general meaning as opposed to the perfective that refers to a specific completion. For our knitting and

[15] There are several hundred biaspectual verbs that do not overtly mark aspect, but most scholars argue that in context a biaspectual verb always expresses only one aspect (see Janda 2007: 90 and Zaliznjak & Šmelev 2000: 10 and citations therein).

tying verbs, this means that the knot is finished. Totality is akin both to Vinogradov's (1972) completion feature and to Langacker's (2008: 111–112) summary scanning.

Both Bondarko (1971) and Dickey (2000) use the feature definiteness to characterize the tendency of the perfective to refer to single, individuated actions, as a verbal parallel to nominal definiteness which refers to single, individuated entities. With respect to the knitting and tying verbs in Table 2, this means that a perfective verb references a specific unique event.

There are some dissenters from the majority opinion that the Russian perfective is marked and the imperfective is unmarked. Galton considered the markedness values for Slavic to be the reverse, following the typological norm of perfective as unmarked and imperfective as marked. Galton's (1976: 289) argument is based primarily on the parameter of expectedness, stating that it is more usual and thus grammatically less marked to view "an occurrence as part of its temporal succession", his characterization of the function of the perfective. Padučeva (1996) argued that the distinction in Russian is equipollent, because both the perfective and the imperfective have positive characteristics, and the complexity of imperfective uses cannot be accounted for by means of a lack of a feature. Zaliznjak & Šmelev (2000: 16–17) are more equivocal: while acknowledging the "real asymmetry" that is captured by the traditional interpretation of perfective as marked and imperfective as unmarked, they conclude that the importance of markedness for Russian aspect has been overrated.

Thus, the prevailing view of scholarship on Russian aspect is that perfective is marked and imperfective is unmarked. We reexamine this assessment in the light of evidence in terms of the parameters in Table 1. We will examine the question of the markedness of perfective vs. imperfective at various levels: the category level of total presence of perfective vs. imperfective, the local lexeme level of three major patterns of aspectually paired verbs ("A" and "B" in Table 2, plus "C" in Table 4 below), and the local level of the future tense subparadigm. The lexeme levels will be addressed in the following subsections, while the question of the local level of the future tense will occupy Section 4.

3.1 The category level of perfective vs. imperfective

Is the consensus view that perfective is marked and imperfective is unmarked supported by frequency data? If so, we would expect imperfective verb forms to be more frequent than perfective verb forms. We can search the Russian National Corpus (ruscorpora.ru, henceforth "RNC") to compare the overall occurrence of all perfective vs. imperfective verb forms. As of January 2021, the RNC contained 321 712 061

words, 26 575 727 of which are perfective verb forms, and 32 459 309 of which are imperfective verb forms. In other words, 45% of verb forms are perfective and 55% are imperfective. It appears that there are indeed more imperfective than perfective verb forms. However, this difference is not very large, and although it is statistically significant, the effect size is an order of magnitude too low to be considered an important difference.[16] The frequency difference therefore lends at best weak support to the claim that perfective is marked and imperfective is unmarked at the level of the entire category of aspect.

As stated above, we examine at least two parameters in evaluating a markedness relationship: frequency and morphological complexity. As we show in the next subsection, the morphology of Russian can point to both perfective and imperfective as marked.

3.2 Local lexeme level of "A" and "B" patterns of aspectually paired verbs

The two predominant patterns of Russian aspectual morphology illustrated in Table 2 show opposite patterns of morphological complexity. Both the "A" and the "B" patterns involve aspectual pairs of verbs, where both verbs have the same meaning and differ only in their aspectual values. In the "A" pattern, the imperfective verb is what we call "simplex" because it has no aspectual morphology, no prefix or suffix that identify it as perfective or imperfective. The perfective verb in the "A" pattern is formed by adding a perfectivizing prefix to the imperfective verb. Thus, in the "A" pattern, the perfective is morphologically more complex. In the "B" pattern, both the perfective and the imperfective verb bear a prefix, and the imperfective is formed by adding an imperfectivizing suffix to the perfective verb. Thus, in the "B" pattern, it is the imperfective that is morphologically more complex.

Given that both the "A" pattern and the "B" pattern give evidence of morphological complexity, albeit in different directions, it makes sense to ask whether frequencies support indications of markedness values. Although the RNC does not tag verbs according to whether they belong to the "A" pattern or the "B" pattern, Janda & Lyashevskaya (2011) undertook a large-scale analysis of nearly six million verb forms in the RNC to identify verbs according to their morphological pattern. The data cited by Janda & Lyashevskaya is disaggregated according to subparadigms (infinitive, past,

[16] A comparison of the numbers of perfective and imperfective against the total number of verbs yields a chi-square value of 391 854, df= 1, a p-value of 0, and a Cramer's V effect size of 0.047. Cramer's V effect size is interpreted as follows: 0.1 = small, 0.3 = medium, 0.5 = large. Cramer's V effect sizes below 0.1 are considered too low to be reportable.

non-past, and imperative), and has been reaggregated to represent the total frequency of each of the two patterns in Table 3. The frequency of past tense forms is provided for an additional comparison in the rightmost column of Table 3.[17]

Pattern	Aspect	Morphological complexity	Total frequency	Frequency of past tense forms
"A" pattern	perfective	simplex + prefix	528 749	317 570
	imperfective	simplex	1 105 655	397 409
"B" pattern	perfective	simplex + prefix	2 618 534	1 654 717
	imperfective	simplex + prefix + suffix	1 698 312	517 965

Table 3. Comparison of the morphological complexity and frequency of perfective and imperfective Russian verbs following the dominant "A" and "B" patterns. Frequency data is cited from Janda & Lyashevskaya 2011.[18]

If we look at total frequency in Table 3, within the "A" pattern, perfective verbs are relatively more complex morphologically, and imperfective verbs are more than twice as frequent. The overall frequency difference in the "A" pattern is both significant and of a reportable size. Within the "B" pattern, it is the imperfective verbs that are relatively more complex morphologically, and perfective verbs are more frequent. The overall frequency difference in the "B" pattern is likewise both significant and of a reportable size.[19] If we examine the frequency of the past tense, all the same observations hold: the differences between past tense frequency and total frequency

[17] We are grateful to an anonymous reviewer who suggested that it would be helpful to cite the frequency of past tense forms in addition to the total frequency, since the past tense is the one tense in which perfective and imperfective verbs share the same inflectional morphology in Russian (see Table 2). As shown in Tables 3, 4, and 5, the overall pattern is the same regardless of whether one compares the total frequency or the frequency of the past tense, with the exception of the verb pair meaning 'understand' in Table 5.

[18] It might seem that the data in Table 3 contradict the overall data cited from the RNC in Section 3.1 because the overall data shows a that imperfective verb forms are somewhat more frequent, whereas aggregation of the data in Table 3 indicates that perfective verbs are more frequent. However, the RNC data cited in Section 3.1 represents all verbs regardless of whether they are paired for aspect (as most verbs are) or unpaired, whereas the data in Table 3 represents only paired verbs according to patterns "A" and "B". There are also unpaired verbs in Russian, in particular the verb *byt'* 'be' which is imperfective, and which is of very high frequency, thus accounting for part of the apparent discrepancy.

[19] Comparison of the numbers of perfective and imperfective against the total number of verbs in the "A" pattern yields a chi-square value of 140053, df= 1, a p-value of 0, and a Cramer's V effect size of 0.168. Comparison of the numbers of perfective and imperfective against the total number of verbs in the "B" pattern yields a chi-square value of 132234, df= 1, a p-value of 0, and a Cramer's V effect size of 0.101.

are significantly different and of a reportable size for both the "A" and the "B" patterns.[20] However, it is important to note that all of the effect sizes for these differences are small (Cramer's V values < 0.2).

To summarize, for the "A" pattern it is perfective that is more complex and less frequent, but for the "B" pattern it is imperfective that is more complex and less frequent.

Both the morphological complexity and the frequency data in Table 3 support an interpretation of local markedness according to which perfective verbs are marked in pattern "A", but unmarked in pattern "B", although the frequency data gives only modest support to this interpretation. While patterns "A" and "B" represent the largest morphologically defined groups of aspectually paired verbs, it is also possible to look at another smaller pattern, "C", where the meanings of the verbs are arguably an important factor.

3.3 Local level of simplex perfective verbs

Tiersma (1982) observed that the semantics of some nouns can lead to local markedness values that reverse the markedness values of the majority of nouns, namely that nouns referring to items usually found in pairs or groups tend to have singular as their marked value, with corresponding higher formal complexity and lower frequency. We suggest that analogous semantic factors can come into play also for verbal aspect. In Russian, most simplex verbs are imperfectives that signal activities (like *vjazat'* 'knit, tie' in Table 2) or states (like *sidet'* 'sit'). However, there are a few simplex verbs that signal achievements and are perfective. Four such perfective simplex verbs that have imperfective correlates are presented in Table 4. We can call this the "C" pattern of morphological marking of aspect in Russian.

	Aspect	Morphological complexity[21]	Total frequency	Frequency of past tense forms
'give'	perfective *dat'*	simplex	323 798	73 641

[20] Comparison of the numbers of past perfective and past imperfective against the total number of verbs in the "A" pattern yields a chi-square value of 30968, df= 1, a p-value of 0, and a Cramer's V effect size of 0.116. Comparison of the numbers of perfective and imperfective against the total number of verbs in the "B" pattern yields a chi-square value of 159684, df= 1, a p-value of 0, and a Cramer's V effect size of 0.154.

[21] There are three imperfectivizing suffixes in Russian: *-yva(j)/-iva(j)* is illustrated in Table 2, *-va(j)* is found in *da-vat'* 'give' and *vsta-vat'* 'stand up', and *-a(j)* is found in *polučat'* 'receive' *rešat'* 'decide' (cf. conjugation *rešaju* 'I decide').

	imperfective *da-vat'*	simplex + suffix	162 109	29 680
'stand up'	perfective *vstat'*[22]	simplex	57 392	40 421
	imperfective *vsta-vat'*	simplex + suffix	24 722	5 896
'decide'	perfective *rešit'*	simplex	105 240	73 740
	imperfective *reš-at'*	simplex + suffix	25 277	2 823
'receive'	perfective *polučit'*	simplex	172 486	68 760
	imperfective *poluč-at'*	simplex + suffix	57 111	13 049

Table 4. The "C" pattern: four perfective simplex verbs and their suffixed imperfective correlates. Frequency data is cited from the RNC.

Giving, standing up, deciding, and receiving are all situations that tend to be understood as momentary, complete, and unique. The semantics of these verbs motivates the interpretation of perfective as relatively more expected and therefore unmarked. This interpretation is supported both by morphological complexity, which is higher for the corresponding imperfectives that are overtly marked by suffixes, and by frequency, which is higher for the perfectives. All of the differences in frequency presented in Table 4 are statistically significant and represent reportable differences.[23] In the case of 'decide' the total frequency difference approaches a medium effect size.

In sum, we see that it is hard to support an overall category-level interpretation of perfective as marked and imperfective as unmarked in Russian, since frequency is inconclusive at that level. Instead, we find more convincing alignments of morphological complexity, frequency, and even expectedness within groups of verbs that have different markedness values for aspect. "A" pattern verb pairs support the interpretation of perfective as marked and imperfective as unmarked. "B" and "C" pattern verb pairs support the interpretation of imperfective as marked and perfective as unmarked.

---

[22] Although etymologically *vstat'* 'stand up' contained a prefix (*vz-*), this prefix has suffered phonological erosion to the point that it is no longer recoverable for contemporary speakers. This verb functions as a simplex stem in modern Russian according to Endresen & Plungian (2011).

[23] The comparisons yield the following values for total frequency (similar values obtain for past tense forms):
'give': chi-square value = 36872, df= 1, p-value = 0, Cramer's V = 0.158.
'stand up': chi-square value = 9019.3, df= 1, p-value = 0, Cramer's V = 0.189.
'decide': chi-square value = 36192, df= 1, p-value = 0, Cramer's V = 0.295.
'receive': chi-square value = 41286, df= 1, p-value = 0, Cramer's V = 0.24.

We find that a finer grained analysis gives us a better analysis. This finding motivates us to investigate whether it is possible to take this line of reasoning one step further and look at a part of the verbal paradigm where there are additional differences in the morphological complexity, namely the future tense. In order to answer this question, however, we must overcome the considerable obstacles that stand in the way of accurately measuring the corpus frequency of the Russian future tense.

4. The Russian future

There is one major gap in the data presented by Janda & Lyashevskaya (2011): that study did not address the future tense in Russian. There is a good reason for this, namely that due to a variety of confounding factors, it is notoriously difficult to measure the occurrence of the future tense in Russian. In this section we present the first attempt at an accurate measure of the corpus frequency of the Russian future tense.

As illustrated in Table 2, inflected verb forms in Russian can express two tenses, one that is past, and one that is not past. The non-past forms of perfective verbs (such as *s-vjaž-et* 's/he will knit, tie' and *pri-vjaž-et* 's/he will tie one thing to another') are associated with future tense (FUT), and we refer to them as "future forms" in this article. The corresponding non-past forms of imperfective verbs are associated with present tense. For imperfective verbs the future tense is expressed by means of a periphrastic form consisting of an auxiliary that expresses person and number combined with the imperfective infinitive form. The auxiliary is identical to the forms of the verb *byt'* 'be', which is the only verb in Russian that can be said to have a true dedicated future form; *budet* when it is not an auxiliary, for example, means 's/he will be' as in (1).

(1)	Zavtra	on	uže	**bud-et**	v	Magadan-e.
	tomorrow	he.NOM	already	**will.be-3.SG**	in	Magadan-LOC.SG

'Tomorrow he **will** already **be** in Magadan.'

[J. Rytxèu. V doline Malen'kix Zajčikov, 1962][24]

[24] All examples in this article are cited from the Russian National Corpus (ruscorpora.ru), the metadata in their passports are given in square brackets.

In Table 2, the periphrastic imperfective future form is illustrated by *budet vjazat'* 's/he will knit, tie' and *budet pri-vjaz-yva-t'* 's/he will tie one thing to another', and we call these "future forms" as well. In this section, we focus on the future forms of both perfective and imperfective verbs such as those in the shaded boxes in Table 2.

For the purposes of this part of our analysis, the most important fact to observe from Table 2 is that an imperfective future form is always both longer and more morphologically complex than the corresponding perfective future form. This is because the perfective future form is merely a conjugated form of the verb, whereas the imperfective future form is a conjugated auxiliary form plus the imperfective infinitive. Isačenko (1960: 444) considers the more morphologically complex imperfective future form to be marked in relation to the perfective future form, despite the fact that his overall assessment is that perfective is marked in relation to imperfective.

Ideally, one would hope to get some global statistics on the distribution of perfective and imperfective future forms from the RNC. Unfortunately, due to various facts of Russian morphology and syntax, it is not easy to extract exact numbers reflecting all and only such future forms. These facts involve homonymy and non-contiguity of periphrastic forms. Tagging available in the RNC does not always successfully identify perfective future forms, and does not identify imperfective future forms at all.

4.1 Homonymy and non-contiguity of future forms

Homonymy is problematic for three reasons, involving syncretism within and across verb paradigms. The first type of homonymy occurs when a perfective non-past second person plural form is homonymous with the second person plural imperative form of the same verb, as in *pogovori-te*, which is analyzed either as [speak.pfv-fut.2.pl] 'you will speak' or [speak.pfv-imp.2.pl] 'speak!' depending on context.[25] Note how these two examples illustrate the two possible interpretations of *pogovorite* 'talk', both from the same author. In (2a) we see the future form, and in (2b) we see an imperative form.

(2)	a.	… mož-et	by-t'	vy	vse-taki	**pogovori-te**	so
		may-3.SG	be-INF	you.NOM	anyhow	**speak.PFV-FUT.2.PL**	with

[25] For some verbs stress disambiguates the perfective future from the imperative, but stress is not marked in the corpus.

svo-im	ženix-om	ob”jasn-ite	emu	moj-u	situacij-u…
own-M.INS.SG	fiancé.INS.SG	explain.PFV-FUT.2.PL	he.DAT	my-F.ACC.SG	situation-ACC.SG

‘…maybe you **will talk** with your fiancé anyhow, and explain my situation to him…’

[Aleksandra Marinina. Poslednij rassvet. 2014]

b. Poslušaj-te,	**pogovori-te**	s	rabotnik-ami	polici-i…	
listen.PFV-IMP.2.PL	**speak.PFV-IMP.2.PL**	with	worker-INS.PL	police-GEN.SG	

‘Hey, **talk** to the police officers...’

[Aleksandra Marinina. Angely na l’du ne vyživajut. 2014]

A second type of homonymy involves biaspectual verbs such as *operirova-t’* [operate.pfv/ipfv-inf] ‘operate’ that can express either aspect, again depending on context, as in *operiruj-ut* [operate.pfv-fut.3.pl] ‘they will operate’ as in (3a) vs. [operate.ipfv-prs.3.pl] ‘they operate’ as in (3b).

(3) a.

Zavtra	ee	**operiruj-ut.**
tomorrow	she.ACC	**operate.PFV-FUT.3.PL**

‘She **will be operated on** tomorrow.’

[Nina Katerli. «Skvoz’ sumrak bytija» // «Zvezda», 2002]

b.

Nu	čto	ja	mog-l-a	ej	vozrazi-t’?
well	what.ACC	I-NOM	can-PST-F.SG	she.DAT	object.PFV-INF
Čto —	bol’šinstv-o	podavljajušč-ee!	— xirurg-ov	**operiruj-ut**	besplatno?
that	majority-NOM.SG	overwhelming-N.NOM.SG	surgeon-GEN.SG	**operate.IPFV-PRS.3.PL**	for.free

‘Well, what could I say to persuade her otherwise? That the vast majority of surgeons **operate** for free?’

[I. Grekova. Perelom, 1987]

Some prefixed motion verbs present a third type of homonymy, since they have two interpretations that are differentiated both by aspect and by semantics (involving two separate verbs), as in *s-xož-u* [roundtrip-walk.pfv-fut.1.sg] ‘I will go and come back’ vs. in *s-xož-u* [down-walk.ipfv-prs.1.sg] ‘I am going down’ which is also frequently part of the idiom *s-xodi-t’ s uma* [literally ‘walk down from mind’] meaning ‘go crazy’. Only the first verb expresses future tense, illustrated in (4a); (4b) expresses present tense.

(4) a. Cirk — èto detsk-ie vospominanij-a i
circus.NOM.SG that childhood-NOM.PL memory-NOM.PL and
položitel’n.ye èmocii! Objazatel’no **sxožu**.
positive-NOM.PL emotion-NOM.PL definitely **roundtrip-walk.PFV-FUT.1.SG**
‘The circus is childhood memories and positive emotions! I **will** definitely **go**.’
[kollektivnyj. Forum: Poxod v cirk, 2010]

b. Ja ponimaj-u, čto potixon’ku **sxožu** s um-a.
I.NOM understand.IPFV-PRS.1.SG that slowly **down-walk.IPFV-PRS.1.SG** from mind-GEN.SG
‘I understand that I’**m** slowly **losing** my mind.’
[Sati Spivakova. Ne vsë, 2002]

The homonymies described above are to some extent mitigated in the disambiguated portion of the RNC, however, manual exploration of this data turns up too much noise to allow for precise measures.

Worse still is the problem of the periphrastic imperfective future, which allows both for insertion of words and various orderings of words, and additionally is confounded by the existence of phrases that can “masquerade” as future forms. For example, in *bud-et snova spraši̭va-t’* [be.fut-3.sg again ask.ipfv-inf] ‘s/he will ask again’ there is the adverb *snova* ‘again’ between the auxiliary and infinitive parts of the future, and (5) shows that it is possible to have not just one intervening word form but many; in this example there are five (including the two clitics *li* ‘whether’ and *že* ‘emphasis’). It is also possible to invert the order of the auxiliary and the infinitive, as in *spraši̭va-t’ bud-et* [ask.ipfv-inf be.fut-3.sg] ‘s/he will ask’.

(5) I **bud-et** li ona mne tak že

and **be.FUT-3.SG** whether she.NOM I.DAT so emphasis

nravi-t'-sja ili net — ja ne znaj-u

please.IMPF-INF-REFL or not I.NOM not know.IMPFV-PRS.1.SG

'And whether I **will like** her as well or not – I do not know.'

[Evgenij Griškovec. OdnovrEmEnno, 2004]

Furthermore, (6) shows that we can encounter multiple intervening words also when we have the reverse word order, with the infinitive first, intervening words, and then the auxiliary verb.

(6) Tak ja dumaj-u, a **sprašiva-t'**,

so I.NOM think.IMPFV-PRS.1.SG but **ask.IPFV-INF**

požaluj, ni u kogo ne **bud-u**.

probably not by who.GEN not be.FUT-1.SG

'I think so, but I **will** probably not **ask** anyone.'

[Alla Bossart. Povesti Zajceva // «Družba narodov», 1998]

One can also find future expression of modals that govern infinitives, yielding both word-order options, as in *možno bud-et sprašiva-t'* [possible be.fut-3.sg ask.ipfv- inf] 'it will be possible to ask' and *sprašiva-t' bud-et možno* [ask.ipfv-inf be.fut-3.sg possible] 'it will be possible to ask'; both word orders are found in (7).

(7) Ved' kogda-nibud' — on obešča-l — **sprašivat'**

after.all someday he.NOM promise.PFV-PST.M.SG ask.IPFV-INF

budet **možno,** **možno** **budet** **sprašivat'**!

be.FUT-3.SG possible possible be.FUT-3.SG ask.IPFV- INF

'After all, someday – he promised – it **will be possible to ask, to ask will be possible**.'

[Dina Rubina. Russkaja kanarejka. Bludnyj syn, 2014]

These modal expressions look like imperfective future forms of the verb *sprašiva-t'* 'ask', but this is not the case. The future form of *byt'* 'be' in these examples is not the auxiliary of the periphrastic future but instead signals the tense that applies to the modal expressions with *možno* 'possible'. Examples like these of future forms of *byt'* 'be' that just happen to be collocated with an imperfective infinitive are common in Russian, and there is no automatic way to disambiguate them in a corpus.

4.2 A sample to represent the overall incidence of future forms

Due to the challenges presented by homonymy and non-contiguity of periphrastic future forms, we have opted to select a group of ten high-frequency perfective and imperfective verb pairs (represented in Table 5 and Figure 1), in order to undertake a targeted study in which we manually check the examples to be certain that we include all and only the future forms of the verbs. We used frequency, plus morphological and semantic criteria to select this set of verbs. For all of these verb pairs, both the perfective and the imperfective verbs appear at a rate of over 100 total attestations (including all inflected forms) per million words (ipm) in the disambiguated subcorpus of the Russian National Corpus.

Since our sample of ten verb pairs does not include either biaspectual verbs or prefixed verbs of motion, the only homonymy that is problematic is the type involving imperative vs. indicative forms. Three of the ten perfective verbs in Table 5 have second person plural future forms that are homonymous with imperatives, namely: *sprosite* 'ask', *polučite* 'receive', and *posmotrite* 'look'. All of the attestations of these forms found in the disambiguated RNC were analyzed by hand to determine which of them were truly future forms, and those future forms were added to the total numbers of all other future forms for those three verbs, thus giving accurate counts. The data in the rightmost column of Table 5 and in Figure 1 are thus based on the total number of perfective future forms adjusted to disambiguate them from imperatives.

Verb pair: Perfective / Imperfective	Gloss	Morphological marking of aspect	Total frequency Perfective / Imperfective	Frequency of past tense forms Perfective / Imperfective	Adjusted estimate of future forms Perfective / Imperfective
u-videt' / *videt'*	'see'	"A": prefixed perfective / primary imperfective	124 683 / 322 185	64 819 / 109 050	20 756 / 1 647
po-dumat' /	'think'	"A": prefixed perfective /	83 115 /	50 524 /	8 023 /

dumat'		primary imperfective	230 969	78 741	2 063
po-smotret' / smotret'	'look'	"A": prefixed perfective / primary imperfective	80 525 / 189 804	38 309 / 62 036	17 914 / 1 455
na-pisat' / pisat'	'write'	"A": prefixed perfective / primary imperfective	96 192 / 146 918	33 836 / 53 850	7 637 / 2 201
s-prosit' / s-praš-iva-t'	'ask'	"B": prefixed perfective / secondary imperfective	166 207 / 61 260	141 176 / 20 462	9 208 / 690
ponjat' / ponim-at'	'understand'	"B": prefixed perfective / secondary imperfective	136 150 / 139 109[26]	73 194 / 31 310	13 023 / 277
prinjat' / prinim-at'	'accept'	"B": prefixed perfective / secondary imperfective	118 645 / 60 591	42 031 / 18 048	9 657 / 375
dat' / da-vat'	'give'	"C" perfective simplex / secondary imperfective	286 575 / 143 974	73 641 / 29 680	35 578 / 779
polučit' / poluč-at'	'receive'	"C" perfective simplex / secondary imperfective	152 984 / 50 738	68 760 / 13 049	15 434 / 1 145
vzjat' / brat'	'take'	suppletive	170 655 / 65 231	73 528 / 15 832	20 756 / 655

Table 5. Sample of verb pairs that demonstrate relative frequencies of perfective and imperfective future forms in Figure 1.

For each imperfective verb in Table 5, a sample of 100 attestations of infinitive forms was extracted and analyzed to determine the rate of genuine future forms, taking into account various word orders and discontinuous periphrastic forms to arrive at an estimate of the percentage of genuine futures. This sample yielded a percentage of genuine futures that could then be applied to extrapolate a good estimate of the actual occurrence of periphrastic future forms for each imperfective verb. In most cases this increased the total number of imperfective futures that were identified, since we were

[26] The total frequencies for the verb pair meaning 'understand' deviates from the aggregate pattern reported for "B" pattern verbs in Table 3 in that there is virtually no difference in frequency. Note, however, that the past tense frequencies for this verb pair do reflect the overall tendency for "B" pattern perfective verbs to be of higher frequency than imperfective verbs.

able to include all examples regardless of how many intervening words separated the auxiliary from the infinitive. Overall, our targeted survey shows that the disambiguated RNC tends to underreport the number of both perfective and imperfective future forms.

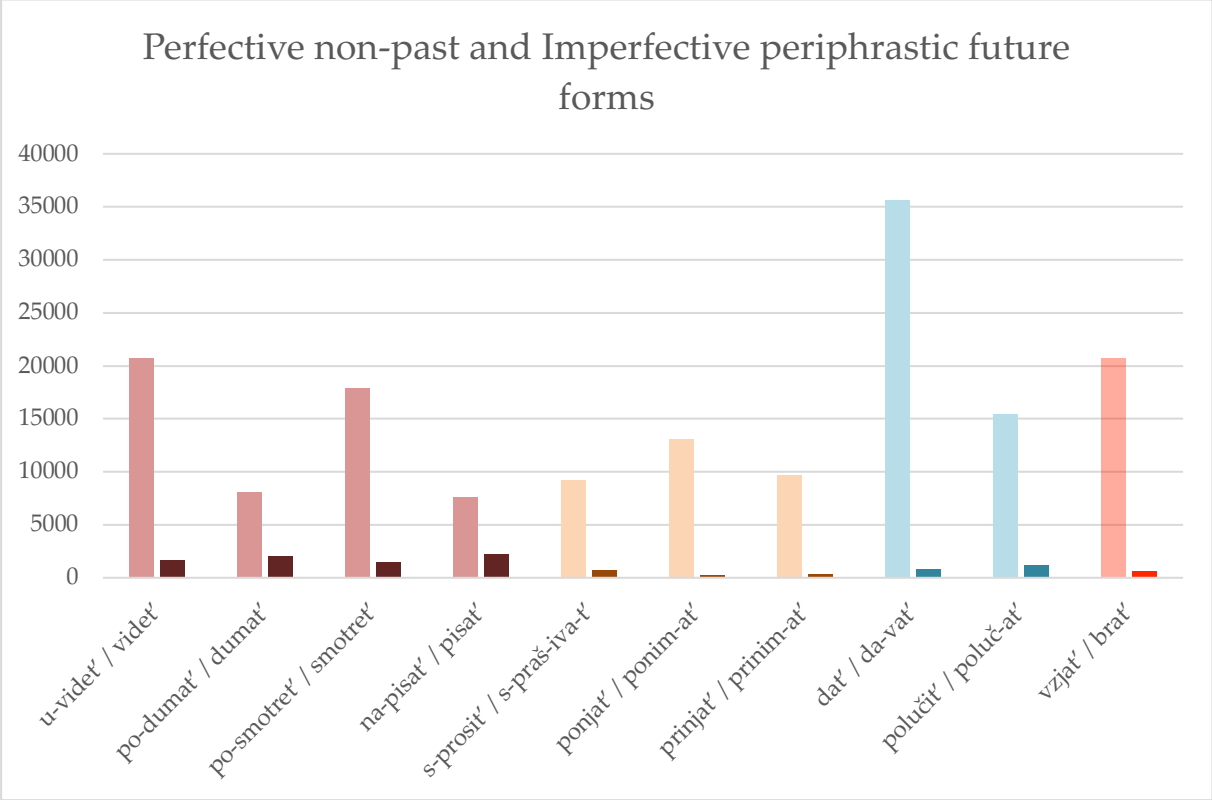

Figure 1. Visualization of data in rightmost column of Table 5: sample of high-frequency verb pairs showing the distribution of perfective non-past and imperfective periphrastic future forms for each verb pair.

Figure 1 visualizes the adjusted estimate of future forms as indicated in the rightmost column of Table 5. The "A" pattern verb pairs are shown in brown, the "B" pattern in green, the "C" pattern in blue, and the suppletive verb pair in red. Within each pair, the lighter hue indicates the perfective future forms, while the darker hue indicates the imperfective future forms. The main point of Figure 1 is to show that the frequency of perfective future forms far exceeds that of imperfective future forms. On average across our ten verb pairs, the perfective future makes up 11.44% of the attestations of perfective verbs, whereas the imperfective future makes up 0.94% of the attestations of imperfective verbs. For our ten verb pairs, the total number of perfective future forms is 157 986, whereas for imperfectives we find 11 287 future forms. The overall estimated ratio of perfective future forms to imperfective future forms is 14:1. The type of aspectual morphology ("A" pattern, "B" pattern, "C" pattern, suppletive) does

not influence this effect. For each individual verb pair, the frequency of perfective futures is many times higher than the frequency of imperfective future forms, and in aggregate the difference is one of an order of magnitude.

4.3 Evaluating perfective vs. imperfective markedness in the Russian future

The difference in frequency between perfective and imperfective future forms is very strong. In order to evaluate frequency differences for the comparisons made in Section 3, we needed statistical tests, and the effect sizes even where reportable were small or approaching medium at best. With respect to future tense forms, the size differences are large: they are not a matter of percentage points but of multiples. And the frequency of future forms aligns with their morphological complexity, as displayed in Table 6.

Aspect	Morphological complexity of future form	Frequency ratio
perfective	conjugated verb form	14
imperfective	conjugated auxiliary verb form + infinitive	1

Table 6. Parameters indicating markedness for aspect in the future tense.

The parameters in Table 6 strongly support the conclusion that for the future tense, perfective is unmarked (less morphologically complex and higher frequency), while imperfective is marked (more morphologically complex and lower frequency). This is a striking conclusion because it is the opposite of the prevailing opinion cited in Section 3 that perfective is marked and imperfective is unmarked.

5. Conclusion

Markedness, also known as encoding asymmetry, is a pervasive fact of language in which three parameters tend to align: expectedness of meaning, complexity (of both form and meaning), and frequency. Observations of markedness span linguistic traditions, and due to the relevance of both the form-meaning relationship and frequency, markedness is highly relevant for cognitive linguistics as a usage-based framework.

In Russian, aspect is expressed at the level of the verb as perfective or imperfective. It is traditionally assumed that the markedness of Russian aspect runs counter to that of other languages with a perfective vs. imperfective distinction, namely that in Russian perfective is marked and imperfective is unmarked. However, overall frequency data is inconclusive: we do not find support for category-level markedness of perfective vs.

imperfective. At the local levels of lexemes, we find more convincing alignments of morphological complexity and frequency that indicate that perfective can behave both as marked and as unmarked. Three patterns of morphological coding of aspect all show alignment of the parameters, though they don't all point to the same markedness values. The "A" pattern has higher morphological complexity for the perfective, which also is of lower frequency, suggesting that perfective is marked. The "B" and "C" patterns have lower morphological complexity for the perfective, which is also of higher frequency, suggesting that perfective is unmarked. But even at these local levels, the significant effect sizes of frequency differences are small.

We take the analysis one step further by examining the encoding asymmetry in the future tense. The future tense in Russian is special for two reasons. First, there is a consistent difference in morphological complexity with respect to aspect: perfective future forms are simply conjugated forms of the verb, and imperfective future forms are more complex, consisting of a conjugated auxiliary verb plus an infinitive. Second, there are many hurdles to measuring the frequency of the future tense in Russian due to confounding factors presented by homonymy, word order, and non-contiguous forms. We present a methodological solution involving the stratification of a sample of verbs according to aspectual morphology, sampling, manual examination of thousands of forms, and extrapolation. This yields the first reasonably accurate estimate of the real incidence of perfective and imperfective future forms and the discovery that perfective future forms are about fourteen times more frequent than imperfective future forms. This measurement supports a remarkable conclusion, namely that within the future tense, perfective is consistently unmarked while imperfective is marked. This conclusion is the opposite of the majority opinion in traditional Russian linguistics.

In sum, we offer support for the theoretical position of Tiersma (1982) that markedness must be understood primarily at the local level. We contribute to the understanding of markedness within the usage-based framework of cognitive linguistics (cf. Diessel 2019: Chapter 11) with a detailed illustration of a case study of the encoding asymmetries presented by Russian aspect and how these asymmetries pattern with relative corpus frequency. We make a methodological contribution to the solution of a difficult issue in determining corpus frequency. And we present the first accurate description of the relative frequency of the Russian future tense for both perfective and imperfective verbs.

References

Andersen, Henning. 1989. Markedness theory – the first 150 years. In Olga M. Tomić (ed.) *Markedness in synchrony and diachrony*. Berlin: Mouton de Gruyter, 11–47.

Andersen, Henning. 2001. Markedness and the theory of language change. *Current Issues in Linguistic Theory* 219: 21–57.

Avilova, Natal'ja S. 1976. *Vid glagola i semantika glagol'nogo slova.* Moscow: Akademija nauk SSSR.

Battistella, Edwin L. 1990. *Markedness: The evaluative superstructure of language.* Albany: SUNY Press.

Battistella, Edwin L. 1996. *The Logic of Markedness*. New York/Oxford: Oxford University Press.

Binnick, Robert I. 1991. *Time and the Verb: A Guide to Tense and Aspect*. New York/Oxford: Oxford University Press.

Bondarko, Aleksandr V. 1971. *Vid i vremja russkogo glagola*. Moscow: Prosveščenie.

Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1983. Markedness, grammar, people, and the world. In Fred R. Eckman, Edith A. Moravcsik, Jessica R. Wirth (eds.) *Markedness*. New York: Plenum Press, 85–106.

Comrie, Bernard. 1989. *Language universals and linguistic typology*, 2nd edition. Blackwell.

Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.

Dickey, Stephen M. 2000. *Parameters of Slavic Aspect*. A Cognitive Approach. Stanford: CSLI Publications.

Diessel, Holger. 2019. *The Grammar Network: How Linguistic Structure is Shaped by Language Use.* Cambridge: Cambridge University Press.

Durst-Andersen, Per. 1992. *Mental Grammar: Russian Aspect and Related Issues*. Columbus, OH: Slavica.

Endresen, Anna and Vladimir Plungian. 2011. Russian morphophonemics in a nutshell: The verb *vstat'* 'stand up'. *Poljarnyj vestnik* 14, 22–44.

Fábregas, Antonio and Laura A. Janda. 2019. When Russian is more perfective than Spanish. *Mundo Eslavo* 18: 20–39.

Galton, Herbert. 1976. *The Main Functions of the Slavic Verbal Aspect.* Skopje: Macedonian Academy of Sciences and Arts.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42: 25–70.

Haspelmath, Martin and Andres Karjus. 2017. Explaining asymmetries in number marking: Singulatives, pluratives, and usage frequency. *Linguistics* 35: 1213–1235.

Isačenko, Aleksandr. 1960. *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim*. Bratislava: Izdatel'stvo akademii nauk.

Jakobson, Roman O. 1971[1932]. Zur Struktur des russischen Verbums. In *Selected Writings II*. The Hague: Mouton, 3–15.

Jakobson, Roman O. 1971[1957]. Shifters, verbal categories, and the Russian verb. In *Selected Writings II*. The Hague: Mouton, 130–147.

Janda, Laura A. 1995. Unpacking Markedness. Mouton de Gruyter, 207–233.

Janda, Laura A. 2007. What makes Russian Bi-aspectual verbs Special. Mouton de Gruyter, 83–109.

Janda, Laura A. and Olga Lyashevskaya. 2011. Grammatical profiles and the interaction of the lexicon with aspect, tense and mood in Russian. *Cognitive Linguistics* 22 (4): 719–763.

Janda, Laura A. and Antonio Fábregas. 2019. Seeing from without, seeing from within: Aspectual differences between Spanish and Russian. *Cognitive Linguistics* 30 (4): 687–718.

Lakoff, George. 1987. *Women, fire, and dangerous things*. Chicago: University of Chicago Press.

Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction.* Oxford: Oxford University Press.

Lewandowska-Tomaszczyk, Barbara. 2007. Polysemy, prototypes, and radial categories. In *The Oxford handbook of cognitive linguistics*, eds D. Geeraerts and H. Cuyckens (Oxford and New York: Oxford University Press), 139–169.

Maslov, Jurij S. 1965. Sistema osnovnyx ponjatij i terminov slavjanskoj aspektologii. *Voprosy obščego jazykoznanija*, 53–80.

Mayerthaler, Willi. 1980. Ikonismus in der Morphologie. *Zeitschrift für Semiotik* 1: 19–37.

Padučeva, Elena V. 1996. *Semantičeskie issledovanija*. Moscow: Jazyki russkoj kul'tury.

Prince, Allen and Paul Smolensky. 2008. *Optimality theory: Constraint interaction in generative grammar.* Chichester: Wiley.

Rosch, Eleanor. 1973a. Natural categories. *Cognitive Psychology* 4: 328–350.

Rosch, Eleanor. 1973b. On the internal structure of perceptual and semantic categories. In Terence E. Moore (ed.) *Cognitive development and the acquisition of language*. New York: Academic Press, 111–144.

Smith, Carlota S. 1991. *The Parameter of Aspect*. Dordrecht: Kluwer.

Tiersma, Peter Meijes. 1982. Local and general markedness. *Language* 59: 832–849.

Timberlake, Alan. 1982. Invariance and the syntax of Russian aspect. In Hopper, Paul (ed.), *Tense-Aspect: Between Semantics and Pragmatics*. Amsterdam: John Benjamins, 305–331.

van Langendonck, Willy. 1989. Ergativity, markedness and prototypes. *Belgian Journal of Linguistics* 4, 173–183.

van Schooneveld, Cornelis H. 1978. *Semantic Transmutations: Prolegomena to a Calculus of Meaning*, vol. 1: *The Cardinal Semantic Structure of Prepositions, Cases, and Paratactic Conjunctions in Contemporary Standard Russian*. Bloomington: Physsardt.

Vinogradov, Viktor V. 1972. *Russkij jazyk*. 2nd ed. Moscow: Vyssšaja škola.

Wierzbicka, Anna. 1967/2018. On the semantics of the verbal aspect in Polish. In: *To honor Roman Jakobson: essays on the occasion of his 70. birthday, 11. October 1966: Vol. 3*, Berlin, Boston: De Gruyter Mouton, 2018, 2231–2249. https://doi.org/10.1515/9783111349138.

Zaliznjak, Anna A. and Aleksej D. Šmelev. 2000. *Vvedenie v russkuju aspektologiju.* Moscow: Jazyki russkoj kul'tury.

II Looking into the Russian future

Daria Kosheleva, Laura Alexis Janda

UiT, The Arctic University of Norway

The relationship between future time and future tense forms in Russian is complex. The forms traditionally attributed to the future tense in certain cases do not refer to future time. Those cases have been previously presented as an inventory not representing a plausible cognitive model and/or attributed to the sphere of modality. In this article, we suggest a data-driven approach applied to the spectrum of meanings of Russian future tense forms. We analyzed corpus data and discovered that 44% of perfective future forms and 22% of imperfective future forms do not unambiguously express future time meaning. Among the non-future time meanings that Russian future tense forms can express are Gnomic, Performative, Implicative, Hypothetical, Alternation, and Stable scenario. Furthermore, we propose that the meanings of the future tense constitute a radial category. Future time reference is the prototypical meaning of the future tense. The remaining meanings comprise extensions connected to the prototypical meaning. We describe the radial category with reference to Langacker's (2008) model of tense and potentiality. Additionally, we explore the interaction of future tense and modality.

Keywords: Russian, future tense, aspect, corpus, radial category

La relation entre temps réel et temps grammatical est complexe en Russe. Les formes grammaticales, traditionnellement attribuées au temps grammatical du futur, ne correspondent pas, dans certains cas, au temps réel du futur. Par le passé, ces cas ont été présentés comme attribués à la modalité. Dans cet article, nous proposons une approche basée sur l'usage, appliquée à la palette de sens des formes grammaticales du futur en Russe. Nous avons analysé des données issues d'un corpus et avons découvert que 44% des formes perfectives du futur et 22% des formes imperfectives du futur n'expriment pas le temps réel du futur avec certitude. Parmi les sens non-futurs que le temps grammatical du futur en Russe peut exprimer nous trouvons le Gnomisme, le Performatif, l'Implicatif, l'Hypothétique, l'Alternance, et le Stable. En outre, nous proposons que les sens du temps grammatical du futur constituent une catégorie de nature radiale. La référence au temps réel du futur est le sens prototypique du temps grammatical du futur. Les autres sens sont des extensions sémantiques liées au sens prototypique. Nous décrivons la catégorie radiale en référence au modèle développé par Langacker (2008). Par ailleurs, nous explorons l'interaction du temps grammatical du futur et de la modalité.

Mots-clés: russe, futur, aspect, corpus, catégorie radiale.

1. Introduction

While the primary function of Russian future tense forms is to refer to events in future time, to a large extent (44% for perfective verbs and 22% for imperfective verbs) future tense forms are used to refer to events that are not unambiguously located in future time. Russian future tense forms can express additional or different meanings, for example, Gnomic, Hypothetical, or Performative. From the perspective of cognitive linguistics, we analyze the use of future tense forms in a database consisting of 1000 perfective and 1000 imperfective examples. We show that meanings that diverge from future time reference are not sporadic, but regular and related.

Normally, events located in future time are refered to by means of forms of the future tense. In Russian there are two forms that can be identified as future tense, differentiated by aspect. The relationship between these forms, aspect, and present tense are presented in Table 1.

Aspect\Tense	Present Tense	Future Tense
Imperfective Aspect	*piš-et* write.IPFV-PRS.3.SG 's/he writes'	*bud-et pisa-t'* be.FUT-3.SG write.IPFV-INF 's/he will write'
Perfective Aspect	–	*napiš-et* write.PFV-FUT.3.SG 's/he will write'

Table 1. Present and future tense forms of Russian verbs.

For imperfective verbs, the future tense form consists of the verb *byt'* 'be' in the future tense combined with the infinitive of the imperfective verb. The imperfective future tense form can also be called complex, periphrastic, or analytical. In terms of inflectional morphology, the perfective future form is morphologically identical to the imperfective present: compare the inflectional endings *piš-et* 's/he writes' and *napiš-et* 's/he will write'. It is because of this morphological identity that the term "non-past" is

often used to describe both the imperfective present tense and perfective future tense forms. The perfective future form is referred to as synthetic or simple.

The Russian future tense forms have other uses, in addition to reference to the future time. Moreover, as shown in (Kosheleva & Janda Ms.), there are quantitative differences. The perfective future tense forms occur in the Russian National Corpus (RNC) 14 times more often than the imperfective future tense forms. In this article, within the framework of cognitive linguistics, we will dwell in more detail on the various meanings of the future tense forms, based on data from the Russian National Corpus and show that these meanings form a radial category.

1.1. Russian future tense in Langacker's framework

We base our understanding of Russian future tense forms on Langacker's model of tense (2008: 301) that consists of Conceived Reality, Reality, Current Reality, Immediate reality, and Non-reality. Future is a projection forward in time from the perspective of Immediate reality. We propose that Russian non-future uses of the future tense can be explained by observing which of the realities and non-realities are adjacent to each other and how they relate to each other in Langacker's model.

Figure 1 (adapted from Langacker 2008: 301; 306) provides a schematic model of tense and potentiality. The various elements of Figure 1 situate future tense uses and their possible modalities. Langacker's model accounts for the relatively solid grounding of past and present in Conceived and Immediate Reality as opposed to the more tenuous grounding of future in Projected Reality. We claim that this difference in grounding motivates the polysemy of the future tense forms that we observe in Russian.

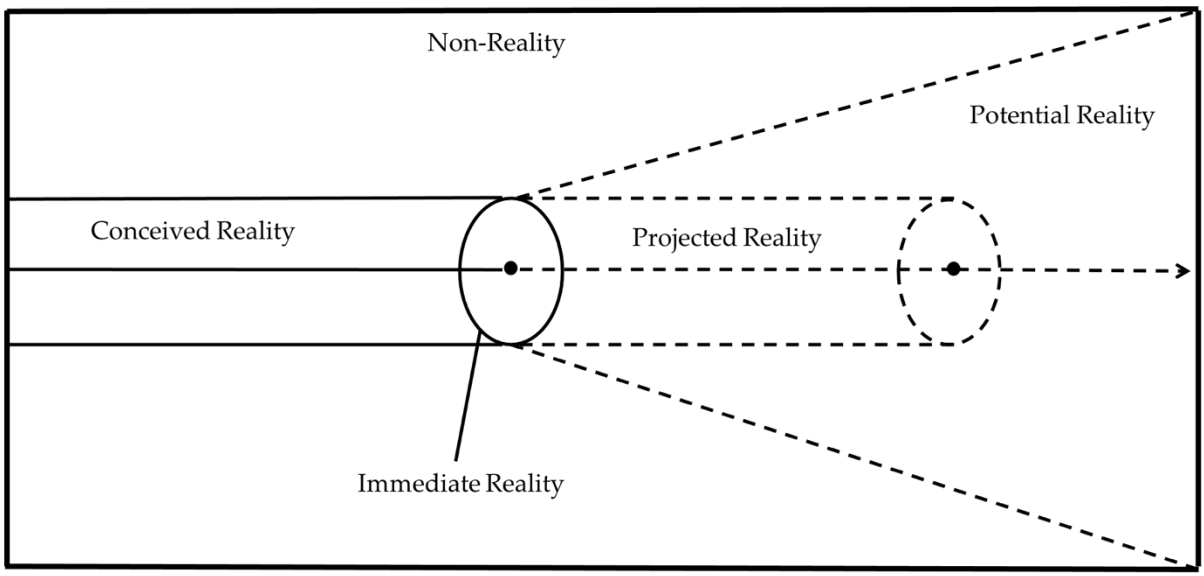

Figure 1. Model of tense and potency. Source: Adapted from (Langacker 2008).

The terms Conceived Reality, Immediate Reality, Projected Reality and Non-Reality refer to different parts of the time-space continuum that is perceived by the speaker.

In the model we see a cylinder that grows through time from past (left) through the present (middle point with a circle) to the future (with dotted lines on the right). The past is represented by Conceived Reality: a speaker (conceptualizer) develops a "version" or conception of reality that is stored in the speaker's mind. The present is situated in Immediate Reality. The cylinder of Projected Reality on the right is the expected future, corresponding to the future meaning of future tense forms. In addition, on the right part of the diagram, there is a cone extending from the present to the future which represents Potential Reality. Potential Reality overlaps with the domain of Non-reality. The cone that expands from the present includes both the cylinder of Projected Reality and other potential but not necessarily expected events in the future. The part of the cone that extends beyond the cylinder represents Potential Reality and is expected to comprise modal meanings expressed by future forms.

Projected Reality is what is expected to happen in the future, and Potential Reality is what could happen in the future. Thus, in these terms, future time coexists with the area of modal influence. In other words, there is a gray zone where there is no firm boundary between the future and Potential Reality (and therefore modality).

In contrast to past and present time, the future is less anchored to an embodied experience. Future tense can be used to speculate about events that may or may not happen, and this speculation becomes increasingly tenuous when we shift from proximate events to ones further removed in time. This lack of anchoring furthermore makes it possible to use future tense for statements that have no connection to time at all, being rather observations of general gnomic truths.

Four adaptations of Langacker's model account for the various uses of future tense forms that we find in Russian, as diagrammed in Figure 2a-d. In Figure 2a Projected reality is profiled (shaded in blue), corresponding to the Future time uses of Russian future tense. Figure 2b profiles both Projected and Potential Reality (shaded in yellow), corresponding to a use of Russian future tense that we term Extended Future because reference is not limited to Projected Reality. Profiling in Figure 2c is limited to a small section of both Projected and Potential Reality (shaded in green), termed Directive because it is used for imperative commands that should be carried out in the

near future (though the outcome is not guaranteed). The most comprehensive profiling is in Figure 2d, which encompasses everything except non-reality. We term this use Gnomic because it refers to eternal facts and consequences. This extension of Langacker's theoretical model is potentially portable to languages in general.

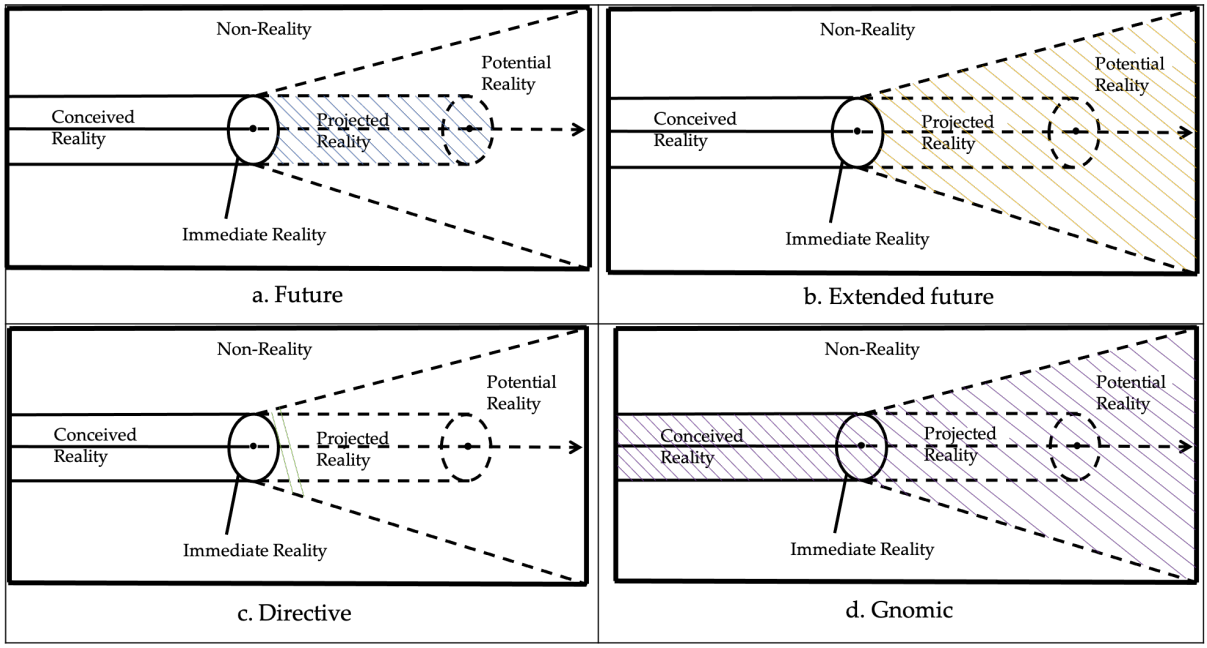

Figure 2a-d. Adaptation of Langacker's model for Russian future tense. Source: Adapted from (Langacker 2008).

In Section 3 we present a comprehensive analysis of the uses of Russian future tense forms that express both future time and non-future time meanings. Our analysis integrates and expands upon previous scholarship, described in Section 1.2 below.

1.2. Non-future uses of the future tense forms in Russian: previous scholarship

Scholars have previously remarked that the morphological forms associated with the future tense in Russian do not always refer to events located in future time. Here we provide an overview of relevant previous scholarship. Unfortunately, each scholar uses a different set of terms for non-future time meanings of Russian future tense forms. To achieve a coherent overview, we use a consistent terminology that we elaborate in more detail in Section 3. According to our system, Extended Future and Gnomic uses overlap with subtypes termed Alternation, Hypothetical, and Implicative. In addition, Extended Future encompasses Performative and Posterior uses (where the reference time of Immediate Reality is shifted back), and Gnomic encompasses Habitual chain and Stable Scenario uses.

Maslov (1990/2004: 515-516) mentions that the future tense (both perfective and imperfective) can be used in a figurative sense to express habitual occurrences. Isačenko (1965/2003: 451) considers separately the meanings of the analytical future and 'simple' future. For the 'simple' form, he uses the term 'perfective present' to highlight the nature of the non-future meanings, both Habitual and other types. As for the analytical future form, Isačenko (1965/2003: 445) says that the analytical future tense usually does not have any additional meaning except for two modal nuances: the unreality of an unfulfilled action (1) and the modal expression of will (2).

(1)	Naprasno	vy	**bud-ete**	**iska-t′**	xot′
	in.vain	you.NOM	be.FUT-2.PL	seek.IPFV-INF	even
	v	odn-om	lic-e	sled-ov	suetlivost-i.
	in	one-N.LOC.SG	person-LOC.SG	trace-GEN.PL	fussiness-GEN.SG

'In vain you **will look** for traces of fussiness in one person.' L. Tolstoy.

(2)	**Bud-u**	ja	s	nim	**vozi-t′-sja**!
	be.FUT-1.SG	I.NOM	with	he.INS	convey-INF-REFL

'I **do** not **want to mess** with him!'

Vinogradov (1947: 463) is in solidarity with Isačenko regarding the analytical form. He notes that in rare cases the future meaning is veiled by a modal shade of indefinite duration, extending into the span of future time (3).

(3)	Nu,	voz′m-i	svo-i	tri	s
	well	take.PFV-IMP.2.SG	own-ACC.PL	three.ACC	with
	polovin-oj,	čto	ty	**bud-eš′**	**dela-t′**?
	half-INS.SG	what.ACC	you.NOM	be.FUT-1.SG	do.IMPFV-INF

'Well, take your three and a half, what **can** you **do**?'

For the perfective future tense, the most striking non-future time meaning is Habitual. In Maslov's work (1990/2004: 521) habituality is manifested through indications of multiple repetition. Isačenko (1965/2003: 451) likewise presents habituality as

repetitive processes and events sometimes introduced by *byvalo/byvaet* ‘it happened/s’ followed by a perfective future tense form. In addition, habituality expressed by perfective future tense forms can be accompanied by other meanings and constructions, including conditional, concessional with a generalized personal meaning, alternation, constructions like *net-net da i* ‘from time to time’ and *voz'mët da i* ‘suddenly’, and reduplication. Vinogradov (1947: 467) finds habituality in chains of future tense verb forms depicting present time, which can also be interpreted as what we call a Salient event (cf. 4; for the definition cf. subsection 3.4.6).

(4)	Živ-ëm	v	odn-om	gorod-e,	počti	
	live.IPFV-1.PL	in	one-M.LOC.SG	city-LOC.SG	almost	
	rjadom,	a	**uvid-iš′-sja**	raz	v	nedel-ju.
	nearby	but	see.PFV-FUT.2.SG-REFL	time.ACC	in	week-ACC.SG

‘We live in the same city, almost nearby, but people **see** each other once a week.’
A. Ostrovsky. Groza.

Both Maslov and Isačenko recognize the type of use that we term Stable Scenarios:

(5)	Čto	**pose-eš′**	to	**požn-eš′**
	what.ACC	sow.PFV-FUT.2.SG	that.ACC	reap.PFV-FUT.2.SG

‘What you **sow**, you **reap**.’

Isačenko presents the Alternation use as a special case of Habitual used when describing a chain of events with the conjunction *to…to…* ‘sometimes X, sometimes Y’:

(6)	To	zajac	**proskoč-et**,	to	**projd-et**	rys′.
	then	hare.NOM.SG	hop.by.PFV-FUT.3.SG	then	pass.through.PFV-FUT.3.SG	lynx.NOM.SG

‘Sometimes a hare **will hop by**, sometimes a lynx **will pass through**.’

Vinogradov (1947: 469) shows Alternation in a slightly different context, as an instance of repetition in the past, though often about possible or habitual actions:

(7)	On	to	**vojd-et**,	to	**vyjd-et**
	he.NOM	then	then go.in.PFV-FUT.3.SG	then	go.out.PFV-FUT.3.SG
	iz	komnat-y	(tak-oj	by-l	neposed-a).
	from	room-GEN.SG	such-M.NOM.SG	be-PST.M.SG	fidget-NOM.SG

'He **would go in** and **out** of the room (he was such a fidget).'

Hypothetical examples behave similarly to Alternations: Isačenko (1965/2003: 453) connects them to habituality (8), while Vinogradov (1947: 469) places them in past contexts (9).

(8)	Utrom	ne	**kup-iš′** –	
	in.morning	not	buy.PFV-FUT.2.SG	
	k	večer-u	vse	**razojd-et-sja**.
	toward	evening-DAT.SG	everything.NOM.SG	disperse.PFV-FUT.3.SG-REFL

'(If)you **don't buy** (it) in the morning, it **will be gone** in the evening.'

(9)	Nača-l	tatar	pokolačiva-t′:	**maxn-et**	
	begin.PFV-PST.M.SG	tatar.ACC.PL	beat.up.IPFV-INF	INF wave.PFV-FUT.3.SG	
	ruk-oj —	ulic-a,	**otmaxn-et**	nazad —	pereuloček.
	hand-INS.SG	street-NOM.SG	wave.off.PFV-FUT.3.SG	behind	alley.NOM.SG

'He began beating up Tatars : if he **waved** his hand (in one direction), a street (would be beaten), if he **waved** his hand in the other direction, an alley (would be beaten).'

Gnomic use of future tense, as well as near Performative use, are mentioned only by Vinogradov (1947: 468), who defines the Gnomic use as expressing a regular permanent result without any time limits:

(10)	I	už	èto	vsegda	**ub′-jut**
	and	already	it	always	kill.PFV-FUT.3.PL

t-ogo,	kto	naprašiva-et-sja.
that-M.ACC.SG	who.NOM	beg.IPFV-PRS.3.SG-REFL

‘(They) **will** always **kill** the one who begs.’ L. Tolstoi.

According to Vinogradov (1947: 467), a Near-performative expresses an immediately forthcoming action:

(11)					
	A	ja	vam	**skaž-u**,	čto…
	and	I.NOM	you.DAT	tell.PFV-FUT.1.SG	that

‘And I **tell** you that…’

Directive meaning is mentioned by both Maslov (as a substitute for the imperative mood) and Vinogradov (as categorical desire, demand, intention, or invitation): cf. examples (12) and (13) respectively.

(12)						
	Pojd-eš′	v	magazin	i	**kup-iš′**	xleb-a.
	go.PFV-FUT.2.SG	in	store.ACC.SG	and	buy.PFV-FUT.2.SG	bread-GEN.SG

‘(You **will**) **go** to the store and **buy** some bread.’

(13)		
	Nu,	**poplyv-em**.
	well	swim.PFV-FUT.1.PL

‘Well, let’s **swim**.’

Separately, it is worth noting the so-called potential (modal) uses found in Vinogradov and Isačenko. Vinogradov approaches potential modality as a possibility of accomplishment (1947: 467):

(14)				
	Ne	**priduma-ju**,	kak	vyj-ti
	not	think.up.PFV-FUT.1.SG	how	go.out.PFV-INF
	iz	èt-ogo	položeni-ja.	
	from	this-N.GEN.SG	situation-GEN.SG	

'I **cannot figure out** how to get out of this situation.'

We see that previous researchers have addressed the issue of non-future uses of the future tense in Russian, but as yet no precise measurement of this phenomenon has been undertaken. The portion of non-future time reference was estimated at around 1/3 of the perfective future verb forms in the pre-corpus era (cf. Forsyth 1970: 120). A corpus-based grammar of Russian provides a detailed description of various categories (order, prohibition, instruction, permission, performative use, near-modal use, opportunity, habituals, etc.) of non-future time uses of future tense forms (cf. Stojnova 2016b). These categories mostly apply to the perfective future tense forms. According to Stojnova (2016a), the imperfective future tense has two main non-future time functions: imperatives (hortative and jussive) and non-referential uses that are not related to the future. Stojnova (2016a: 248) also points out that there are marginal non-future uses of the imperfective future that could be described in the same terms as the perfective non-future uses. Stojnova[27]'s studies are very detailed and informative, but her comparative review is based on random samples of only 100 corpus hits for each aspect (perfective and imperfective future tense forms). We consider this amount of data to be insufficient since it is likely that some submeanings may remain out of sight.

We bring several new perspectives to scholarship on the Russian future by measuring the occurrence of future tense forms and subjecting them to thorough semantic analysis, and by comparing the behavior of perfective as opposed to imperfective future tense forms. We show that close to half of perfective future tense forms do not unambiguously express future time, and the same is true for almost one in four imperfective future tense forms. We also give a breakdown of what other meanings are expressed by future tense forms and how this differs across perfective vs. imperfective aspect. We find connections between the Russian material and Langacker's model of tense and potency (see Figure 1). We propose that future tense meanings constitute a radial category with the prototypical meaning of Future time and various extensions related to it.

1.3. Russian future tense and modality

Another matter that arises with respect to the future tense is its relationship with modality. Can it be argued that the future tense in Russian intersects with modality? And if so, to what extent? In the Russian linguistic tradition, we find that there is no

[27] For further discussion, we refer interested readers to additional works cited by Stojnova in 2016a and 2016b.

common opinion on this matter, so we must address a variety of accounts. Arutynova (2010: 10) states that the future is always interacting with modal categories because when people think about the future, they are trying to guess what will happen but do not always succeed. There are some who strongly believe that future tense is shaped by modality and can be interpreted as a manifestation of modality (Klimonow 2011). This idea corresponds to the part of Figure 1 where the future cylinder is located inside of Potential Reality. However, the rest of the cone is not in focus. By contrast, some view modality as an "independent" element in the system (Petruxina 2009; Petruxina & Li 2015) claiming that future tense can be used without modal meanings. In this case, the focus is on Projected Reality alone. In other words, the future is purely the extension of Langacker's cylinder absent the cone of Potential Reality.

Stojnova (2016a) suggests that it is possible to ascribe modal meaning to almost all (except for the habitual) non-prototypical uses of the future (especially perfective) tense: performatives, imperative-like constructions, generic uses. For the imperfective future, Stojnova adds that it is difficult to draw a line between the uses with and without modal connotations. Overall, according to Stojnova the following conditions facilitate modal interpretation: 1) negative context; 2) conditional context; 3) participants and/or situations with non-referential status. Stojnova's theory correlates with Langacker's model (cf. Figure 1). The future tense does not provide a very reliable connection to a specific moment of time: notice that the border of the Projected Reality is a dotted rather than a solid line. The above-mentioned conditions weaken the grounding of the situation in a specific moment of time even more and move the situation into the domain of Potential Reality.

In addition, a middle ground is represented by different interpretations of the future, modality, and their (partial) interaction (Radbil 2011; Wiemer et al. 2020). Those cases at least to some extent can be viewed from the perspective of the balance between the cylinder and the cone in Langacker's model (cf. Figure 1).

Before we move to the future tense meanings and their interaction with modality, we need to define the types of modality. Here we engage the works of Kratzer (1981), van der Auwera and Plungian (1998), Klimonow (2011), Petruxina and Li (2015), and Wiemer et al. (2020) to define epistemic, volitive, potential, and basic modality. In epistemic modality utterances, the external participant (i.e., the speaker) assesses the degree of reliability of a proposition. Volitive modality manifests the internal participant's desire (intention) to perform the action. Potential modality denotes the skills or abilities that allow the participant to perform the action. Basic modality expresses the attitude of the internal participant (i.e., the subject) to the action. By

contrast, Radbil (2011) does not distinguish between types of modality; he introduces a distinction between two types of future: "future as a fact" (i.e. no modal meaning) and "future as modality" (i.e. the confidence that the event will happen).

In Section 3, we investigate the extent to which perfective and imperfective future forms express future time meaning, and what else they express when they do not unambiguously express future. In addition, in Section 4, we examine the data from the above-mentioned articles that focus on Russian (Petruxina & Li 2015; Wiemer et al. 2020; Klimonow 2011; Radbil 2011) to determine what kinds of meanings discovered in our database from the RNC (cf. Section 2) co-occur with various types of modalities.

2. Database of Perfective and Imperfective Future

In this section we present the database that serves as the basis for our analysis. The database consists of two datasets of future tense forms of perfective and imperfective verbs in samples from the RNC.[28]

The first dataset is of perfective forms. Sentences containing perfective future tense forms were extracted from the RNC and pseudorandomized. The first one thousand examples were analyzed by hand. Thirty-nine examples were flagged as "noise" because they did not illustrate the perfective non-past; these examples were misidentified as future in the RNC annotation, but they are actually examples either of imperfective verbs or of biaspectual verbs in imperfective usage. An additional thirty-nine examples were drawn from the pseudorandomized data to bring the total to one thousand.

The second dataset contains examples of imperfective future forms extracted from the RNC. The imperfective future tense forms are the future form of the verb *byt'* together with the imperfective infinitive (see Table 1). In order to avoid the issues concerning periphrastic (and often non-contiguous) forms described in (Kosheleva & Janda Ms.), the corpus search was restricted to imperfective future forms consisting of '*budu* + infinitive' at a distance of 1. Using the same procedure as for the perfective dataset, the downloadable sample was pseudorandomized and then analyzed by hand. Even though we restricted the conditions of the search, ninety-five examples had to be excluded as noise, in most cases because the auxiliary verb was semantically attached

[28] All of the data and annotations described in this article can be accessed at https://doi.org/10.18710/MHWRGE.

to an adverbial rather than to an infinitive. An additional ninety-five examples were culled to bring the total to 1000.

The data was classified into four major semantic groups and further annotated for information pertaining to semantics and modality. The four major classes are Future, Extended future, Gnomic and Directive. The Future class describes examples that unambiguously express future time. In this case, the future tense is used to locate events in the future relative to Speech Time. The Extended future presents some uncertainty because it refers to events that can be anchored to the past and/or present. Gnomics refer to events that are not grounded in time. Directives refer to actions that are expected to be executed immediately after the utterance is pronounced. These classes are not completely autonomous: they are related to each other and to the prototype (the Future class). In addition, the examples from each class can bear additional properties (Stable scenario, Habitual chain etc.). Together they form the radial category presented in Section 5.

Figure 3 shows the distribution of examples from our sample across the four semantic classes for the perfective dataset. The biggest class is Future, then Extended future and Gnomic are of nearly the same size. Directives are the smallest class with only 12 perfective examples.

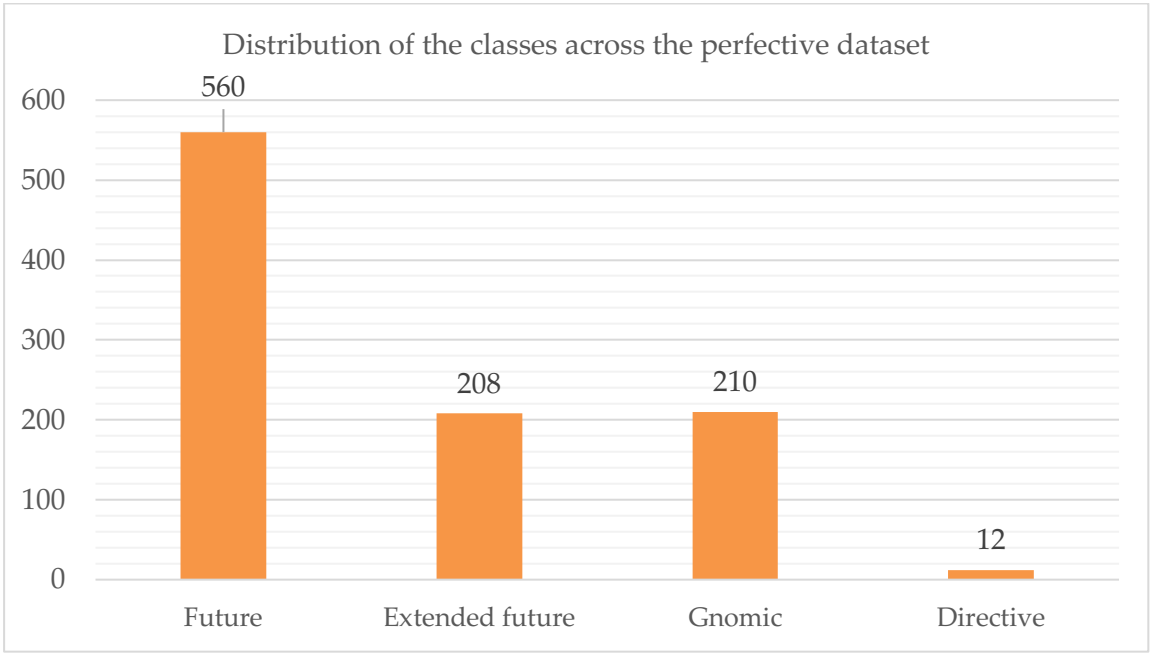

Figure 3. Distribution of semantic classes in the perfective dataset

The distribution of the examples across classes for the imperfective dataset is shown in Figure 4, which is organized similar to Figure 3.

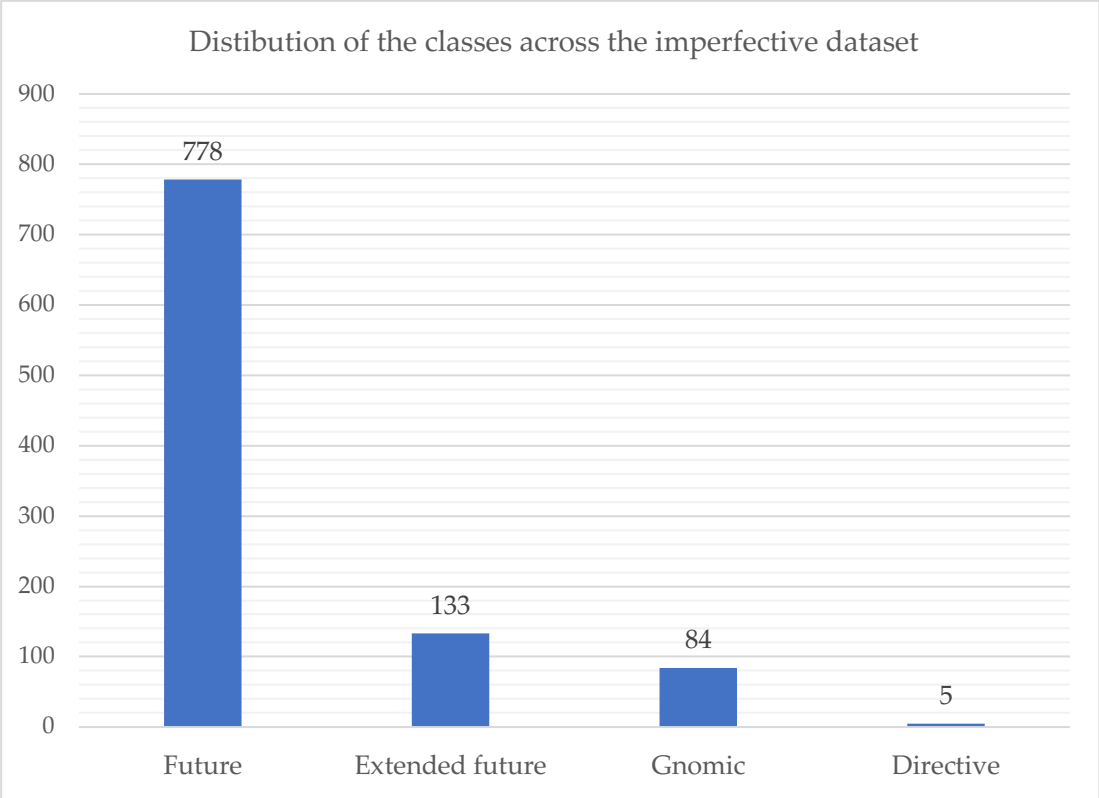

Figure 4. Distribution of semantic classes in the imperfective dataset

The Future class strongly dominates in the imperfective dataset. Extended future and the slightly smaller Gnomic classes are of comparable size. Directives constitute a minority. Each of the classes is described in more detail in Section 3.

3. Classification

In this section we examine the datasets in greater detail looking for individual clusters of examples that are similar to each other. Based on those clusters, we propose a model for the classification of future tense meanings in Russian. We propose that this classification forms a radial category. In Sections 3.1—3.5 we go through each class and describe the types of examples found in this class for both perfective and imperfective future verb forms.

3.1. Future time meaning

560 of 1000 examples of perfective future forms unambiguously express Future, as in (15).

(15) Let čerez dvesti-trista vse sam-o

year.GEN.PL across two.hundred-three.hundred everything.NOM.SG self-N

obrazu-et-sja, — uteša-l Čexov, i

take.shape.PFV-FUT.3.SG-REFL console.IPFV-PST.M.SG Chekhov.NOM.SG and

ljud-i tesni-l-i-s' k nemu tolp-oj.

people-NOM.PL press.IPFV-PST-PL-REFL toward he.DAT crowd-INS.SG

'In two or three hundred years everything **will sort itself out**, Chekhov consoled, and people crowded around him.'

[Aleksandr Kušner. Počemu oni ne ljubili Čexova? // «Zvezda», 2002]

One of the attestations in our perfective dataset is of a biaspectual verb, where the interpretation is clearly Future:

(16) 8 fevral-ja v amerikansk-om Solt-Lejk-Siti

8 February-GEN.SG in American-N.LOC.SG Salt-Lake-City.LOC.SG

startu-et XIX zimn-jaja Olimpiad-a.

start.BIASP-FUT.3.SG XIX winter-F.NOM.SG Olympics.NOM.SG

'On February 8th the XIX Winter Olympics **kicks off** in Salt Lake City in America.'

[Pavel Černikov. Rossija v cifrax, 2002]

The imperfective future tense forms in our database are more consistent than the perfective ones in terms of referring to a point in the future relative to the speech time: 778 examples belong to that category. Example (17) illustrates a prediction for the events that are going to happen in the following year:

(17) V budušč-em god-u **bud-et** **prodolža-t'-sja**

in next-M.LOC.SG year-LOC.SG be-FUT.3.SG continue.IPFV-INF.REFL

rost vredonosn-yx programm dlja Linux,

growth.NOM.SG harmful-GEN.PL program.GEN.PL for Linux

i	vysok-a	verojatnost’	t-ogo,		čto
and	high-F	likelihood.NOM.SG	that-N.GEN.SG		that
v	2002-m	pojav-jat-sja	i	poluč-at	
in	2002-M.LOC.SG	appear.PFV-FUT.3.PL-REFL	and	receive.PFV-FUT.3.PL	
rasprostraneni-e	virus-y	dlja	Palm,	Pocket PC,	
spread-ACC.SG	virus.NOM.PL	for	Palm.GEN.SG	Pocket PC.GEN.SG	
sotov-yx	telefon-ov.				
cellular-GEN.PL	telephone-GEN.PL				

‘Next year, the growth of malware **will continue**, and it is highly likely that in 2002 viruses for Palm, Pocket PC and cell phones will appear and spread.’

[S. Potresov. God virusnogo bespredela, 2001]

There are three examples that refer to the future but have an additional meaning which comes not from the form but from the verb itself: *znat’* ‘know’[29]. The main function of these examples is to threaten the hearer. The threat is accompanied by a subordinate clause that can optionally be attached by conjunctions *čto* ‘what’ and *kak* ‘how’. The action by which the hearer is threatened has already happened at least once, hence the knowledge about it is present, not future. The threat is the repetition of this past action in the future:

(18)

Ingušsk-uju	milici-ju	zdes’	ne	ljub-jat:	
Ingush-F.ACC.SG	police.ACC.SG	here	not	love.IPFV-PRS.3.PL	
“Bud-ut	**zna-t’,**	čto	tak-oe	čečensk-ij	
be-FUT.3.PL	know.IPFV-INF	what	such-N.NOM.SG	Chechen-M.NOM.SG	
ženask-ij	batal’on!”				
female.NOM.SG	battalion.NOM.SG				

[29] The potential interchangeability of the perfective and imperfective future forms goes beyond the scope of this article; see (Janda et al. 2019).

‘The Ingush police is not popular here: “They **(will) know** better than to mess with a Chechen women’s battalion!”’

[Elena Samojlova. «Ljubogo menta pokolotit’ mogu!», 2002]

In (18) the women have already committed some kind of threatening action that inspired fear. And it is the knowledge that they are capable of this kind of action that belongs to the future.

In all the examples in the Future class, we observe a prediction of an event that is to take place in the future. These are events that are excluded from baseline Reality (Langacker 2019: 5) but instead are grounded in Projected Reality according to Langacker’s (2008: 306) model. While Future meaning is exactly what we would expect a future tense form to express, it is perhaps surprising that nearly half of the attestations in our perfective sample and almost one-fourth of the attestations in our imperfective sample cannot be clearly classified as Future. Recall that Forsyth (1970: 120) made a comparable but smaller estimate of one third.

3.2. Extended future

208 of the perfective future tense forms in our dataset describe an event that cannot be unambiguously assigned to Future due to lack of certainty about its completion or the fact that future tense forms can refer to events that are actually past or present. There are 133 attestations of imperfective future tense forms that belong to the Extended future group. We find the same subgroups in both the perfective and imperfective datasets.

There are two ways in which uncertainty is introduced, namely through the use of Implicative and Hypothetical expressions, covered in Sections 3.2.1 and 3.2.2. Posterior future, addressed in Section 3.2.3 is a reference to an event that takes place after a past event, but may be situated in any subsequent portion of the timeline (past, present, or future). Performatives in Section 3.2.4 describe events contiguous with the present moment. Section 3.2.5 describes Alternation found in the imperfective dataset.

3.2.1. Implicative

The largest group of Extended future examples falls into the category we label “Implicative” (Karttunen 1971). Implicative verbs contribute an additional layer of meaning, e.g. *smoč* ‘manage to’ adds a “degree of difficulty”. In these uses, future tense forms indicate not future events, but future possibilities of events depending on the presupposition that the situation described by a future tense verb form will help to

facilitate an action. Future possibilities of events operate with a greater degree of the modal component.

These uses can be interpreted according to Talmy's (2000, vol. 1: Chapter 7) model of force dynamics. The interaction with the force includes resistance, overcoming, blockage and removal of such blockage. Examples with Implicatives can act as motivating forces (Agonists) and the removal of barriers (Antagonists).

In the perfective dataset we find seventy-two examples of verbs with Implicative meanings followed by an infinitive. Two verbs account for over half the data, with twenty-one examples of the verb *smoč'* 'manage to', and twenty examples of the verb *prijtis'* 'have to'. Other verbs that occur more than once are *pozvolit'* 'allow' (six examples), *stat'* 'begin/become' (five examples), *udat'sja* 'succeed' (three examples), and *sumet'* 'succeed' (two examples).

(19) a. So vremen-em agentstv-o **smož-et**

with time-INS.SG agency-NOM.SG manage.PFV-FUT.3.SG

prevrati-t'-sja v krupn-uju prodjusersk-uju

turn.into.PFV-INF.REFL in large-F.ACC.SG production- F.ACC.SG

firm-u, raspolagajušč-uju ser'ëzn-ymi sredstv-ami.

company-ACC.SG endowed.with-F.ACC.SG serious-INS.PL fund-INS.PL

'Over time, the agency **will manage to** turn into a large production company with serious funds'.

[Artur Šumkov. Kinoèkonomika ne budet èkonomnoj, 2002]

b. Da, mal'čik-i, segodnja vam

yes boy-NOM.PL today you.DAT

prid-et-sja užina-t' s sosisk-ami…

have.to.PFV-FUT.3.SG-REFL dine.IPFV-INF with sausage-INS.PL

'Yes, boys, today you **will have to** dine on sausages…'

[Elena Pavlova. Vmeste my ètu propast' odoleem! 2004]

c. Sobyti-e bud-et togda, kogda

event.NOM.SG be-FUT.3.SG then when

my **zastav-im** mèr-a rasskaza-t',

we.NOM force.PFV-FUT.1.PL mayor-ACC.SG tell.PFV-INF

počemu "Xard-bank-u" otda-l-i nedvižimost'

why Hard-Bank-DAT.SG give.away.PFV-PST-PL property.ACC.SG

stoimost'-ju 70 mln. za 14 mln.

value-INS.SG 70 mln for 14 mln

'The event will happen when we **force** the mayor to tell why Hard-bank was given property worth 70 million for 14 million.'

[Sergej Nikolaev. Raz vzryv, dva zaderžanie, 2003]

The examples above contain both of the most frequent verbs from the sample: *smoč'* 'manage' and *prijtis'* 'have to', as well as a less frequent verb *zastavit'* 'force', which occurs only once in the perfective dataset. These are not predictions, but rather statements of the relative confidence of the speaker that the events are likely to occur in the near future. The implicative element adds dynamics of force to the main verb expressed by an infinitive. The examples, however, vary in structure and the force is applied in various manners. In (19a) the infinitive refers to the subject in nominative case, and here that subject (Agonist) will be empowered by a future situation in which there are no barriers (Antagonist). In (19b) the logical subject (in dative case) is also the same for the implicative verb and the infinitive that it governs; an unnamed Agonist applies force to the logical subject. Example (19c) illustrates a modal verb and infinitive with different subjects: *my* 'we' for *zastavim* '(we) will force' and *mèr* 'mayor' as the logical subject for *rasskazat'* 'tell'.

Two important Russian imperfective verbs with implicative meanings present paradigm gaps that prevent them from appearing in future forms: **budu moč'* '(I) will be able' and **budet prixodit'sja* 'will have to' do not exist. Given this fact, this group could not be expected to be numerous. There are only three examples. The verbs *starat'sja* and *umet'* (20) are followed by infinitives:

(20) Skoro v naš-ej stran-e každ-yj

soon in our-F.LOC.SG country-LOC.SG each-M.NOM.SG

bud-et **ume-t'** čita-t' i pisa-t'!

be-FUT.3.SG know.how.IPFV-INF read.IPFV-INF and write.IPFV-INF

'Soon everyone in our country **will be able to** read and write!'

[I. Grekova. Fazan, 1984]

Other elements, such as the conjunction *kak* 'how' can add implicative flavor:

(21) — Kak že ja **bud-u** **govori-t'**,

how EMPH I.NOM be-FUT.1.SG talk.IPFV-INF

kogda vy vs-e vrem-ja perebiva-ete…

when you.NOM all-N.ACC.SG time-ACC.SG interrupt.IPFV-PRS.2.PL

'How **am** I **supposed to talk** when you interrupt all the time…' (lit. 'how I will talk')

[J. O. Dombrovskij. Xranitel' drevnostej, 1964]

In the broader context, in (21) the speaker has already been talking for a while and one of the listeners has been trying to stick a word in all the time. The speaker is outraged and tries to convey the idea that he would like the hearer to stop interrupting in the nearest future. This idea is expressed in the form of a rhetorical question.

3.2.2. Hypothetical

Our perfective dataset contains fifty examples classed as Hypothetical, in forty-four of which a hypothesis is introduced by means of *esli* 'if', as in (22). According to Fauconnier (1985: Chapter 3), *if* is a space-builder that sets up a mental space that is relatively subjective in relation to Conceived Reality (cf. Langacker 2008: 528). *Esli* 'if' takes the example from the cone of Projected Reality to its border with Non-Reality.

(22) Esli **propust-iš'** xot' odn-o zanjati-e po

if miss.PFV-FUT.2.SG even one-N.ACC.SG lesson-ACC.SG along

masterstv-u, sčitaj, čto ty otčislen-a.

mastery-DAT.SG consider.IPFV.IMP.2.SG that you.NOM expelled-F

‘If you **miss** even one acting lesson, consider yourself expelled.’

[Sati Spivakova. Ne vsë, 2002]

Each of the remaining six examples classed as Hypothetical is unique in the perfective dataset, although there are some patterns, and all of them are introduced by space-builders. In four of them a hypothesis is introduced by a phrase such as *somnevajutsja, čto* ‘they doubt that’, *pri uslovii, čto* ‘under the condition that’ (23), and *možet* as an abbreviated form of *možet byt’* ‘perhaps’.

(23)

Po	ocenk-am	Raytheon,	firm-a	mog-l-a	
along	estimate.DAT.PL	Raytheon.GEN.SG	company-NOM.SG	can.IPFV-PST-F.SG	
by	izgotavliva-t’	èt-i	raket-y	pri	
COND	manufacture.IPFV-INF	this-ACC.PL	missile-ACC.PL	at	
stoimost-i	po	400	tys.	doll.	
cost-LOC.SG	along	400	thousand	dollars	
za	každ-uju,	**pri**	**uslovi-i,**		
for	each-ACC.SG	at	condition-LOC.SG		
čto	armi-ja	**zakaž-et**	1000	raket.	
that	army-NOM.SG	order.PFV-FUT.3.SG	1000	missile.GEN.PL	

‘According to Raytheon estimates, the company could manufacture these missiles at a cost of $ 400 thousand for each, **provided that** the army **orders** 1,000 missiles.’

[Vladimir Korovin. Novosti za rubežom // «Vozdušno-kosmičeskaja oborona», 2002]

The main clause of this example belongs to Non-Reality, signaled by the conditional marker *by*. The clause introduced by *pri uslovii, čto* uses a perfective future tense form to describe a possible facilitating event. In the remaining two instantiations it appears that *esli* has been elided: cf. example (24).

(24)

Ne	**progolosu-em —**	uxudš-it-sja	finansirovani-e	vs-ex
not	vote.PFV-FUT.1.PL	deteriorate.PFV-FUT.3.SG-REFL	financing-NOM.SG	all-GEN.PL

rasxod-ov na social'n-ye nužd-y, čto

expense-GEN.PL on social-ACC.PL need-ACC.PL that

neizbežno skaž-et-sja na avtoritet-e lev-yx

inevitably result.PFV-FUT.3.PL-REFL on authority-LOC.SG left-GEN.PL

sil i rezul'tat-ax očeredn-yx vybor-ov.

force.GEN.PL and result-LOC.PL next-GEN.PL election-GEN.PL

'(If) we do not **vote** – the financing of all expenses for social needs will deteriorate, which will inevitably have an impact on the authority of the forces on the left and the results of the next elections.'

[Vladimir Fedotkin. Vlast' i oppozicija, 2003]

Seven of the examples in the perfective dataset are classed as both Implicative and Hypothetical, such as (25) which contains the hypothetical space-builder *esli* 'if' in combination with the implicative verb *udat'sja* 'manage'.

(25) Xorošo, **esli** stran-e **uda-st-sja** pereži-t'

good if country-DAT.SG manage.PFV-FUT.3.SG survive.PFV-INF

nynešn-juju "stabilizaci-ju" i ona ne perejd-ët

current-F.ACC.SG stabilization-ACC.SG and she.NOM not go.over.PFV-FUT.3.SG

v poln-uju i okončatel'n-uju degradaci-ju.

in full-F.ACC.SG and final-F.ACC.SG degradation-ACC.SG

'It is good **if** the country **manages** to survive the current "stabilization" and it does not go into complete and final degradation.'

[Aleksandr Xramčixin. Kompleks polnocennosti, 2003]

The number of Hypothetical uses of the imperfective future is comparable to the perfective: thirty-nine (vs. fifty examples for perfectives). The space-builder *esli* 'if' remains the dominant means of expression in the group (thirty-one examples):

(26) **Esli** ja ne **bud-u** **protira-t'**

if	I.NOM	not	be.FUT-1.SG	wipe.IPFV-INF	
zvezd-y	každ-yj	večer, —		duma-l	on, —
star-ACC.PL	each-M.ACC.SG	evening.ACC.SG		think.IPFV-PST.M	he.NOM
oni	objazatel'no	potuskne-jut.			
they.NOM	surely	fade.PFV-FUT.3.PL			

'**If** I **don't** [lit. **will** not] **wipe** the stars every night,' he thought, 'they will surely fade.'

[Sergej Kozlov. Pravda, my budem vsegda? 1969-1981]

Other ways of expressing hypotheticality include various space-builders and the elision of *esli*, replacing it with a dash. The space-builders *budto* 'as if', *koli* 'if' and *eželi* 'if' are represented once each in our data. There are five examples where the space-builder 'if' is elided, one of which is presented here:

(27)

Bud-ut	**obiža-t'**,	**pristava-t' —**		
be.FUT-3.PL	offend.IPFV-INF	molest.IPFV-INF		
prixod-i	i	žaluj-sja		smelo.
come.IPFV-IMP.2.SG	and	complain.IPFV.IMP.2.SG-REFL		bravely

'(If) they [lit. **will**] **offend**, **molest** (you) – come and complain without fear.'

[È. G. Kazakevič. Zvezda, 1946]

Lastly, similar to the situation in our perfective dataset, two imperfective examples fall into two categories at the same time: Hypothetical and Posterior future, as in (28). The clarification of the Posterior future element comes in Section 3.2.3.

(28)

…dvoe	iz	nix	zajavi-l-i,		čto
two	from	they.GEN	announce.PFV-PST-PL		that
vernu-l-i-s'	by	v	"P. O. R. T. O. S.",		**esli**
return.PFV-PST-PL-REFL	COND	in	P.O.R.T.O.S.		if
organizaci-ja	snova	**bud-et**	**dejstvova-t'**.		
organization-NOM.SG	again	be.FUT-3.SG	operate.IPFV-INF		

'…two of them said they would return to P.O.R.T.O. S. if the organization [lit. **will function**] **were to operate** again.'

[Andrej Andreev. Buduščee prinadležit nam! 2003]

In (28) *esli* builds a mental space where the organization that the speaker is talking about is functioning. In this mental space the people ('they') are happy to return.

3.2.3. Posterior future

We begin this section with another hybrid example, this one combining Implicative use (signaled by *smoč'* 'manage') with the Posterior future[30], conditioned by the setting of what was said in the past:

(29) My ... vsegda govori-l-i, čto po finansov-ym
we.NOM always say.IPFV-PST-PL that along financial-DAT.PL
pričin-am ne **smož-em** sdela-t' èt-o
reason-DAT.PL not manage.PFV-FUT.1.PL do.PFV-INF this-N.ACC.SG
vovremja i bez pomošč-i zapadn-yx stran.
on.time and without help-GEN.SG western-GEN.PL country.GEN.PL

'We ... have always said that for financial reasons we **would not be able to** do this on time and without the help of Western countries.'

[Dmitrij Litovkin. Sroki podviga perenosjatsja, 2002]

There are forty-three examples of Posterior future in the perfective dataset, most of them showing the same pattern as the example above, where the main clause contains a past tense finite verb form, and the future form appears in a subordinate clause.

In addition, we present a perfective example that combines Hypothetical using *esli* 'if' with the Posterior future, occasioned by the fact that the document was sent in the past to be used in a hypothetical future scenario.

(30) ... kajzer-om Vil'gel'm-om II, v 1914 god-u

[30] One interpretation could be that Posterior Future is simply Future. However, since Posterior Future describes an event, where the reference time of Immediate reality is shifted to the past, we put it into the Extended future class.

emperor-INS.SG		Wilhelm.INS.SG	II	in	1914	year-LOC.SG		
prisla-vš-im		v	Peterburg		dv-e	not-y		
send.PFV-PST.ACT.PTCP-M.INS.SG		in	Petersburg-ACC.SG		two-F.ACC	note-ACC.PL		
ob	ob”javleni-i	vojn-y,		odn-u		na	slučaj,	
about	declaration-LOC.SG	war-GEN.SG		one-F.ACC.SG		on	case.ACC.SG	
esli	Rossi-ja	**otkaž-et-sja**		ostanovi-t’		mobilizaci-ju ...		
if	Russia.NOM.SG	refuse.PFV-FUT.3.SG-REFL		stop.PFV-INF		mobilization-ACC.SG		

‘…emperor Wilhelm II in 1914 had sent to Petersburg two declarations of war, one in case that Russia **refuses** to stop its mobilization…’

[Maksim Sokolov. 21.IX – 27.IX // «Izvestija», 2002.09.27]

Example (28) in the previous section additionally illustrates Posterior future because the verb *zajavili* ‘said/declared’ puts the whole situation (including the mental space) into the past, as in reported speech. In the imperfective dataset, the Posterior future is the largest subgroup of Extended future: it includes seventy-three examples.

Posterior future sometimes refers to events that were supposed to happen in the past at some point after the moment of speaking but might not ever have happened at all, as in (31). In other cases, the event has not happened yet or is going to continue in the future, cf. (32). However, most examples do not clearly indicate a difference between an event that happened after another event in the past or is still expected in the future, and one can only guess this from the context: cf. example (33).

(31)	On	žda-l,	čto	ja	**bud-u**	**provaliva-t’-sja,**
	he.NOM	wait.IPFV-PST.M	that	I.NOM	be.FUT-1.SG	fail.IPFV-INF.REFL
	i	xote-l,	čtoby	ja	provali-l-sja	
	and	want.IPFV-PST.M	that	I.NOM	fail.PFV-PST.M-REFL	
	kak	možno	medlenn-ej	i	interesn-ej.	
	as	possible	slow-COMPAR	and	interesting-COMPAR	

‘He expected [lit. that I **will fail**] me **to fail**, and wanted me to fail as slowly and interestingly as possible.’

[Fazil’ Iskander. Trinadcatyj podvig Gerakla, 1966]

In (32) the narrator describes a situation in which someone expected him to fail, and this happened when he was in school. From a broader context, it is clear that the time when he might have failed has already passed because we learn further on that he succeeded. So, the failure never took place and the time period where it could have happened is already over.

(32) Poda-l-i čt-o-to tak-oe bel-oe, ja
serve.PFV-PST-PL something-ACC like-N.ACC.SG white-N.ACC.SG I.NOM
prinja-l èt-o za mann-uju kaš-u, no
take.PFV-PST.M this-N.ACC.SG for semolina-F.ACC.SG cereal-ACC.SG but
kogda poprobova-l, to ponja-l, čto ja
when tried.PFV-PST.M then realize.PFV-PST.M that I.NOM
bud-u es-t’ èt-o vs-ju svo-ju žizn’,
be.FUT-1.SG eat.IPFV-INF this-N.ACC.SG all-F.ACC.SG own-F.SCC.SG life.ACC.SG
po tri raz-a každ-yj den’.
along three.ACC time-GEN.SG each-M.ACC.SG day.ACC.SG

‘They served something white, I took it for semolina, but when I tried it, I realized that I [lit. **will**] **would eat** it all my life, three times every day.’

[Natal’ja Skljarova. Esli by u medvedja bylo ruž’e, 2002]

At the moment of producing sentence (32) the speaker was obviously alive and planning to continue to eat the food that she tasted for all the foreseeable future. So, the eating event began in the past and continues indefinitely.

(33) Èt-o označa-l-o, čto otnyne specialist-ov
This-N.NOM.SG mean.IPFV-PST-N that from.now specialist-ACC.PL
po èt-oj disciplin-e **bud-ut** **gotovi-t’**

along	this-F.DAT.SG	discipline-DAT.SG	be.FUT-3.PL	prepare.IPFV-INF
v	gosudarstvenn-yx	vysš-ix	učebn-yx	zavedeni-jax.
in	state-LOC.PL	high-LOC.PL	educational-LOC.PL	institution-LOC.PL

'This meant that from that moment on specialists in this discipline [lit. **will**] **would be trained** in state higher educational institutions.'

[Marija Kozlova. «Advokat», 2004.12.01]

In (33) we do not know whether the training process has changed over time or not, so the duration of the event is unknown.

3.2.4. Performative

Our perfective dataset contains thirty-seven examples of Performative Extended future. Performatives are defined as illocutionary acts that can be executed by uttering a sentence (see Searle 1989: 536). In the Russian tradition, non-past Performatives are viewed as referring to a speech act that coincides with uttering the verb that names this act (cf. Vinogradov 1947; Zaliznjak 2015).

We also include Near-performative uses in this class. Near-performatives are not Performatives stricto sensu[31]. These are approximately simultaneous actions: they happen either just before (34) or just after (35) another action. The expression of Performatives and Near-performatives is mostly restricted to verbs that describe speaking (*skažu* 'I will say', *nazovu* 'I will name', *opišem* 'we will describe') and other actions connected to verbal argumentation such as directing the focus of the hearer (*zametim* 'we will note', *podčerknem* 'we will emphasize', *rassmotrim* 'we will examine'). We also find verbs used metaphorically to refer to discourse actions such as *dobavim* 'we will add' (speaking points), *ostanovimsja* 'we will stop' (meaning that we will spend time discussing certain points), *privedu* 'I will bring' (with direct objects *primer* 'example' and *vyderžku* 'excerpt' meaning that the speaker is inserting items into a discussion). We have two examples of Performative uses that involve other types of (nonverbal) actions: *pokažu* 'I will show' in a frame where the speaker

[31] For the discussion of the verbs functioning as performatives in a public discourse see Dickey (2000, Chapter 6).

is offering documents to an official, and *pogljažu* ‘I will take a look’ in a frame where the speaker performs actions and describes them while talking to children.

(34)	My	že	**ostanov-im-sja**	na	bolee
	we.NOM	EMPH	stop.PFV-FUT.1.PL-REFL	on	more
	uporjadočenn-yx		process-ax.		
	predictable-LOC.PL		process-LOC.PL		

‘We **will focus** on the more predictable processes.’

[Sergej Dorenko. Levye sily – perezagruzka, 2003]

In the text preceding (34), the author of an article about political processes is describing a situation and listing possible scenarios for these processes. The speaker has already decided to write about the more predictable processes well before this sentence is actually written and read. For this reason, the action of focusing attention signaled by *ostanovimsja* ‘we will focus’ can be viewed as already completed.

(35)	O	plan-ax	**skaž-u,**	čto	xoč-u
	about	plan-LOC.PL	say.PFV-FUT.1.SG	that	want.IPFV-PRS.1.SG
	privez-ti	domoj	medal-i—	na	pamjat’.
	bring.PFV-INF	home	medal-ACC.PL	on	memory.ACC.SG

‘As for the plans, I **will say** that I want to bring the medals home, as a souvenir.’

[Oleg Lisogor: «Čto tolku mečtat’? Rabotat’ nado!», 2002]

In (35) the action of saying referred by the verb in the main clause is simultaneous with the speaker’s utterance of the content described in the subordinate clause.

Twelve examples in the imperfective dataset are marked as Performatives. Here the imperfective uses demonstrate the same pattern as the perfective ones: most of them are first person singular forms. Verbal argumentation is the primary meaning of the verbs in the Performative category, where we observe verbs like *rezjumirovat’* ‘summarize’, *obsuždat’sja* ‘discuss’, *vrat’* ‘lie’, *govorit’* ‘speak’, *pisat’* ‘write’.

(36)	Ja	ne	**bud-u**	**rezjumirova-t’**	rezul’tat-ov

I.NOM	not	be.FUT-1.SG	summarize.IPFV-INF		result-GEN.PL
izlož-enn-ogo			issledovani-ja	i	perexož-u
present.PFV-PST.PASS.PTCP-N.GEN.SG			research-GEN.SG	and	go.over.IPFV-PRS.1.SG
prjamo	k	vyvod-am.			
directly	toward	conclusion-DAT.PL			

'I **will** not **summarize** the results of the study and (I) proceed directly to the conclusions.'

[A.N. Leont'ev. Biologičeskoe i social'noe v psixike čeloveka, 1981]

In Example (36)[32], the speaker decides to skip the talk about one part of his research and this decision immediately (performatively) results in avoiding it and proceeding to the next section.

3.2.5. Alternation

In the perfective dataset all Alternations belong to Gnomic (see Section 3.4.4). In the imperfective, however, there is one example where the first part of the Alternation introduced by *čem* happens in the present while the second part beginning with *tem* is in contrast with the first and is supposed to be fulfilled in the future:

(37)						
	I	**č-em**	ničtožn-ee	mo-ja	rol'	v
	and	what-INS	insignificant-COMPAR	my-F.NOM.SG	role.NOM.SG	in
	nastojašč-ej	žizn-i,	**č-em**	bescvetn-ee	sam-yj	
	real-F.LOC.SG	life-LOC.SG	what-INS	colorless-COMPAR	very-M.NOM.SG	
	fon	mo-ego	suščestvovani-ja,	**t-em**		
	background.NOM.SG	my-N.GEN.SG	existence-GEN.SG	that-INS		
	jarč-e	**bud-et**	**sija-t'**	mo-e		
	bright-COMPAR	be.FUT-3.SG	shine.IPFV-INF	my-N.NOM.SG		

[32] We realize that the interaction of negation with imperfective aspect could play a potential role in the classification. This interaction goes beyond the scope of this article.

sentimental'n-oe,	mo-e	ščedr-oe,	mo-e
sentimental-N.NOM.SG	my-N.NOM.SG	generous-N.NOM.SG	my-N.NOM.SG
velikodušn-oe	i	prekrasn-oe	solnc-e.
benevolent-N.NOM.SG	and	beautiful-N.NOM.SG	sun-NOM.SG

'And **the** more insignificant my role in real life, **the** more colorless the very background of my existence are, **the** brighter my sentimental, my generous, my benevolent and beautiful sun **will shine**.'

[I. F. Annenskij. Vtoraja kniga otraženij, 1909]

3.3. Directive

Whereas Performatives announce an action taken by the speaker, Directives deliver instructions to immediately perform an action that should be nearly simultaneous with the utterance. Our category of Directives is represented by twelve perfective examples, all of which function similarly to imperatives (for more details on the difference between these two forms see Stojnova 2016a), though they are realized morphologically in various ways, such as by means of finite forms as in (38)[33].

(38)

Voz'm-eš'	mikrofon.	**Zakat-iš'**	glaz-a
take.PFV-FUT.2.SG	microphone.ACC.SG	roll.PFV-FUT.2.SG	eye-ACC.PL

kartinno.	Èt-o	ty	i
picturesquely	this-N.ACC.SG	you.NOM	and
bez	mikrofon-a	ume-eš'.	
without	microphone-GEN.SG	know.how.IPFV-PRS.2.SG	

'**Take** the microphone. **Roll** your eyes picturesquely. You can even do it without a microphone.'

[Zapis' LiveJournal, 2004]

[33] In order to save space and for readability reasons, we do not present here the broader context that is available in the RNC. In all cases where there is an ambiguity, we have performed a more detailed analysis of the context.

Sometimes finite forms are combined with hortative markers like *pust'* 'let, may' as in (39).

(39) **Pust'** t-a že učast'
may this-F.NOM.SG EMPH fate.NOM.SG
postign-et det-ej prezident-a Buš-a.
befall.PFV-FUT.3.SG child-GEN.PL president-GEN.SG Bush-GEN.SG
'**May** the same fate **befall** the children of President Bush.'
[Aleksandr Proxanov. Prokuratura – kastet s programmnym upravleniem, 2003]

There are five examples of imperfective Directives and they are a diverse group. Two Directives are used in combination with the hortative marker *davajte* 'let us' (40), which is not attested in the perfective dataset (for the use of *dajte*/*davajte* see Janda, Lyashevskaya 2011: 738, 741).

(40) …a my **davaj-te** **bud-em** **volnova-t'-sja,**
and we.NOM give.IPFV.IMP-2.PL be.FUT-1.PL worry.IPFV-INF-REFL
čtoby glupost-ej ne natvori-t'
that stupidity-GEN.PL no create.PFV-INF
na svo-em učastk-e.
on own-M.LOC.SG site-LOC.SG
'…and **let us worry** about not doing stupid things on our site.'
[Vasilij Grossman. Žizn' i sud'ba, 1960]

The broader context of (40) tells the hearer about the division of the responsibilities: one group of people is doing one specific thing and the other is instructed to worry about their own site.

3.4 Gnomic

The single biggest deviation from using both perfective and imperfective future tense forms to refer to future time are the Gnomic uses. While in Extended future the future remains the main meaning of the verb forms (that is, this class does not deviate

significantly from the main meaning in terms of quality), and Directives are few in number, the class of Gnomic uses stands out in terms of quantity *and* quality. In total, there are 210 examples with perfective verb forms and eighty-four with imperfective verb forms that belong to Gnomic. In these examples, the future tense forms refer not to a future event, but to situations that are not grounded in time.

These examples are distinct from the Future type described above (3.1 and 3.2) in that they do not describe events that can be located in Projected Reality. The Gnomic events do not derive their definiteness from a specific temporal location (for more on the comparison of definiteness and tense in terms of grounding see Langacker 2008: 78). Instead, these events may appear at one or multiple locations in reality (including Potential Reality). They derive their definiteness from Conceived Reality available to the speaker (Langacker 2008: 301). In other words, they are anchored to a generalized situation that is accessible to the speaker. We use the term Gnomic to highlight the lack of a specific temporal location. In addition to third person singular, second person references also support Gnomic readings (as we will see further on in examples 43 and 45). Example (41) is extracted from a passage on various possible useful applications of aluminum foil in everyday life:

(41) Žaren-aja kuric-a, zavernu-t-aja

fried-F.NOM.SG chicken-NOM.SG wrap.PFV-PST.PASS.PTCP-F-NOM.SG

v fol'g-u i ulož-enn-aja

in foil-ACC.SG and place.PFV-PST.PASS.PTCP-F.NOM.SG

v plotno zakryvaj-ušč-uju-sja

in fully cover.IPFV-PRS.ACT.PTCP-F.ACC.SG

kastrjul'k-u, **ostan-et-sja** dolgo tepl-oj.

saucepan-ACC.SG remain.PFV-FUT.3.SG long warm-F.INS.SG

'Fried chicken wrapped in foil and placed in a tightly closed saucepan **will remain** warm for a long time.'

[M. Volodina. O fol'ge, 2002]

In (41), the situation that grounds the event is any instance of a fried chicken wrapped in foil. Given this situation, the speaker can observe that it instantiates a Gnomic

potential, namely that the chicken will stay warm. Metonymy plays a role here as well: a single event is picked out to represent a whole class of events.

(42) is an example of an imperfective Gnomic use:

(42)	Sobak-a	ne	**bud-et**	**es-t'**	
	dog.NOM.SG	not	be.FUT-3.SG	eat.IPFV-INF	
	t-ogo,	č-em	ja	pita-ju-s'.	
	that-N.GEN.SG	what-INS	I.NOM	feed.on.IPFV-PRS.1.SG-REFL	
	'A dog **will** not **eat** what I eat.'				
	[Sergej Dovlatov. Inaja žizn', 1984]				

In (42) there is neither a specific dog, nor do we expect that there will be any dog in the future. It is just common knowledge that normally dogs eat something better than what the speaker has.

Gnomic uses present various perspectives that connect to a variety of characteristics. These include the stability of salient features across time, modality and hypotheticality. The pattern of submeanings among Gnomic uses with imperfective verbs partly follows but also partly deviates from the pattern observed for perfective verbs. We take up each submeaning in turn in the following subsections.

3.4.1. Stable scenarios

Some uses are Gnomic because they are grounded in encyclopaedic knowledge about how Conceived Reality functions. This can include generalizations[34] based on the experience of the speaker and the Gnomic use can serve to deliver advice to a hearer.

(43)	— Dlja	nas,	pčel,	v	skoš-enn-oj	trav-e
	for	us.GEN	bee.GEN.PL	in	cut.PFV-PST.PASS.PTCP-F.LOC.SG	grass-LOC.SG

[34] In Russian grammars, second and sometimes first-person verb forms used without a pronominal subject accompanying them usually fall under the definition of general personal sentences (see Isačenko (1965/2003: 415) and Russian Grammar 1980 §2251).

prok-u	nikak-ogo.	Nektar	iz	nee
use-GEN.SG	none-M.GEN.SG	nectar.ACC.SG	from	she.GEN
ne	**voz'm-eš'**,—	prodolža-l-a	star-aja	Pčel-a.
not	take.PFV-FUT.2.SG	continue.IPFV-PST-F	old-F-NOM.SG	Bee-NOM.SG

'— For us bees, there's no use in cut grass. You **can't take** nectar from it — continued the old Bee.'

[Viktor Kologriv. Medovyj lug // «Murzilka», 2002]

In (43), the bee has observed on previous occasions that cut grass is useless and conveys this as a stable fact to her interlocutor, one that is potentially infinitely reproducible. Possibility or lack thereof gives a modal flavor to such statements, and the use of second person singular without a subject conveys a generic statement based on the speaker's knowledge of the world. There are fifty-two examples of perfective Stable scenarios including three examples where Stable scenario is combined with Habitual chain or Hypothetical.

Thirteen imperfective examples are marked as Stable scenarios. The speaker's experience or knowledge of the world and human nature provides the ground to generalize and give advice or instruction.

(44)

Ne	suščestvu-et		krizis-a		vozrast-a —	suščestvu-et	
not	exist.IPFV-PRS.3.SG		crisis-GEN.SG		age-GEN.SG	exist.IPFV-PRS.3.SG	
strax,	čto	čt-o-to			ne-zaplanirov-ann-oe		
fear.NOM.SG	that	something-N.NOM.SG			not-plan.PFV-PST.PASS.PTCP-N.NOM.SG		
sluč-it-sja			v	tvo-ej	žizn-i,	i	ty
happen.PFV-FUT.3.SG-REFL			in	your-F.LOC.SG	life-LOC.SG	and	you.NOM
ne	**bud-eš'**		**zna-t',**		čt-o	dela-t'.	
not	be.FUT-3.SG		know.IPFV-INF		what-ACC	do.IPFV-INF	

'An age crisis does not exist: there is a fear that something unplanned will happen in your life and you **will** not **know** what to do.'

[Marija Vardenga. Galina Tjunina. Fragmenty belogo stixa, 2002]

Taking into account the broader context available in the RNC, the example (44) is a philosophical passage not limited to specific unplanned sudden future events; it relates a common and constant situation that many people find themselves in.

3.4.2. Implicative

Five perfective Gnomic uses directly involve implicative verbs as auxiliaries to infinitive main verbs, as in (45).

(45) Za prevyšeni-e limit-a vam

for exceeding-ACC.SG limit-GEN.SG you.DAT

prid-et-sja doplačiva-t’ iz rasčet-a

have.to.PFV-FUT.3.SG-REFL pay.up.IPFV-INF from accounting-GEN.SG

15-30 cent-ov za každ-yj

15-30 cent-GEN.PL for each-M.ACC.SG

kilometr probeg-a sverx norm-y.

kilometer.ACC.SG distance-GEN.SG above limit-GEN.SG

‘For exceeding the limit you **have to** pay extra at the rate of 15-30 cents for each excess kilometer.’

[Denis Litošik. Avtomobil’ naprokat // «Avtopilot», 2002.05.15]

Here a rental car agent is stating a fixed rule about what happens when a client drives over a set number of kilometers. Note that if perceived without the broader context, (45) could refer to a single situation in the future: there would be a specific hearer (*vy* ‘you’) getting specific instructions.

No implicative uses are attested in the imperfective dataset.

3.4.3. Hypothetical

In Hypothetical Gnomic uses, we see an ‘if…then…’ semantic structure that can be occasioned only by future tense forms, as in this example, or enhanced by words like *esli* ‘if’. Here there is no grounding in a specific time, but a prediction based on overall experience:

(46) Tiraž malen'k-ij, **vygon-jat** ix iz
edition.NOM.SG small-M.NOM.SG chase.away.PFV-FUT.3.PL they.ACC from
odn-oj, tipografi-i, oni napečata-jut-sja v drug-oj.
one-F.GEN.SG printery-GEN.SG they.NOM print.PFV-FUT.3.PL-REFL in other-F.LOC.SG

'The edition is small, (if) they **get kicked out** of one publishing house, they print them in another.'

[D. Volkov, V. Sungorkin. Kuxnja upravljaemoj demokratii, 2003]

In (46), no one is actually going to kick the people out; however, the speaker hypothesizes that if that is to happen, it won't be a problem because these people are able to find a different place to publish their leaflets or flyers. In total, there are nine Hypothetical examples attested in the perfective dataset.

Two imperfective examples are both Stable scenarios and Hypothetical, as (47) illustrates with the space-builder *esli* 'if' followed by a covert piece of advice.

(47) Xorošo, **esli** v vaš-em ugolk-e
good if in your-M.LOC.SG corner-LOC.SG
otdyx-a **bud-et** **prisutstvova-t'** tak-oj
rest-GEN.SG be.FUT-3.SG be.present.IPFV-INF such-M.NOM.SG
tradicionn-yj èlement, kak vod-a.
traditional-M.NOM.SG element.NOM.SG like water-NOM.SG

'(It is/will be) good **if** there [lit. **will be**] **is** such traditional element like water in your garden retreat.'

[Valerija Iršenkova. Svoj ugolok ja ubrala cvetami, 2003]

According to the broader context, this speaker is addressing someone who is designing a garden and wants to remind them to include a creek or fountain in their plans. Instead of giving advice directly and overtly, the speaker uses an imperfective future tense form as a subtle way to hint at a suggested option. Here the Gnomic meaning performs the function of a politeness strategy. There are seven more imperfective Hypothetical Gnomic uses, two of which are introduced by the space-builder *pri uslovii, čto* 'provided that'.

The linking of one event to another is a common although not specific characteristic of Gnomic uses apparent also in the next two subsections.

3.4.4. Alternation

This subtype in the Gnomic class is represented by three examples in the perfective dataset. No Gnomic Alternations have been found among the imperfective examples.

Pairs of events that are linked by experience can be formally linked in grammatical constructions such as '*to*…, *to*…' and '*čem*…, *tem*…' as in (48).

(48) A **č-em** dol'-še **soxran-it-sja** prostranstv-o

and what-INS long-COMPAR preserve.PFV-FUT.3.SG-REFL space-NOM.SG

igr-y, **t-em** lučše dlja razviti-ja rebenk-a.

play-GEN.SG that-INS better for development-GEN.SG child-GEN.SG

'**The** longer an opportunity for play **is preserved**, **the** better it is for the child's development.'

[Anna Fen'ko. Nevroz po povedeniju// «Kommersant-Vlast'», 2002]

3.4.5. Habitual chain

The Habitual[35] chain type has been described by Dickey (2000: 55-56) and Bondarko (1971: 197-208). Habitual chains describe the speaker's knowledge about two or more events that typically take place in a given sequence. Habitual chains are also by definition stable situations.

(49) Byva-et, **spil-jat** v tajg-e ogromn-yj

happen.IPFV-PRS.3.SG cut.down.PFV-FUT.3.PL in taiga-LOC.SG huge-M.ACC.SG

kedr, **privez-ut** v poselok, a iz

cedar.ACC.SG bring.PFV-FUT.3.PL in village.ACC.SG and from

pust-ogo vnutri stvol-a medved' vyleza-et.

[35] Bybee et al. (1994:141) describes Gnomic uses as timeless situations that hold forever. Some may argue that Habitual chains are not Gnomic due to their repetitive meaning. However, since the repetitive potential of Habitual chains is continuous, we argue that they can be recognized as a variant of Gnomic use.

empty-M.GEN.SG	inside	trunk-GEN.SG	bear.NOM.SG	crawl.out.IPFV-PRS.3.SG

‘It happens that they **cut down** a huge cedar in the taiga and **bring** it to the village, and a bear crawls out of the hollow trunk.’

[Gennadij Snegirev. Medved’ // «Murzilka», 2003]

In the example (49) the two events are cutting down and bringing the tree to the village, which form a fixed sequence regardless of when they take place. *Byvaet* or *byvalo* ‘it happens’ indicates the potentially infinite number of times that the event can take place. The Habitual chain use of the Gnomic often appears with sequences that took place repeatedly in the past. There are twenty-seven perfective examples indicating Habitual chains, including one Habitual chain combined with Alternation and eight Habitual chains that are at the same time Stable scenarios.

Gnomic Habitual chains can sometimes contain imperfective future tense forms as well. Our dataset contains four examples of imperfective Habitual chains.

(50)

Potom	Viktor	Pavlovič	dolgo	**bud-et**
then	Viktor.NOM.SG	Pavlovich.NOM.SG	long	be.FUT-3.SG
side-t’	nepodvižno,	**potom**	načn-et	kiva-t’
sit.IPFV-INF	motionless	then	begin.PFV-FUT.3.SG	nod.IPFV-INF
golov-oj,	kak-to	pokorno,	po-starčeski	tosklivo.
head-INS.SG	somehow	docilely	old-fashioned.way	sadly

‘**Then** Viktor Pavlovich **will sit** motionless for a long time, **then** he will begin to nod his head, somehow docilely, with old-fashioned melancholy.’

[Vasilij Grossman. Žizn’ i sud’ba, 1960]

The person described in (50) has specific behavioral patterns, which are sequenced by means of *potom* ‘then’.

3.4.6. Salient event

A Salient event highlights a sudden or exceptional event that stands out as a figure in contrast to the background of what is usual (cf. single events in Dickey 2000: 57; Bondarko 1971: 213). There are seven examples of perfective Salient events and two

examples that combine the meanings of Salient event and Hypothetical. In our database, all Salient events belong to Gnomic.

(51) Obyčno tak-ie «perebo-i ritm-a»
usually such-NOM.PL rupture-NOM.PL rhythm-GEN.SG
sluča-jut-sja, kogda, kak grom sredi
happen.IPFV-PRS.3.PL when like thunder.NOM.SG in.middle.of
jasn-ogo neba, **mel'kn-et**
clear-N.GEN.SG sky-GEN.SG flash.PFV-FUT.3.SG
zagadočn-yj son, kotor-yj
mysterious-M.NOM.SG dream.NOM.SG which-M.ACC.SG
tak i xoč-et-sja nazva-t' veščim.
so and want.IPFV-PRS.3.SG-REFL name.PFV-INF prophetic-M.INS.SG

'Usually ruptures in one's rhythm take place when, like a stroke of lightning, one **gets a flash** of a mysterious dream that one feels must be prophetic.'

[Aleksandr Volkov. Miry Stivena Xoukinga // «Znanie – sila», 2003]

Here in (51) the background is the regular routines of a person's life that are suddenly interrupted by a dream. This example also illustrates the description of a Stable scenario – something that "usually" happens.

There is only one imperfective example that resembles a Salient event:

(52) Derevjann-yj ili parketn-yj pol nužno
wooden-M.ACC.SG or parquet-M.ACC.SG floor.ACC.SG must
objazatel'no ukrepi-t' (inače on **bud-et**
necessarily strengthen.PFV-INF otherwise he.NOM be.FUT-3.SG
skripe-t') i pod linoleum-om),
squeak.IPFV-INF and under linoleum-INS.SG

zadela-t’ treščin-y.

repair.PFV-INF crack-ACC.PL

‘A wooden or parquet floor must be strengthened (otherwise it **will squeak** under the linoleum), (one also) needs to repair the cracks.’

[Elena Volkova. Tot samyj linoleum, 2002]

Example (52) offers advice on how to do a good job and is interrupted by a parenthetical clause that describes an undesirable alternative.

3.4.7. Gnomic uses restricted in the future

Many Gnomic examples could have a potential endpoint. However, there are several examples in the imperfective dataset where the speaker chose to specify a closing boundary for the event. These examples describe generalized events lacking temporal grounding that are bounded at some point in the future.

(53) Teper’ èto tol’ko vremenn-aja razluk-a

now this only temporary-F.NOM.SG separation-NOM.SG

duš-i i tela, vremja, kogda

soul-GEN.SG and body-GEN.SG time.NOM.SG when

tel-o **bud-et** **otdyxa-t’** i **rassypl-et-sja**

body-NOM.SG be.FUT-3.SG rest.IPFV-INF and crumble.PFV-FUT.3.SG-REFL

v prax, togda kak duš-a

in dust.ACC.SG when as soul-NOM.SG

bud-et oživa-t’ vse bolee i

be.FUT-3.SG come.alive.IPFV-INF all more and

bolee, kak vse razgora-jušč-ee-sja plamja,

more as all burn.IPFV-PRS.ACT.PTCP-N.NOM.SG-REFL flame.NOM.SG

do **dn-ja**, kogda bud-et vosstanovl-en-a

to day-GEN.SG when be.FUT-3.SG restore.PFV-PST.PASS.PTCP-F

cel'nost',	kogda	my	voskresn-em	i
integrity.NOM.SG	when	we.NOM	be.resurrected.PFV-FUT.1.PL	and
zaživ-em	žizn-'ju	Boži-ej	voveki.	
begin.to.live.PFV-FUT.1.PL	life-INS.SG	God's-F.INS.SG	forever	

'Now this is only a temporary separation of the soul and body, the time when the body **will be at rest** and **crumbling** to dust, while the soul will come to life more and more, like all the flaming fire, **until the day** when integrity is restored, when we are resurrected and live the life of God forever.'

[mitropolit Antonij (Blum). Strastnaja sedmica, 1980]

In (53) the author describes processes that are happening and will continue to happen in the future for a while until a specific event is supposed to take place ('until the day when…').

3.5. Ambiguous biaspectual examples

Biaspectual verbs have non-past forms that can be interpreted either as a present tense form of an imperfective verb or as a future tense form of a perfective verb. In our database we find examples of non-past forms of biaspectual verbs that can be interpreted either as perfective future tense forms with a Gnomic meaning, or as imperfective present tense forms. While no amount of context can definitively distinguish between these two options, it is usually the case that one of the interpretations is more likely.

(54)

V	èt-ot	[podrostkov-yj]	period,		
in	this-M.ACC.SG	teenage-M.ACC.SG	period.ACC.SG		
sčita-et	Vygotskij,	proisxod-it			
consider.IPFV-PRS.3.SG	Vygotsky.NOM.SG	happen.IPFV-PRS.3.SG			
glubok-oe	preobrazovani-e	voobraženi-ja:	iz		
profound-N.NOM.SG	transformation-NOM.SG	imagination-GEN.SG	from		
sub″ektivn-ogo	ono	**preobrazu-et-sja**	v	ob″ektivn-oe.	
subjective-N.GEN.SG	it.NOM	transform.IPFV-PRS.3.SG-REFL	in	objective-N.ACC.SG	

transform.PFV-FUT.3.SG-REFL

‘In this [teenage] period, Vygotsky believes, a profound transformation of the imagination takes place: from the subjective, it **is transformed** into objective.’

[E. P. Krupnik. Voprosy psixologii, 2003]

In (54) the famous psychologist is referring to what usually happens in the life of an adolescent using the biaspectual verb *preobrazuetsja* ‘transforms’. Either Vygotskij is making a Gnomic statement about what always happens, or he is making a statement about a transformation that is currently happening from the internal perspective of the adolescent period.

4. Future tense meanings and modality

This section presents a digression on how future can interact with modality. To show which future tense meanings are combined with different types of modality, we examined the examples presented in the works of Petruxina and Li (2015), Wiemer et al. (2020), Klimonow (2011), and Radbil (2011) and compared them with our classification of future tense meanings Here we present a brief overview of the future – modality interaction. We found examples with Future, Extended future, and various Gnomic meanings; no Directives are attested in the abovementioned articles.

Most of the examples express unextended Future. The dominant modality is the epistemic modality. Klimonow (2011) provides a rather simple example (55) of a prototypical Future which he labels as epistemic modality:

(55)	Boris	**pereplyv-et**	Volg-u.
	Boris.NOM.SG	swim.across.PFV-FUT.3.SG	Volga-ACC.SG

‘Boris **will swim across** the Volga.’

In addition, Klimonow (2011) shows the interaction of Future and potential modality. Petruxina and Li (2015) introduce the notions of volitive modality. Radbil puts the Future meanings on a scale between “future as a fact” and “future as (undifferentiated) modality”. The distribution of modal meanings across utterances with Future meaning is presented in Table 2.

	Petruxina & Li	Wiemer et al.	Klimonow	Radbil
Epistemic	✓	✓	✓	(✓)
Volitive	✓			
Potential			✓	
Undifferentiated modality				✓

Table 2. Distribution of modality in combination with the Future meaning.

Extended Future is represented by such extensions as Hypotheticals, Posterior futures and Performatives. Hypothetical examples are presented in the works of Petruxina and Li (2015) and Radbil (2011). Petruxina and Li describe if-statements as potential, possible action (56).

(56) Esli Bonapart **bud-et** **id-ti** tak, to
if Bonaparte.NOM.SG be.FUT-3.SG go.IPFV-INF so then
čerez tri, ot sil-y četyre
after three.ACC from force-GEN.SG four.ACC
nedel-i dostign-et porog-a mo-ego dom-a.
week-ACC.PL reach.PFV-FUT.3.SG doorstep-GEN.SG my-M.GEN.SG house-GEN.SG

'If Bonaparte **continues to move forward** (lit. '**will go**') like this, then in three, maybe four weeks he will reach my doorstep.'

Posterior future, mentioned only by Petruxina and Li (2015), loses modality because technically the action in the utterance has already happened. Performatives can have a volitive modal component (Petruxina & Li 2015), or a combination of volitive and basic modality (Klimonow 2011). Radbil (2011) finds modality in performative utterances to be bleached: future as modality transforms into future as a fact. A short summary of the relationship between various Extended futures and modality is shown in Table 3.

	Petruxina & Li	Klimonow	Radbil
Volitive	Performative	Performative	
Potential	Hypothetical		
Basic		Performative	
No modality	Posterior		Hypothetical;
			Performative

Table 3. Distribution of modality across the Extended future.

In addition to Future and its extensions, Gnomic uses are widely presented in the discussed literature. Example (57) is a statement about the human ability to overcome difficulties; it combines Gnomic use with potential modality (Klimonow 2011).

(57) Čelovek **vyderž-it** mnog-oe.

person.NOM.SG withstand.PFV-FUT.3.SG much-N.ACC.SG

'A person can withstand (lit. '**will withstand'**) a lot.'

As shown in Table 4, Gnomic uses of future tense forms without any additional properties are attested to have the following types of modality: potential (Klimonow 2011), deontic (Petruxina & Li 2015), epistemic (Wiemer et al. 2020) and undifferentiated modality (Radbil 2011).

	Petruxina & Li	Wiemer et al.	Klimonow	Radbil
Epistemic		✓		
Potential			✓	
Deontic	✓			
Undifferentiated modality				✓

Table 4. Distribution of modality in combination with Gnomic meaning.

We attested three additional specifications of Gnomic: Habitual chains, Salient events and Stable scenarios. The distribution of these Gnomic subtypes is shown in Table 5.

	Petruxina & Li	Wiemer et al.	Klimonow
Potential		Habitual chain	Salient event; stable scenario
Circumstantial	Stable scenario	Habitual chain	

Table 5. Distribution of modality across the specialized Gnomic uses.

Below is an example of Habitual chain with attributed circumstantial modality (Wiemer et al. 2020).

(57) A byva-et čto ot pečal-i posle
and happen.IPFV-PRS.3.SG that from grief.GEN.SG after
poter-i xozjain-a i sobak-a **umr-et** srazu
loss-GEN.SG owner-GEN.SG and dog-NOM.SG die.PFV-FUT.3.SG immediately

'Sometimes (lit. 'it happens so that') the dog **will immediately** die from grief after the loss of the owner.'

Various types of modalities and future tense meanings can combine quite freely: e.g., utterances with potential modality can have either Gnomic or (Extended) Future meaning. We do not attest a pattern that could explain the motivation behind various future tense meanings via modality. The uncertain nature of both future tense and modality creates the space for variation. However, this variation and overlap does not conflict with our classification presented in Section 3.

5. Radial category of meanings for future tense forms

We found attestations of the elements of the same semantic classification in both perfective and imperfective future tense forms and we used similar semantic classifications for both, in accord with Stojnova's (2016b: 248) thesis that the perfective and imperfective future tense forms tend to express the same meanings. As we have shown in detail in Section 3, the category of future tense is polysemous. According to Lewandowska-Tomaszczyk (2007: 142), polysemous entities have a prototypical meaning and a radial category. A radial category is understood here as a

network of related meanings structured around the prototypical meaning (Lakoff 1987: 91). The representation of our classification as a radial category provides a coherent account of the meanings of the future tense across aspect. To show the differences and the similarities of the meanings exhibited by the two future tenses, we present two radial categories in Figures 5a and 5b.

Figures 5a and 5b visualize the radial categories of meanings expressed by future tense forms in Russian.

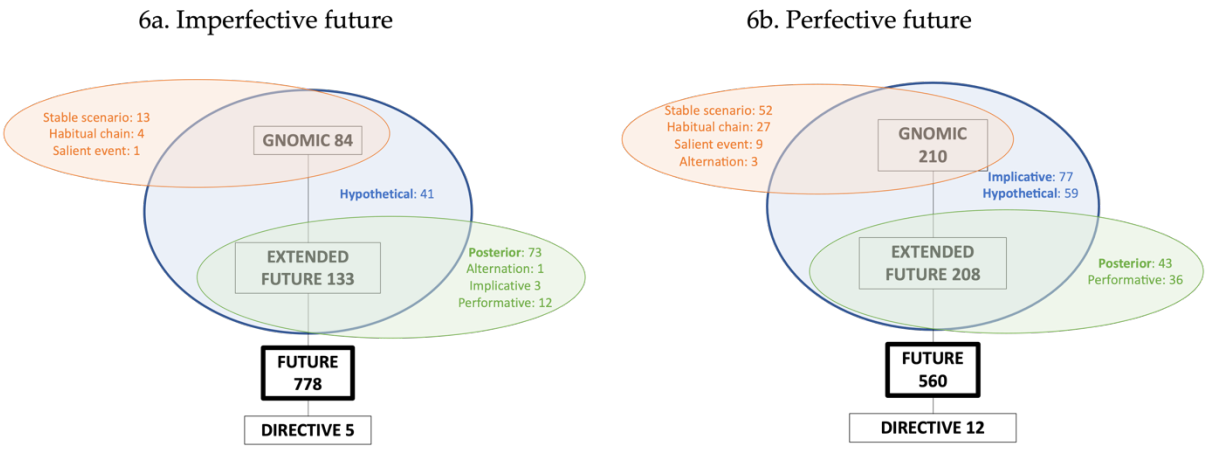

Figure 5a. Radial category of meanings expressed by imperfective Russian future forms. Figure 5b Radial category of meanings expressed by perfective Russian future forms. Major meanings are on the vertical axis, with Future as the prototype. Ovals represent submeanings, with boldface to indicate those that can overlap.

The radial categories of future tense meanings are multilayered. The first layer is represented by the four major meanings of the Russian future tense forms, which are arranged in square boxes, with a thick line around Future to indicate its status as the prototypical meaning in the network. All four of the major meanings are instantiated by both perfective and imperfective verbs. Directive is relatively marginal and therefore distant from the rest of the meanings. The remainder of the vertical axis shows relative temporal grounding of the meanings, with Future as the most grounded, followed by Extended future, which is partially grounded, and at an even further remove Gnomic, which lacks temporal grounding. The ovals are the second layer. They represent the various submeanings presented in Section 3. The number to the right of each meaning indicates the number of examples (out of a thousand) that illustrate the use of the meaning. Note that some examples express more than one

submeaning: for example, in the perfective dataset, eight examples combine Habitual chain with Stable scenario.

We find that the range of meanings is not entirely identical across the two aspects. While most of the submeanings are attested for both perfective and imperfective verbs, Salient events are found only with perfective verbs (with the exception of one example that can be potentially viewed as a Salient event). Bold face indicates submeanings that can combine with other submeanings, multiply motivating the use of future forms. Implicative and Hypothetical are shared by the Extended future and Gnomic meanings and can combine both with Posterior uses of the Extended future and with Stable scenario Gnomic uses.

The layered diagrams of the Russian future forms in Figures 5a and 5b capture the schematic simplicity of the radial category along with the complex interaction of both major uses and submeanings. In addition, Figures 5a and 5b show that perfective and imperfective futures are very alike despite the minor differences expressed in the numbers of examples with the attributed meanings.

6. Conclusion

A sample of 1000 examples of perfective future forms and 1000 imperfective future forms was manually analyzed for meanings and submeanings. Only 56% of perfective future tense forms unambiguously express Future time meaning, with the remainder expressing Extended future, Gnomic, and Directive meanings. Nearly 78% of imperfective future tense forms express Future time meaning, and the remainder follow a pattern similar to that of the perfective forms, differing in the use of some submeanings. We present a layered radial category that captures the complex interactions among the major meanings and the submeanings of the future forms. The definition of Extended future, Gnomic, and Directive uses constitutes a theoretical extension of Langacker's (2008) model and is potentially valuable for the analysis of tense in other languages.

Overall, in relation to imperfective future forms, we find that perfective future tense forms are more varied in their expression of meaning. Gnomic uses of future tense forms stand out as particularly important, both for a theoretical understanding of future tense forms and for language pedagogy, which should focus more on perfective forms and their Gnomic uses.

We also demonstrate that many future tense forms can be used in modal settings. However, the various types of modality do not directly correspond to the variety of

meanings of the future tense and thus are insufficient to fully motivate the radial category of Russian future tense. There is no clear pattern to the distribution of modality vs. future and non-future time meanings. We cannot explain the non-future and extended meanings entirely by recourse to modality. Modalities and (non-)future meanings can be combined freely. On the one hand, there are cases of one type of usage receiving more than one modal interpretation (by one or more researchers). For example, Future Performatives can be viewed as volitive or basic by Klimonow (2011) or have no modality according to Radbil (2011). On the other hand, the same type of modality can be represented by more than one type of use. For example, potential modality is combined with Future Hypothetical, Future, Gnomic Habitual chain, Gnomic, Gnomic Salient events, and Gnomic Stable scenarios.

References

Arutjunova, Nina D. 2011. Buduščee v jazyke. Logičeskij analiz jazyka. Lingvofuturizm. Vzgljad jazyka v buduščee. Moscow: Indrik.

Auwera, Johan van der, and Vladimir A. Plungian. 1998. Modality's semantic map. Linguistic Typology 2. 79–124.

Bondarko, Aleksandr V. 1971. Vid i vremja russkogo glagola. Moskva: Prosveščenie.

Bybee, Joan L., Revere Dale Perkins, and William Pagliuca. 1994. The evolution of grammar: Tense, aspect, and modality in the languages of the world. Vol. 196. Chicago: University of Chicago Press.

Dickey, Stephen M. 2000. Parameters of Slavic aspect. A cognitive approach. Stanford CA: CSLI Publications.

Fauconnier, Gilles. 1985. Mental Spaces. Cambridge/London: MIT Press.

Forsyth, James. 1970. A grammar of aspect: Usage and meaning in the Russian verb. Cambridge University Press.

Janda, Laura A., and Olga Lyashevskaya. 2011. Aspectual pairs in the Russian national corpus. Scando-Slavica 57.2. 201-215.

Janda, Laura A., Anna Endresen, and Robert J. Reynolds. 2019. Aspectual opposition and rivalry in Russian are not discrete: new evidence from experimental data. *Russian Linguistics* 43.3. 249-271.

Isačenko, Aleksandr. 1965/2003. Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim.: Morfologija. Moscow: Jazyki slavjanskoj kul′tury.

Karttunen, Lauri. 1971. Implicative verbs. Language 47(2). 340-358.

Klimonow, Wladimir D. 2011 Vzaimodejstvie modal′nosti i aspektual′nosti v sovremennom russkom jazyke. Vestnik Nižegorodskogo universiteta im. N.I. Lobačevskogo 6(2).

Kosheleva, Daria, and Janda, Laura A. Why markedness is always local. Manuscript.

Kratzer, Angelika. 1981. Partition and revision: The semantics of counterfactuals. Journal of Philosophical Logic 10(2)s. 201-216.

Lakoff, George. 1987. Women, fire, and dangerous things: What categories reveal about the mind. University of Chicago press.

Langacker, Ronald W. 2008. Cognitive grammar: A basic introduction. New York: Oxford University Press.

Langacker, Ronald W. 2019. Levels of Reality. *Languages* 4(2), 22.

Lewandowska-Tomaszczyk, Barbara. 2007. Polysemy, prototypes, and radial categories. *The Oxford handbook of cognitive linguistics*. 139-169.

Maslov, Jurij. 1990/2004. *Izbrannye trudy. Aspektologija. Obščee jazykoznanie*. Moskva: Jazyki slavjanskoj kul′tury.

Petruxina, Elena V. 2009. Russkij glagol. Kategorija vida i vremeni. Moskva: MAKS Press.

Petruxina, Elena V., and Li Čžuxong. 2015. Modal'nost' glagol'nyx form buduščego vremeni v russkom jazyke. *Vestnik Moskovskogo universiteta*, Serija 9, Filologija 4. 72-86.

Radbil, Timur. B. 2011. Buduščee kak fakt i buduščee kak modal′nost′ v paradigme «real′nost′-tekst ». Logičeskij analiz jazyka. Lingvofuturizm. Vzgljad jazyka v buduščee. 254-262.

Searle, John R. 1989. How performatives work. Linguistics and philosophy 12(5). 535-558.

Stojnova, Natalia M. 2016a. Nefutural'nye upotreblenija form buduščego vremeni. Materialy dlja proekta korpusnogo opisanija russkoj grammatiki. http://rusgram.ru (2 October, 2018).

Stojnova, Natalia M. 2016b. Analitičeskaja konstrukcija buduščego vremeni. Materialy dlja proekta korpusnogo opisanija russkoj grammatiki. http://rusgram.ru (2 October, 2018).

Švedova, Natalia, et al. 1980. Russkaja grammatika. Moskva: Nauka.

Talmy, Leonard. 2000. Toward a Cognitive Semantics. Cambridge, MA: MIT Press.

Vinogradov, Viktor V. 1947. *Grammatičeskoe učenie o slove.* Moskva: Russkij jazyk.

Wiemer, Björn, et al. 2020. Četyrexugol′nik snjatoj utverditel′nosti v russko-pol′skom sopostavlenii. *VII Meždunarodnaja konferencija Komissii po aspektologii Meždunarodnogo komiteta slavistov.* Moskva: Izdatel′stvo RGPU im. AI Gercena.

Zaliznjak, Anna A. 2015. Prezens soveršennogo vida v sovremennom russkom jazyke. *Dekostruktion and Konstruktion*. Festschrift für Ulrich Schweier zum 60. Geburtstag. München: Kubon & Sagner, 293-316.

III Russian future: an inside and an outside perspective

Daria Kosheleva

Department of language and culture, UiT The Arctic University of Norway

daria.kosheleva@uit.no

Abstract

Using the example of the future tense in Russian, in this article I show that relative frequency is important for native speakers, when they need to choose between perfective and imperfective aspect. At the same time, frequency does not have a similar effect on non-native speakers. In the experiment, native speakers and non-native learners of Russian completed the following task: they were asked to change the tense of a verb form in a sentence from past to future. The results show that both groups of respondents deviate from the expected answer when completing the task. Their answers differ qualitatively and quantitatively. I explore the factors that may motivate the respondents' choice and present a statistical model of the obtained data using mixed-effect logistic regression analysis. In addition, I highlight the most common deviations for both types of speakers.

1. Introduction

Russian verbal aspect is one of the most problematic parts of the curriculum for foreign language learners of Russian. For a learner, it might be difficult to use the correct aspect of the verb for referring to past events; however, expressing future time reference presents even more challenges.

In this article, via an experiment, I intend to point out the factors that contribute to the choice of the future form for the non-native speakers of Russian, what kind of deviations they produce when they use future tense forms, and compare these results to the performance of the native speakers.

Traditionally, Russian language instructors teach learners that there are two verb forms that express future meaning. As a rule, the imperfective form is given first. It is a periphrastic form comprised of the verb *byt'* 'be' in the future tense and an infinitive of an imperfective verb (*budu delat'* 'I will be doing'). The second is the perfective non-past form *sdelaju* 'I will do', which has the same inflectional morphology as the imperfective present[36]. In addition, the Russian language provides an extensive variety of other means for expressing future time reference, the most notable of which is probably the verb *stat'* accompanied by an imperfective infinitive (for a detailed analysis and comparison of periphrastic constructions with *byt'* 'be' and *stat'* 'become' cf. Stojnova 2019). Additional means are usually described in full in reference grammars (cf. Levine 2009, Timberlake 2004, Wade 2020).

Non-native speakers of Russian tend to overuse the imperfective future form (Swan 2017: 825). One reason for overuse of the imperfective future may lie in the fact that from the non-native perspective it is easier to produce an analytical form: all you need is the future paradigm of the verb *byt'* 'be' (i.e., just six forms) and an (imperfective) infinitive. Another notorious deviation is the combination of *budu* '(I) will' with a perfective infinitive: **budu sdelat'* 'I will do'.

Olshevskaya (2018) proposes an overview of learners' mistakes made in connection with verbal aspect. In her overview based on data from the Russian Learner Corpus (the RLC, web-corpora.net/rlc), Olshevskaya reports three types of mistakes made by learners of Russian while using the future tense. First, learners use the imperfective present instead of the perfective future, e.g., the imperfective *vozvraščaetsja* 'returns' instead of the perfective *vernëtsja* 'will return'. Second, the future auxiliary *budu* is combined with the perfective infinitives containing alternations *-ima-*/*-ja-*, *-yva-*/*-iva-*, -

[36] The perfective future can also be referred to as the non-past form or the perfective present (see the review in Kosheleva & Janda 2022).

nu-/-n-: **budet podnjat'* 'will uplift.PFV' instead of *budet podnimat'* 'will uplift.IPFV'. Note that this combination can also be interpreted as choosing the "wrong" imperfective auxiliary for the "correct" perfective aspect. And third, learners add the excessive *budu* 'will' auxiliary to the constructions with *stanu* 'become', e.g., **budet stat' bogatym* instead of *stanet bogatym* 'will become rich.'

The future tense forms are difficult not only for non-native speakers to learn but also for children to acquire, including children whose native language is Russian. Future tense as a category is acquired relatively late, after the past and the present tense forms have been fully established in the child's speech (Gvozdev 1961: 182-183). For a detailed review see Polinsky (2006: 18). Interestingly, Cejtlin (2000: 149) notes that children make the same mistake as non-native speakers: they use the auxiliary verb *byt'* 'be' in the future tense together with the perfective infinitive. Turian &Altenberg (1991: 219) show that the combination of *budu* and the perfective infinitive can also occur in the speech of a bilingual child whose first language (Russian) is going through attrition.

Adult native speakers of Russian do have certain preferences in the use of the future tense forms. As pointed out in (Kosheleva & Janda, under revision), in the Russian National Corpus the ratio between the perfective and imperfective future forms is 14:1, i.e., the perfective forms are much more frequent.

In this article, I present an experiment that addresses some of the major challenges that the future tense presents for learners of Russian and compare learners' performance with that of native speakers. I limit the discussion primarily to speakers' choice of aspect, setting aside for the time being morphological and orthographical errors. Morphological errors comprise a set of issues that emerge due to the overall rich morphology of the Russian language. Inflection and spelling in Russian are therefore a separate challenge that is not directly connected to the choice of the correct verb form in terms of tense and aspect. For example, in the experiment I got answers such as *napadast* instead of *napadet* 's/he will attack', *sproset* instead of *sprosit* 's/he will ask', and *naduvet* instead of *naduet* 's/he will inflate'. I treat these examples as correct future tense forms despite the morphological and/or orthographical errors.

My study confirms the discovery by Janda et al. (2019: 270) that Russian native speakers rely heavily on the relative frequency of aspectual verb forms. This frequency effect, as well as particular challenges faced by non-native (L2) speakers in choosing aspectual forms in the future tense, will be discussed in this article.

The article is structured as follows. The remainder of the article is divided into five sections. I set out the design of the experiment, both the materials and the procedure, in Section 2. The participants are described in Section 3. The analysis of the obtained data is presented in Section 4. I offer discussion in Section 5 and conclusion in Section 6.

2. Experimental design

The experiment makes it possible to investigate what factors are important for both types of speakers, native and non-native, when they are choosing a future tense verb form. In the experiment, the participants are tasked with changing the proposed past tense verb form into a future tense verb form in the corresponding aspect. The targeted form is placed in a context in a sentence and each sentence is illustrated with a picture.

2.1 Materials

Each participant received thirty-two tasks with one verb each. An overview of all the tasks in the experiment, showing how the selection of verbs gives a stratified representation of morphological types for marking of aspect in Russian is presented in Table 1. The column "Type" gives the information on the morphological type for marking the aspectual pair. There are three suffixes that form secondary imperfective verbs in Russian: *-vaj-*, *-ivaj-*, and *-aj-*. Each suffix is represented in eight verb pairs. In addition, there are eight pairs of verbs with high frequency, where aspect is marked with prefixes *po-* and *za-*, and a suppletive pair.

In Table 1, for each pair of verbs, the perfective verb is given first. For example, in the pair *umyt' / umyvat' 'wash', umyt'* is perfective and *umyvat'* is imperfective. Columns "G1" and "G2" are for Group 1 and Group 2 respectively, showing which aspect was used for the first and the second set of tasks. The tasks are given in the rightmost column. Only one sentence is given per verb pair, the verbs are separated with a slash. All sentences are short, between two and six words, and always contain a subject and a verb. Thus, I try to reduce the possible number of factors which might influence the choice of aspect in the context. And besides, the fewer words in the sentence, the easier it is for the non-native speakers of Russian to complete an already difficult task. It should be recognized that such short, simplified examples are rare in a corpus, therefore the sentences in the experiment are constructed by the author (who is a native speaker of Russian). Nineteen sentences contain a direct object (the word order is SVO). The subjects of the sentences are distributed as follows: thirteen are feminine, seventeen are masculine, one subject is neuter, and one is plural.

Type	Verb (PFV/IPFV)	G1	G2	Sentences in the experiment
-vaj-	*umyt' / umyvat'* 'wash'	IPFV	PFV	*Katja umyla / umyvala lico.* 'Katja washed (her) face.'
-vaj-	*dopit' / dopivat'* 'drink up'	IPFV	PFV	*Petja dopil / dopival kofe.* 'Petja drank up the coffee.'

-vaj-	*nadut' / naduvat'* 'inflate'	IPFV	PFV	*Katja nadula / naduvala puzyr'.* 'Katja inflated a bubble.'
-vaj-	*zalit' / zalivat'* 'extinguish'	IPFV	PFV	*Student zalil / zalival komp'juter čaem.* 'The student spilled the tea on the computer.'
-vaj-	*prožit' / proživat'* 'reside'	PFV	IPFV	*Sem'ja prožila / proživala v kvartire neskol'ko mesjacev.* 'The family resided in the apartment for several months.'
-vaj-	*odet' / odevat'* 'get somebody dressed'	PFV	IPFV	*Mama odela / odevala devočku v krasivoe plat'e.* 'Mom dressed the girl in a beautiful dress.'
-vaj-	*vypit' / vypivat'* 'have a drink'	PFV	IPFV	*Petja vypil / vypival nemnogo vodki.* 'Petja drank some vodka.'
-vaj-	*uspet' / uspevat'* 'be in time for'	PFV	IPFV	*Petja uspel / uspeval na poezd.* 'Petja caught the train.'
-ivaj-	*otporot' / otparyvat'* 'rip off'	IPFV	PFV	*Mama otporola / otparyvala pugovicu.* 'Mom ripped off the button.'
-ivaj-	*istolkovat' / istolkovyvat'* 'interpret'	PFV	IPFV	*Učitel'nica istolkovala / istolkovyvala novye slova.* 'The teacher interpreted the new words.'
-ivaj-	*peredelat' / peredelyvat'* 'make over'	PFV	IPFV	*Kompanija peredelala / peredelyvala kinoteatr v magazin.* 'The company converted the cinema into a store.'
-ivaj-	*rassejat' / rasseivat'* 'dissipate'	IPFV	PFV	*Solnce rassejalo / rasseivalo tuči.* 'The sun scattered the clouds.'
-ivaj-	*uderžat' / uderživat'* 'hold'	PFV	IPFV	*Petja uderžival / uderžal Katju na kraju obryva.* 'Petja kept Katja on the edge of the cliff.'
-ivaj-	*osmotret' / osmatrivat'* 'examine'	IPFV	PFV	*Šerlok Xolms osmotrel / osmatrival sledy.* 'Sherlock Holmes examined the footprints.'
-ivaj-	*zapisat' / zapisyvat'* 'write down'	PFV	IPFV	*Katja zapisyvala / zapisala plan v tetrad'.* 'Katja wrote the plan down in a notebook.'

-ivaj-	*sprosit' / sprashivat'* 'ask'	IPFV	PFV	*Mal'čik sprosil / spрašival učitel'nicu.* 'The boy (asked) the teacher.'
-aj-	*sgoret' / sgorat'* 'burn up'	PFV	IPFV	*Obed sgorel / sgoral.* 'The lunch burned up.'
-aj-	*sžeč' / sžigat'* 'burn'	IPFV	PFV	*Babuška sožgla / sžigala pis'ma.* 'Grandma burned the letters.'
-aj-	*napast' / napadat'* 'attack'	PFV	IPFV	*Tigr napal / napadal na čeloveka.* 'The tiger attacked a person.'
-aj-	*požat' / požimat'* 'shrug'	IPFV	PFV	*Petja požal / požimal plečami.* 'Petja shrugged (his) shoulders.'
-aj-	*snjat' / snimat'* 'rent'	PFV	IPFV	*Sem'ja snjala / snimala kvartiru.* 'The family rented an apartment.'
-aj-	*vybrat' / vybirat'* 'choose'	IPFV	PFV	*Petja vybral / vybiral dorogu.* 'Petja chose the road.'
-aj-	p*omoč' / pomogat'* 'help'	IPFV	PFV	*Petja pomog / pomogal stariku.* 'Petja helped the old man.'
-aj-	*otgresti / otgrebat'* 'shovel away'	PFV	IPFV	*Marija otgrebla / otgrebala sneg.* 'Maria shoveled away the snow.'
po-	*kupit' / pokupat'* 'buy'	IPFV	PFV	*Katja kupila / pokupala novuju sumku.* 'Katja bought a new bag.'
po-	poslušat′ / slušat′ 'listen'	PFV	IPFV	*Mal'čik poslušal / slušal muzyku.* 'The boy listened to the music.'
sup.	*brat' / vzjat'* 'take'	IPFV	PFV	*Petja vzjal / bral karandaši.* 'Petja took the pensils.'
po-	*podarit' / darit'* 'present'	PFV	IPFV	*Petja podaril / daril Kate cvetok.* 'Petja gave Katja a flower.'
po-	*pospat' / spat'* 'sleep'	IPFV	PFV	*Mal'čik pospal / spal.* 'The boy slept.'
po-	*postroit' / stroit'* 'build'	PFV	IPFV	*Katja postroila / stroila dom.* 'Katja built a house.'
po-	*pozvonit' / zvonit'* 'call'	IPFV	PFV	*Petja pozvonil / zvonil.* 'Petja called.'

za-	*zaplatit' / platit'* 'pay'	PFV	IPFV	*Katja zaplatila / platila za pokupki.* 'Katja payed for the purchases.'

Table 1. Tasks in the experiment.

2.2 Procedure

The experiment was conducted in the form of an internet-based survey. Potential participants received a link to the survey, and could fill it out for an unlimited amount of time. Under the conditions of the experiment, the participants could not use any help (e.g., a dictionary, a grammar, or another person). However, there was no means to control if this was indeed the case.

The survey comprises the following sections: the introduction (including the information about the participants and consent); the instructions; and the tasks themselves. Below I describe each section in detail.

When the participant opens the link, they see an introductory block. The introduction contains a short description of the survey, stating the general topic of the experiment, the terms of data use and the consent form. If the speakers consent to participating in the experiment, they are further asked to share information about themselves. The demographic information about the participants includes their age range (18–30, 31–45, 46–60, 61–75 and 75+ following Labov 1972), gender (female, male, or not specified) education (secondary, secondary professional, or higher education) and their relation to linguistics. In addition, the participants are asked to indicate their birth month (January – June vs. July – December) to randomly divide the people into two groups (Group 1 and Group 2), each of which received a different set of tasks (see Table 1).

Since the experiment is for both native and non-native speakers of Russian, I also ask for the native language of the respondent. To control for heritage speakers, I ask whether at least one of the parents of the respondent speaks Russian. The non-native respondents are additionally asked about their proficiency in Russian. After thorough consideration, I chose 'years studied' as a metric. This is not an ideal metric because people learn languages with difference pace, and may interpret, what the learning process means differently. Nevertheless, this option is most acceptable within the frame of this experiment. Other options would be, for example, self-assessment or asking the respondent to provide a language certificate. First, self-assessment of language proficiency can lead to under- or overestimation of the learner's ability to use a foreign language (MacIntyre et al. 1997), so if I ask just for the CEFR level, I might also get inaccurate results; in addition, not every learner knows the CEFR system, and I would like to avoid giving unnecessary explanations. Second, language proficiency tests for the Russian language, such as TORFL (Test of Russian as a Foreign Language), are not

as popular as similar tests for English (TOEFL, IELTS). The test requirement would significantly reduce the number of potential participants in the experiment.

The instructional component is purposefully simplified, so that respondents with different levels of Russian language proficiency can understand what they need to do: "In the tasks on the next page, you need to put (change / rewrite) the verbs in brackets in the future tense form." The participants are given a sentence containing one verb in past tense and they need to change the tense of the verb. Before the participants begin completing the tasks, they see the following model. Two examples with perfective and imperfective verbs respectively, as well as the expected answers, are given as below:

Prompt for a perfective verb:

Katja (sdelala) domašnee zadanie. 'Katja did the homework.'

Correct answer:

sdelaet 'will do'

Prompt for an imperfective verb:

Katja (delala) domašnee zadanie. 'Katja was doing the homework.'

Correct answer:

budet delat' 'will be doing'

After reading the instructions, the respondents proceed to the tasks section where they are reminded not to use any external help, i.e., a dictionary or a machine translation service. Each respondent receives a set of 32 short sentences. Every sentence is illustrated with a picture and contains one past tense verb form. As mentioned above, the respondents are randomly divided into two groups. Groups 1 and 2 test the same 32 verb pairs but with complementary distribution of verbal aspect: for every perfective verb in Group 1 there is a corresponding imperfective verb in Group 2, and vice versa. The tasks are organized in this way so that I can test both perfective and imperfective aspects for each verb pair without giving any participant more than one aspect for a given verb pair. Figure 1 illustrates a task from the experiment. The sentence contains a relatively frequent verb *vybirat'* 'choose' (50.97 ipm in the RNC), a subject, and a direct object. The sentence is paired with a picture of a person who must choose a direction (and supposedly walk along the road). The picture is intended to illustrate the sentence and help respondents who might have difficulty understanding it.

Figure 1. Petja vybiral dorogu. 'Petja was choosing the road.'

All three components (the introduction, the instructions, and the tasks) are given in Russian. I decided not to translate everything into English because it would create an uneven situation for the participants with twenty-two different native languages. While English is a commonly recognized international language, I could not assume that all participants would have equally good understanding of English. If they learn Russian as a foreign language, that can mean that Russian is the only foreign language that they know. To provide equal conditions for everyone, and to ensure that the experiment is not affected by another language, the decision was made not to use English. This decision was made to conform to The Norwegian National Ethics Committee Guidelines (NESH guidelines 2016: 15) stating that "Researchers must ensure that the participants have actually understood the information".

The use of Russian in the consent form and all the explanations was a deliberate choice. It was impossible to predict from the outset what languages would be native to the L2 participants and to translate the questionnaire into each mother tongue. The participants were able to give feedback at the end of the survey. I claim that the choice of the language of instruction was both practically and ethically justified. On the one hand, to complete the survey successfully, one needs quite good competence in Russian, equivalent to at least CEFR level A2+ or B1. And for most items, the picture illustrating the meaning of the sentence should help. Not all the verbs in the questionnaire are expected to have been mastered at the B1 level, but this was part of the design of the experiment, namely, to find out whether L2 speakers can guess what form they should use based only on the morphology of the past tense form. If a person could not understand the disclaimer that served as a consent form, they would not be able to complete the questionnaire. It was furthermore obvious if a person did not understand the task: for example, four participants just wrote down infinitives or nonsense words that do not resemble any actual Russian verbs. These participants were excluded from analysis. Only one participant commented that Russian instructions might pose a barrier for someone with a lower level of proficiency to fill out the survey.

3. Participants

Initially, 173 people responded to the call to fill out the survey, however I had to eliminate several participants. Among the persons claiming to be non-native speakers of Russian, I eliminated heritage speakers; these are people who answered positively to the question about whether their parents spoke Russian. Eleven participants reported that at least one of their parents spoke Russian. Secondly, I eliminated participants who clearly did not understand the tasks. If the person did not answer anything (e.g., put a ".") or all the answers were the same (e.g., infinitives only, no future tense forms), their responses were not relevant to the experiment. Four participants fell into this category.

As a result, I received a total of 160 completed surveys. The distribution of the answers is presented in Table 2. There are 78 native speakers out of which 48 speakers completed the Group 1 task. There are 82 answers given by non-native speakers; 41 of them worked with the Group 1 questionnaire and another 41 worked with the Group 2 questionnaire.

	Group 1	Group 2	Total (Group 1 + Group 2)
Native Russian speakers	48	30	78
Non-native speakers	41	41	82
Total (Native + Non-native)	89	71	160

Table 2. The distribution of the participants in terms of "nativeness" and the Group of tasks they were assigned to.

I briefly describe the main characteristics of the answers for each group in subsection 4.2 below. The full anonymized dataset is available at https://dataverse.no/privateurl.xhtml?token=f04bd79a-299e-4e5e-80fb-0efc02257e4f. In subsections 3.1 and 3.2, I present the more detailed information about the participants.

3.1 Native speakers of Russian

Most of the participants who stated that Russian is their native language are female (54 individuals). Male participants comprise about one third (21 individuals) and three individuals did not want to state their gender.

I organized the possible age ranges in the following way: 1) 18—30; 2) 31—45; 3) 46—60; 4) 61—75. These are traditional age groups established by (Labov 1972: 22). Table 3 gives summary information about gender and age characteristics of the participants. Quite often, the prototypical people interested in learning something and eager to fill in a questionnaire about it, are young and female. My sample is also not without this bias towards young females who are the largest group (37 people) involved in the survey. The consequence of this may be better results (i.e., more "correct" = expected answers) in the sample compared to the overall picture in real life.

Age group		Gender	Type of questionnaire

	Number of participants	Male	Female	Not stated	Group 1	Group 2
18—30	57	17	37	3	32	25
31—45	13	1	12	0	8	5
46—60	4	1	3	0	4	0
61—75	4	2	2	0	4	0

Table 3. The distribution of the native speakers according to Age group, Gender, Type of questionnaire.

The participants have various educational backgrounds. Most of them – 73 individuals – have a university degree (BA, MA, or PhD). Three individuals have secondary education (high school) and two reported that they have a specialized secondary education (which, in the Russian educational system, requires more training).

3.2 Non-native speakers of Russian

Non-native speakers (or learners) of Russian are asked the same set of questions as native speakers with two additional questions about their native language and their knowledge of Russian.

There are 47 female participants, 34 male participants and one person wished not to reveal their gender. The non-native participants' age ranges between 18 and 75+: the 18—30 group is the biggest with 65 people; the runner up is the 31—45 group comprised of 11 people. There are four participants aged between 46—60 years. One person is between 61 and 75 years old and one more person is more than 75 years old. Table 4 shows the distribution of the non-native respondents in terms of age, gender, and the set of tasks they received.

Age group	Number of participants	Gender			Type of questionnaire	
		Male	Female	Not stated	Group 1	Group 2
18—30	65	26	38	1	32	33
31—45	11	5	6	0	5	6
46—60	4	1	3	0	4	0
61—75	1	1	0	0	0	1
75+	1	1	0	0	0	1

Table 4. The distribution of the non-native speaker according to Age group, Gender, Type of questionnaire.

Most of the non-native participants hold a university degree: 74 people reported that they have a BA, MA, or PhD. Six people have received specialized secondary education and two people graduated from high school.

Figure 2 shows that the participants come from all parts of the world: the graph shows the total numbers of the participants with various mother tongues. The biggest group are native English speakers with thirty participants, followed by Vietnamese with ten participants, and French with nine participants. In total, native speakers of twenty-two languages participated in the questionnaire.

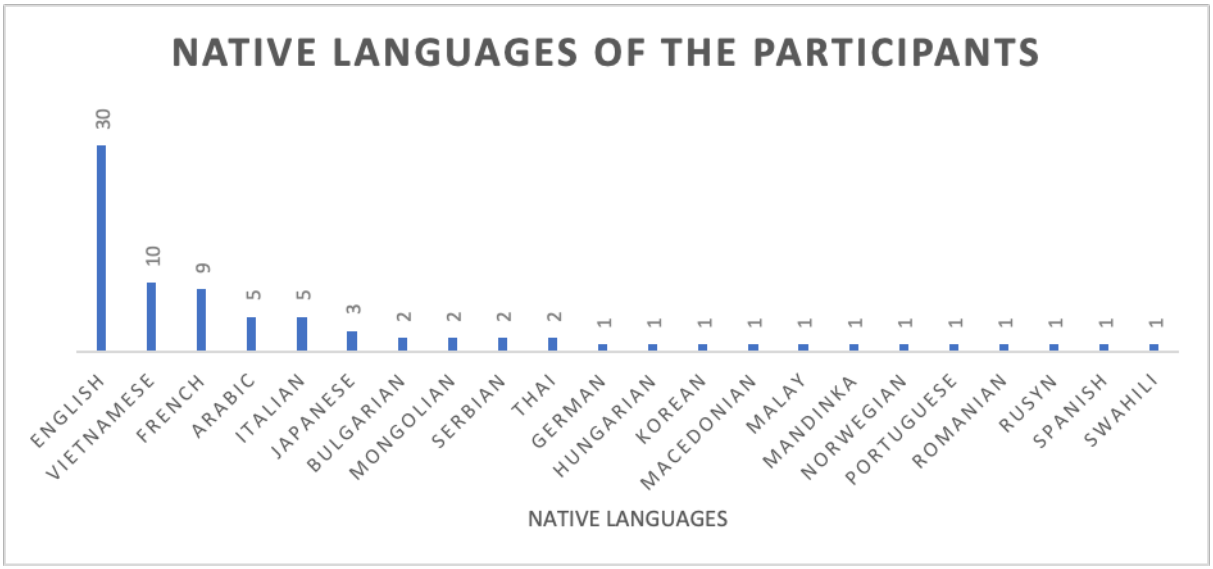

Figure 2. The number of the non-native respondents according to their native languages.

Figure 3 visualizes the information about the number of participants who have been learning Russian for different periods of time. The x-axis shows the number of years, and the bars indicate the number of respondents for each option. The respondents received various amounts of training in Russian: from six months up to fifty-three years. The mode for time spent learning Russian is four years (19 participants). The average time of training is five-and-a-half years.

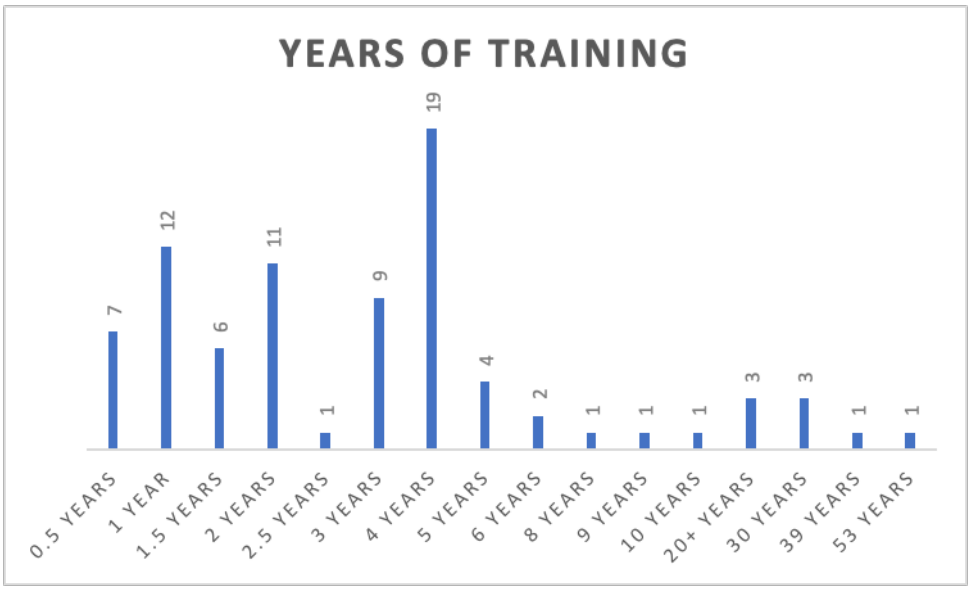

Figure 3. Received years of training in Russian.

Participants have been learning Russian in a variety of places. Since the question "where did you learn Russian" was not obligatory, not everyone replied to it and the responses vary in specifics and form. Most participants have been learning Russian at various universities (Higher School of Economics in Russia, Harvard University, University of

Genoa etc.), schools (particularly in Vietnam) and language courses. Some people reported that they had been learning the language on their own.

Based on the obtained information, it is safe to assume that the respondents have different levels of language proficiency varying from intermediate to highly advanced and near-native.

4. Data analysis

In this section, I provide the statistical analysis of the data obtained in the experiment (subsection 4.1) followed by a description of various deviation patterns in the data produced by the native and non-native learners (subsection 4.2).

4.1 Statistical analysis: mixed effect logistic regression

I undertook a mixed effects logistic regression analysis of the experimental data, following the procedure detailed in Gries (2021: Chapter 6.4). Table 5 presents the variables and their levels, described in more detail below. The results of the analysis are summarized in Table 5 and their visualization is presented in Figure 4.

Type of Variable	Name of Variable	Levels	Comments
Response (binary)	Match	yes, no	Does the form produced match the aspect in the stimulus?
Fixed effect predictor (categorical)	OrigAspect	PF, IPFV	What was the aspect in the stimulus? PF = Perfective, IPFV = Imperfective.
Fixed effect (numerical)	RelFreq	-3.655938 3.655938	The logarithm of the relative frequency of the verb presented in the stimulus as compared with the frequency of the verb of the opposite aspect? The variable is a numeric ranging between (-3.655938; 3.655938).
Fixed effect (categorical)	SpeakerType	native, non-native	Was the participant a native speaker of Russian or not?
Random effect	ID_individual	160	Who was the participant?
Random effect	ID_stimulus	32 verb pairs	Which stimulus was presented?

Table 5. The variables for the statistical analysis.

- Match compares the original aspect with the aspect chosen by the respondent. The value is "yes" if the aspect in the stimulus is the same as the aspect in the response, and "no" if it is the opposite aspect. If the respondent used the incorrect construction '*budet* + perfective infinitive', Match was coded as "yes" because the aspect in the response matched the aspect in the stimulus.
- OrigAspect indicates the aspect of the verb given in the stimulus. The verbs in the stimuli are either imperfective (IPFV), or perfective (PF). I did not give the respondents any Biaspectual verbs.
- RelFreq presents the relative frequency of a given verb as opposed to its aspectual partner verb. In other words, this factor indicates whether the verb presented in the stimulus is more frequent than the verb of the opposite aspect. For example, if the verb in the stimulus is *otparyvat'* 'rip off' (IPFV), then the verb that it is compared to is *otporot'* 'rip off' (PF). First, I look for the frequency of both verbs in the Russian National Corpus. I divide the frequency of the verb in the stimulus by the frequency of the verb of the opposite aspect. In the final step, I take the natural logarithm of the resulting number. This procedure makes it possible to represent relative frequency on a symmetrical scale (see more in Janda & Reynolds 2019: 480-481). RelFreq ranges between (-3.655938; 3.655938). Zero on this scale means that the two verbs are of equal frequency. A positive number indicates that the given verb is more frequent than its aspectual partner verb, whereas a negative number indicates that it is less frequent. As an example, let us consider the verb pair *otporot'* / *otparyvat'* 'rip off'. The total number of words in the Russian National corpus is 283,431,966 (in the old version). There are 97 occurrences of the perfective verb *otporot'* 'rip off' and 23 occurences of the imperfective verb *otparyvat'* 'rip off'. These two numbers correspond to 97 / 283.431966 = 0.34 ipm and 23/ 283.431966 = 0.08[37] ipm respectively. For the perfective *otporot'*, the frequency of the stimulus divided by the frequency of the verb of the opposite aspect is (0.34 / 0.08 =) 4.22. For the imperfective verb *otparyvat'* the relative frequency is (0.08 / 0.34 =) 0.24. Finally, I calculate the natural logarithms of the relative frequencies. The natural logarithm of the relative frequency for the perfective stimulus *otporot'* is 1.44 and the natural logarithm of the relative frequency for the imperfective stimulus *otparyvat'* is -1.44. This procedure was carried out for all the verb pairs in the stimuli.
- SpeakerType indicates whether the participant was a native speaker of Russian (78 participants) or a non-native learner of Russian (82 participants).
- ID_individual gives a code for each individual participant.

[37] The numbers are rounded to the last two digits in this text. Exact numbers were used in all calculations.

- ID_stimulus gives a code for each individual stimulus.

In designing the model, I took the following considerations into account. The most important research question to be addressed by the model is whether native speakers are more influenced by relative frequency than non-native speakers. This question motivates modeling an interaction between relative frequency and speaker type. The original aspect is also a possible main effect since imperfective past tense events are more likely to be rendered as perfective future events than vice versa. To control for any effects that can be attributed to either individual participants or stimuli, I include those as random effects. This motivates the following formula for the regression model:

Match ~ 1 + OrigAspect + RelFreq * SpeakerType + (1 | ID_individual_c) + (1 | ID_stimulus_c)

This formula can be rendered in prose as: "I model Match as predicted by OrigAspect, RelFreq and SpeakerType as main effects, with an interaction between RelFreq and SpeakerType, and with random intercepts for ID_individual and ID_stimulus".

The formula was evaluated in R using the lme4 package and returned the following results visualized in Table 6 and Figure 4:

	Match		
Predictors	*Log-Odds*	*CI*	*p*
(Intercept)	2.79	2.22 – 3.35	**<0.001**
OrigAspect [PF]	0.93	0.71 – 1.15	**<0.001**
RelFreq	0.94	0.79 – 1.09	**<0.001**
SpeakerType [nonnative]	-1.06	-1.76 – -0.35	**0.003**
RelFreq * SpeakerType [nonnative]	-1.42	-1.61 – -1.23	**<0.001**

*Table 6.Results of mixed effects logistic regression for Match ~ 1 + OrigAspect + RelFreq * SpeakerType + (1 | ID_individual_c) + (1 | ID_stimulus_c)*

The intercept represents the prediction of the "yes" value of Match when the original aspect is imperfective, the relative frequency is 0 (identical frequency of both perfective and imperfective forms), and the form was produced by a native speaker of Russian. The Log-Odds is 2.79 in favor of the perfective aspect. The rest of the Log-Odds are calculated by adding or subtracting all the other Log-Odds from the intercept (see Table 6). The main effects and the interaction are very significant. The p-value <0.001 for two main effects (original aspect and relative frequency) and the interaction, and p-value = 0.003 for the speaker type. The 95% confidence intervals (CI) are narrow and do not

cross zero, which shows that the obtained numbers are reliable. The intercept indicates a high probability for the Match as "yes", and this probability further increases both when the original aspect is perfective and when the relative frequency of the original aspect is higher than that of the opposite aspect. However, the prediction of "yes" for Match decreases when the speaker type is non-native, and when this factor interacts with relative frequency.

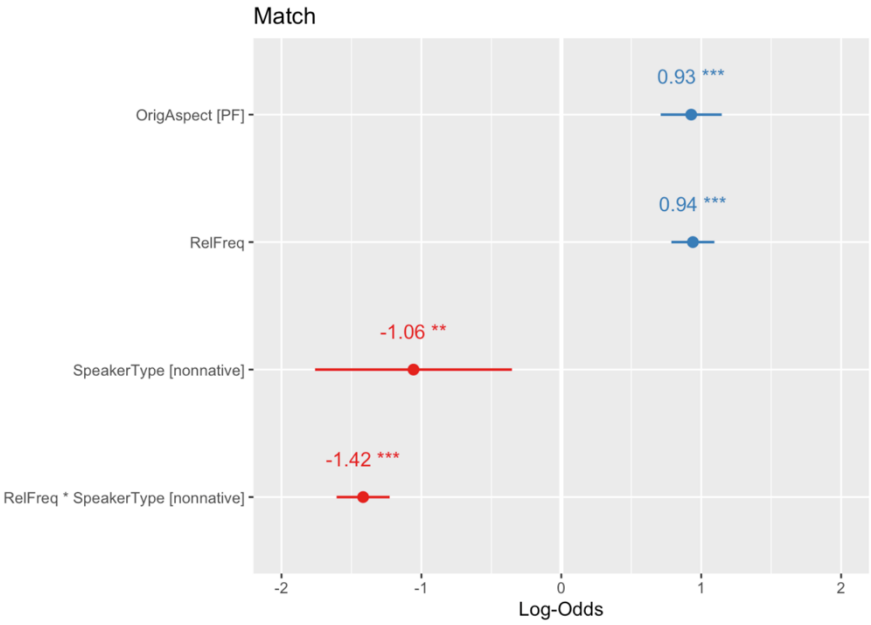

Figure 4.Figure 4. The visualization of the logistic regression analysis presented in Table 6.

Overall, the model is highly significant (LR-test = 434.14, df = 4, $p < 0.0001$) and its relative likelihood (of > 3e+92) makes it vastly more plausible than a null model (only with the random effects). These results indicate that all the predictor variables in the formula were highly significant, and their confidence intervals did not cross zero. Perfective aspect and higher Relative frequency both are associated with a "yes" value for Match, whereas the Speaker type of "non-native" as well as the interaction of Relative frequency with Speaker type "non-native" is more likely to be associated with a "no" value for Match.

Variance inflation factor values are all 4.05 or lower, indicating that there was no collinearity among the variables and a test for overdispersion showed no significant problem there either. A drop1() test was performed to see whether any of the variables could be deleted, but it was shown that all the variables in the model are very significant. The C-score is 0.897, which indicates an excellent model (over 0.8).

Overall, in 81.75% responses Match = "yes", so this is the baseline for evaluating the predictions of the model. Table 7 is a confusion matrix which compares the observed values for the predicted variable Match (whether the original aspect matches with the aspect in the answer) to the predicted values. The observed values are shown in the rows,

and the predicted values are shown in the columns. So, in 454 cases the model predicted a “no” that was observed as a “no”, but in 480 cases the model predicted a “yes” that was observed as a “no”.

Observed \ Predicted	“no”	“yes”
“no”	454	480
“yes”	135	4051

Table 7. Confusion matrix for the predicted variable Match.

The regression model accuracy is 87.99%, which is above the baseline (81.75%) and the difference is significant (3.619569e-34, in other words $p<0.001$).

Figure 5 visualizes the model predictions for the interaction between relative frequency and the type of the speaker. This interaction yields the largest deviation (-1.42) from the intercept. At the bottom of Figure 5 is a “rug”, a row of ticks showing the distribution of the datapoints for relative frequency. These ticks are distributed symmetrically because both aspects of the verb pairs were represented in the stimuli. That is, the most important information in Figure 5 is concentrated above the area with the highest concentration of the ticks. The two lines represent the model predictions for the two types of speakers, each with a “ribbon” representing the 95% confidence interval of the prediction. The black line represents the native speakers; the light grey line shows the non-natives. The native speakers of Russian rely highly on the relative frequency of the verb. Most of the relative frequencies fall between -1.75 and 1.75, which corresponds to the middle of the graph. The predicted probability of getting a “yes” Match in this range increases from 0.8 to 1 as the relative frequency grows. On the contrary, the non-native speakers show a mostly flat relationship to relative frequency, with perhaps a small decline in the probability of getting a “yes” Match.

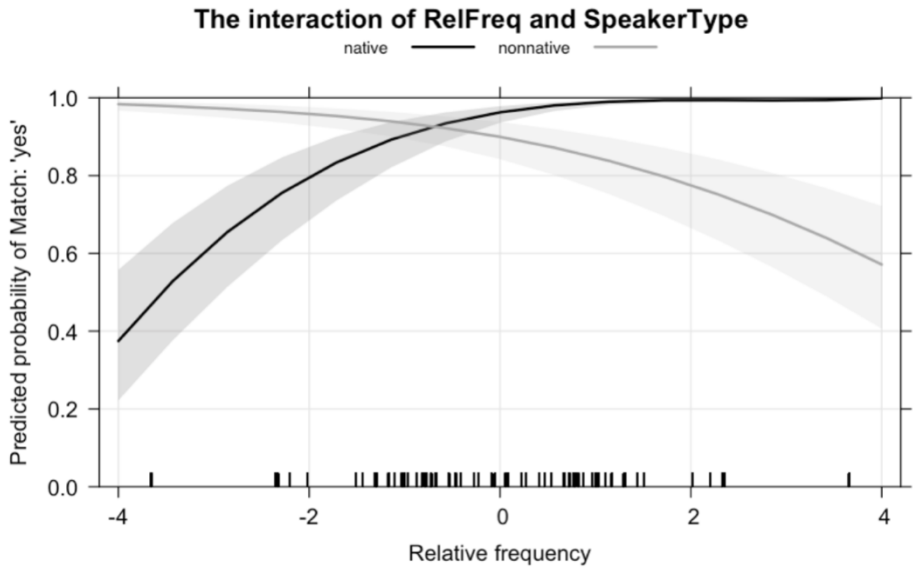

Figure 5. Model predictions of a matched aspect for the interaction of Relative Frequency and Speaker Type.

In sum, the model shows that the aspect of the verb in the stimulus (OrigAspect, perfective or imperfective), the relative frequency (-3.655938; 3.655938), the type of the speaker (native or non-native) and the interaction between the relative frequency and the type of the speaker all contribute to the choice of the aspect of the verb in the answers of the respondents. Native speakers have a strong tendency to choose the aspect that is of higher frequency. Learners of Russian are not sensitive to frequency in this way.

4.2 Deviation types

In this section, I turn to concrete examples of deviations in the responses from the expected answers. Regarding the aspect of the verb form, native speakers choose "wrong" aspect in 436 out of 2496 responses. There are only 23 responses in which the respondent chose imperfective aspect when the aspect of the stimuli was perfective. The majority of the aspect deviations (413 out of 436 non-matching responses) made by native speakers occur when the stimuli are in the imperfective aspect and the response contains a verb form in the perfective aspect.

The non-native speakers have given 498 responses containing the "wrong" aspect. Their behavior (as shown by statistical analysis) is different, and they are not so consistent: there are 191 responses with the perfective aspect to the stimuli with imperfective aspect, and 307 responses containing imperfective future forms when the aspect of the stimuli is perfective.

The obtained responses contain not only "mistakes" regarding the aspect of the verb form. There are several other types of deviations, some of which are shared by both types of speakers and some of which are specific to the non-native speakers. The responses of the native speakers do not contain any unique types of deviations.

Both native and non-native speakers of Russian exhibit the following types of deviations:

- usage of synonymous verbs (both perfective and imperfective);
- preservation of the past tense;
- various misspellings and typos;

illustrated by the examples below.

In Example (1 a—c), a respondent used an imperfective synonym *žit'* 'live' of the perfective verb *prožit'* 'live' in the stimulus.

(1) a. Sem'ja **prožila** v kvartire neskol'ko mesjacev. 'The family **lived** in the apartment for several months.'
b. Sem'ja **proživët** v kvartire neskol'ko mesjacev. 'The family **will live** in the apartment for several months.'
c. Sem'ja **budet žit'** v kvartire neskol'ko mesjacev. 'The family **will be living** in the apartment for several months.'

All examples in this section have the same structure. (1.a) is the stimulus, (1.b) is the expected response, (1.c) is the observed response. The future form of the imperfective non-prefixed verb *budet žit'* 'will be living' was used in the response (1.c) instead of (1.b). There are 16 cases of this deviation type reported for the native speakers, and 48 cases for the non-native speakers. If the synonymous verb in the response was of the same aspect as the verb in the stimulus, such pairs do not count as an aspectual mismatch in the analysis described in subsection 4.1.

In Example (2), the respondent simply copied the past tense form of the verb *rasseivat'* 'scatter':

(2) a. Solnce **rasseivalo** tuči. 'The sun **was scattering** the clouds.'
b. Solnce **rasseivalo** tuči. 'The sun **was scattering** the clouds.'
c. Solnce **budet rasseivat'** tuči. 'The sun **will be scattering** the clouds.'

Example (2.a) shows the stimulus, (2.b) is the observed response; (2.c) is the expected response. The respondent merely copied the form *rasseivalo* 'was scattering' from the stimulus rather than producing a future tense form. There are four responses from the native speakers and 49 responses from the non-native speakers for this type of the deviation. All of them count as aspectual matches since the original aspect is preserved. In addition, there is one response that belongs to a non-native speaker, who used a perfective verb in the past tense form. The stimulus contained an imperfective verb, so this response is recorded as a mismatch, since the aspect in the response is different from the aspect in the stimulus.

Example (3) shows a misspelling. The only verb form that the respondent could have meant by *soščët* is *sožžet*. In addition to the misspelling, the verb form in response differs from the verb form in the expected answer in terms of aspect: the expected response

(3.b) has imperfective aspect, whereas the actual response (3.c) is in the perfective aspect.

(3) a. Babuška **sžigala** pis'ma. 'Grandma **was burning** the letters.'
b. Babuška ***budet sžigat'*** pis'ma. 'Grandma **will be burning** the letters.'
c. Babuška ***soszët*** pis'ma. 'Grandma **will burn** the letters.'

In the native speakers' group, misspellings occurred mainly when conjugating infrequent verbs, such as *otparyvat'/otporot'* 'rip off' and *sžigat'/sžeč'* 'burn'. There are 48 native responses containing misspellings and typos. None of the native misspellings count as mismatches.

Non-native speakers have more challenges when conjugating Russian verbs in the future tense. The total number of responses among non-native speakers containing this type of deviation is 337. In most cases, misspellings do not affect the status of a response as a mismatch. I.e., if the verb form in the stimulus and the form in the response differ in terms of aspect, the response is counted as a mismatch.

However, it is not always clear if the deviation is just a mixture of typos and misspellings, or that the respondent does not know how to conjugate the verb properly. There are eighteen unclear non-native responses, all of them are counted as mismatches.

The abovementioned deviations are shared by both types of speakers. In addition, the non-native speakers exhibit four other distinct deviation types.

First, in 289 responses non-native speakers used the present tense instead of the future tense. In 44 out of 289 responses, the verb in the stimulus was in perfective aspect, so these responses were counted as mismatch. In Example (4.c), the respondent used the present tense form of the imperfective verb *proživat'* 'live'. The verb form *prožila* 'lived' in the stimulus (4.a) is the past tense form of the perfective verb *prožit'* 'live'. The expected answer is shown in (4.b)

(4) a. Sem'ja **prožila** v kvartire neskol′ko mesjacev. 'The family **lived** in the apartment for several months.'
b. Sem'ja **proživet** v kvartire neskol′ko mesjacev. 'The family **will live** in the apartment for several months.'
c. Sem'ja **proživaet** v kvartire neskol′ko mesjacev. 'The family **lives** in the apartment for several months.'

Next, non-native speakers of Russian use the future auxiliary (*budu*) in combination with the past tense verb form (cf. Example 5). The (5.a) is the stimulus, (5.b) is the expected answer; and (5.c) is the response with the future auxiliary combined with the past tense form of the verb *snjat'* 'rent'. There are thirteen responses in this group; the respondents are native speakers of English, Mandinka, Mongolian, and Norwegian.

(5) a. Sem'ja **snjala** kvartiru. 'The family **rented** an apartment.'
b. Sem'ja **snimet** kvartiru. 'The family **will rent** an apartment.'

c. Sem'ja **budet snjala** kvartiru. 'The family **will rent (lit. will rented)** an apartment.'

Third, the non-native responses exhibit the use of the future auxiliary *budu* in combination with the present tense verb form (cf. Example 6).

The (6.a) is the stimulus, (6.b) is the expected answer; and (6.c) is the response with the future auxiliary combined with the present tense form of the verb *otgrebat'* 'shovel away'. There are only three such responses given by one person.

(6) a. Marija **otgrebala** sneg. 'Maria **was shoveled** the snow away.'
b. Marija **budet otgrebat'** sneg. 'Maria **will shovel** the snow away.'
c. Marija **budet otgrebaet** sneg. 'Maria **will (is) shoveling** the snow away.'

Last but not the least popular, the most salient source of aspect mismatch are the responses with the notorious "overuse" of the future auxiliary *budu* with perfective infinitives (Swan 2017: 825). This mismatch is illustrated with Example (7). The (7.a) is the stimulus, (7.b) is the expected answer; and (7.c) is the response with the imperfective future auxiliary combined with the infinitive of the perfective verb *umyt'* 'wash'. There are 202 responses of this type, all of which are counted as mismatches. This type of deviation can be open to interpretation. The combination of imperfective auxiliary and perfective infinitive can be seen as a deviation in terms of the "type of the future": synthetic or analytical, and in this case, there is no mismatch. However, I am under the impression that the speaker did not recognize the verb aspect in the stimulus, and interpreted the combination as an imperfective future, i.e., a mismatch.

(7) a. Katja **umyla** lico. 'Katja **washed** her face.'
b. Katja **umoet** lico. 'Katja **will wash** her face.'
c. Katja **budet umyt'** lico. 'Katja **will wash** her face.'

Finally, some responses contain several types of deviations. For instance, Example (8) shows a synonymous verb used in the present tense instead of the future.

(8) a. Petja **dopival** kofe. 'Petja **was finishing** the coffee.'
b. Petja **budet dopivat'** kofe. 'Petja **will be finishing** the coffee.'
c. Petja **popivaet** kofe. 'Petja **sips** the coffee.'

(8.a) is the stimulus, (8.b) is the expected answer; and (8.c) is the response with the present tense form of the verb *popivat'* 'sip' instead of *dopivat'* 'drink up/finish'.

As the responses demonstrate, native speakers are not an exception when it comes to deviating from the expected answers. Non-native speakers share several types of deviations with native speakers. Those include using synonymous verbs, past tense verb forms, and some issues with normative orthography. However, non-native speakers have more deviations in their responses. Also, non-native responses exhibit additional deviations, such as using present tense forms, or combinations of *budu* with the past or present tense form.

5. Discussion

Native speakers' performance depends highly on the relative frequency: the more frequent the verb in the aspectual pair is, the bigger the chance that it will be used in the response. This is not exactly surprising: frequency plays a central role in acquisition, usage, and development of language (Bybee & Hopper 2001, Bybee 2003, 2006). Native speakers have life-long input that they can rely on. The native language can be viewed as a corpus that is stored in the head of the speaker (Taylor 2012). The native speaker can use this corpus to judge what is "natural" or "unnatural" use of the future tense, and then base their responses in the survey on these judgements.

However, the frequency effect for the native speakers may at least partially be the consequence of the design of the experiment. As pointed out by Dickey (ms., 11-13), this preference can be explained by how people conceive past and future events through different mechanisms: past events are processed through episodic memory (Tulving 2002: 3) and future events are conceived as simulation, prediction, intention, and planning (Szpunar et al. 2014 cited by Dickey ms.: 11). According to Dickey (ms.: 11), the imperfective aspect in the future tense is used less because future tends to be a prediction of a completed event rather than a complex of either ongoing processes or repeated or habitual series. It is not entirely clear how the cognitive mechanisms responsible for the memory interact with tense in a foreign language.

The results that I received in the experiment may not look bright for the language learners. It seems that many learners would never achieve native-like performance because of the lack of exposure to the language. It is not realistic to expect people to spend thousands of hours mastering their language skills. A learner must prioritize what part of grammar they need most at their current stage of learning. On the other hand, if a learner invests time and effort, they can achieve native-like fluency. And our goal as instructors is to help them do that by designing materials that consider the mistakes and deviations highlighted in the experiment.

6. Conclusion

The article sheds light on issues that speakers of Russian face when they use the future tense, especially the challenges connected with aspect. Both native and non-native speakers of Russian have difficulties in matching verbal aspect when a stimulus is given in the past tense and the response should be future. Although this is an experimental task, it is not entirely unnatural, given that predictions of future events are commonly made based upon past observations.

The logistic regression analysis has shown that native and non-native speakers of Russian rely on different strategies when dealing with the future tense. Relative

frequency plays an important role for the native speakers. Native speakers tend to choose the aspect of higher frequency. Non-native speakers are not sensitive to the frequency of the verb in the same way as native speakers, and it is possible that in this sense non-native speakers may never achieve truly native-like fluency even at levels of high proficiency. The deviation patterns exhibited by non-native speakers suggest that they would rather avoid conjugating perfective verbs, opting instead to use the future auxiliary (*budu*) even with perfective verbs where it is ungrammatical. If the non-native speakers are not sure of the answer, it is important for them to note the presence of the future tense indicator in the most "obvious" way. By contrast, the native speakers have a stronger preference for the perfective verb form even if this choice prompts them to deviate from the task of matching the aspect.

References

Kosheleva, Daria & Janda, Laura. 2022. Looking into the Russian future. *CogniTextes. Revue de l'Association française de linguistique cognitive* https://doi.org/10.4000/cognitextes.2087

——— Under revision. Why markedness is always local.

Bybee, Joan, & Hopper, Paul (Eds.). 2001. *Frequency and the emergence of linguistic structure* (Vol. 45). John Benjamins Publishing.

Bybee, Joan. 2003. "19 Mechanisms of Change in Grammaticization: The Role of Frequency." *The handbook of historical linguistics*, 602.

———. 2006. "From usage to grammar: The mind's response to repetition." *Language*, 711-733.

Cejtlin, Stella. 2000: Цейтлин, Стелла. *Язык и ребенок: Лингвистика детской речи.* Москва: Владос.

Dickey, Stephen. Manuscript. *Tense and aspect.*

Gries, Stefan. 2021. *Statistics for Linguistics with R*. De Gruyter Mouton.

Gvozdev, A. 1961: Гвоздев, Александр. *Вопросы изучения детской речи.* АПН РСФСР.

Janda, Laura, Endresen, Anna & Reynolds, Robert. 2019. "Aspectual opposition and rivalry in Russian are not discrete: new evidence from experimental data." *Russian Linguistics: International Journal for the Study of the Russian Language*: 43 (3), 249-271.

Janda, Laura, & Reynolds, Robert. 2019. "Construal vs. redundancy: Russian aspect in context." *Cognitive Linguistics*, *30*(3), 467-497.

Labov, William. 1972. *Sociolinguistic patterns.* University of Pennsylvania Press.

Levine, James. 2009. *Russian grammar*. McGraw-Hill.

MacIntyre, Peter, Noels, Kimberly, & Clément, Richard. 1997. "Biases in self-ratings of second language proficiency: The role of language anxiety." *Language learning,* 47(2), 265-287.

Norwegian National Ethics Committee. 2016. *Guidelines for research ethics in the social sciences, humanities, law, and theology.*

Olshevskaya, Maria. 2018: Ольшевская, Мария. "Использование Русского учебного корпуса в преподавании РКИ: вид глагола". *Международный аспирантский вестник. Русский язык за рубежом, (1), 13-18.*

Polinsky, Maria. 2006. "Acquisition of Russian: Uninterrupted and incomplete scenarios." *Glossos*, 8, 1-64.

Stojnova, Natalia. 2019: Стойнова, Наталья. "Будущее несовершенного вида и инфинитивное сочетание с глаголом стать как конкурирующие конструкции в современном русском языке." *Русский язык в научном освещении* 1: 58-83.

Swan, Oscar E. 2017. "Sketching the Russian tense-aspect system for verbs that form simple aspect pairs." *Slavic and East European journal* 61.4: 825-843.

Szpunar, Karl, Spreng, Nathan, & Schacter, Daniel. 2014. "A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition." *Proceedings of the National Academy of Sciences*, 111(52), 18414-18421.

Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge University Press.

Tulving, Endel. 2002. "Episodic memory: From mind to brain." *Annual review of psychology*, 53(1), 1-25.

Turian, Donna, & Altenberg, Evelyn. 1991. "Compensatory strategies of child first language attrition." In Seliger, Herbert & Vago, Robert (eds.) *First Language Attrition*. Cambridge: Cambridge University Press, 207-226.

Wade, Terence. 2020. *A comprehensive Russian grammar*. John Wiley & Son