

Lidia Federica Mazzitelli

## Belarusian Modals of Necessity. A Corpus-Based Analysis

### 1. Introduction

The aim of my contribution is to offer a quantitative, corpus-based description of the morpho-syntactic and semantic properties of modals of necessity – which I define, following Besters-Dilger *et al.* (2009), as polyfunctional expressions of modality – in contemporary standard written Belarusian.

Existing studies of Belarusian modals are not numerous (see § 3. below), and none of them has a corpus-based perspective. As a result, we do not know much about their discourse properties: their distribution across different textual genres, their preference for the expression of one particular type of modality, and their overall frequency of use. The aim of the present study is to shed some light on this topic by analyzing the *Belacorpus* (Mazzitelli 2021), a corpus of contemporary written standard Belarusian texts containing ca. 1,5 million words. In particular, I will tackle the following questions: what are the morphosyntactic properties of Belarusian modals, and what types of modality do they express (dynamic, deontic, epistemic)? What is their frequency of use of in the corpus, and across different textual genres? Is the use of some modals limited to given types of discourse? As I show in the paper, Belarusian modals indeed do show differences as to the type of modality they can express and as to the textual types in which they preferably occur.

The paper is organized as follows: in § 1. I introduce the Belarusian language and the *Belacorpus*, the corpus used for this study. In § 2. I discuss the terminological definition of 'modals', and in § 3. I briefly review previous studies concerning Belarusian modals of necessity. In § 4. I present and discuss the results of the corpus analysis, and in § 5. I draw some conclusions and present future research desiderata.

#### 1.1. Standard Belarusian

(Standard) spoken Belarusian is, nowadays, one of the Slavic languages facing serious endangerment. According to the data collected by Hentschel, Kittel (2011), Belarusian is spoken on a daily basis by ca. 5% of the population, with the vast majority of Belarusians using rather Russian (intermixed to various extents with Belarusian elements) and the so-called Belarusian-Russian mixed language, or *trasjanka*. Its use is largely confined to the symbolic function; in the past (and still today to various extents) Belarusian has been actively ousted out of the public sphere by explicitly pro-Russian state policies (Brüggemann 2010; Hentschel *et al.* 2015; Zeller, Levikin 2016).

The quite unstable status of spoken Belarusian is mirrored in its peculiar situation as a language with two written standards: the so-called *taraškevica* (originally codified by Branislaŭ Taraškevič in 1918, nowadays only used by Belarusian-speaking media outlets based abroad and in private correspondence in Belarus, where its use is not allowed in the public sphere) and the so-called *narkamaŭka* (state-sanctioned, heavily russified, codified first in 1933 and then again in 2008). The two standard languages show differences in both lexicon and morphosyntax, see Bazhutkina (2020). In this paper, I analyze texts in both standards: there are no significant differences in the behaviour of modals between *taraškevica* and *narkamaŭka*, with the notable exception of the modal uses of *meč* 'have', which are almost only found in texts written in *taraškevica* (§ 4.2 below).

### 1.2. The Corpus

The data for this study are taken from the *Belacorpus*, a corpus of contemporary written Belarusian texts I compiled in 2010 (Mazzitelli 2021). The corpus contains ca. 1,5 million words from texts of various genres (fiction, non-fiction, specialized press, popular press, legislative texts), all produced in the period 1987 to 2010 (TABLE 1). The corpus is available online (<[https://github.com/Belarusian-Corpus/belacorpus\\_public](https://github.com/Belarusian-Corpus/belacorpus_public)>), upon request. At the moment, there is no query interface: the raw texts can be downloaded in .txt format and searched using a software such as AntConc (Anthony 2022). The corpus texts are both in *narkamaŭka* and *taraškevica*; in the remainder of this paper, when quoting examples, I follow the original orthography.

TABLE 1.  
Composition of the *Belacorpus*

Section	Words	%	Texts
National newspapers	643.939	44,3	126
Fiction (novels, short stories)	361.841	24,9	33
Legislative texts	121.283	8,3	22
Non-fiction (philosophy, sociolinguistics)	113.601	7,8	43
Non-fiction (books: history)	106.448	7,3	4
Magazines	67.045	4,6	11
Memoires	25.078	1,7	1
Local newspapers	15.283	1,1	7
<b>Total</b>	<b>1.454.518</b>	<b>100</b>	<b>247</b>

The *Belacorpus* is not the only electronic corpus of standard Belarusian available to researchers, the other two being the *Corpus Albaruthenicum* and the *Parallel'nyj korpus russko-belorusskij*. The *Corpus Albaruthenicum* is hosted by the Belarusian National Technical University (<<http://grid.bntu.by/corpus>>), and comprises of 350,027 words from 74 texts; it only covers one textual genre, namely scientific texts. The *Parallel'nyj korpus russko-belorusskij* is part of the Russian National Corpus (<<https://ruscorpora.ru/>>); it comprises of 2, 839, 268 words, with both original Belarusian texts and translations from Russian into Belarusian. The original Belarusian texts are 89, for a total of 1,920,705 words; the texts are produced in a timeframe spanning from 1862 to 2011, and belong to two genres, fiction and journalistic prose. Even though both these corpora are excellent resources, they lack diversity in terms of represented textual genres; moreover, the *Parallel'nyj korpus* contains texts from older stages of the Belarusian language, which I do not consider in this investigation. I decided therefore to use exclusively the *Belacorpus* as a source of data.

## 2. Modals in Typological Perspective

Following Besters-Dilger *et al.* (2009), I understand modals as a prototype-based category, which admits core (prototypical) and non-core (peripheral) members. Core modals are defined as follows:

A modal is a polyfunctional expression of modality. It always occurs with main verbs in the predicate position and opens one and only one argument position, which is filled by a lexical verb stem. A modal does not select its own nominal argument but influences the encoding of the arguments of the verbal form (Besters-Dilger *et al.* 2009: 169).

The very first characterizing property of modals is to be expressions of modality. Following Bybee *et al.* (1994), I consider modality as a semantic domain encompassing the notions of necessity, possibility and volition (but, contrarily to Bybee *et al.*, I do not consider evidentiality part of the domain of modality; see the discussion in Aikhenvald 2004: 7ff); in this paper, I only focus on the notion of necessity. Three types of modality can be distinguished: dynamic, deontic, and epistemic (van der Auwera *et al.* 1998: 80ff). Dynamic and deontic modality concern the causes of the participant's obligation to be engaged in the event denoted by the predicate. Expressions of dynamic modality concern events which generate an external (1a) or internal need (1b) of the participant; expressions of deontic modality denote an obligation, dictated by external sources of authority (2). Epistemic modality "refers to a judgment of the speaker: a proposition is judged to be uncertain or probable relative to some judgment" (*ibidem*: 81), see (3).

- (1) a. *To get to the station, you **have to** take bus 66.*  
 b. *Boris **needs** to sleep ten hours every night for him to function* (van der Auwera *et al.* 1998: 80).
- (2) *Citizens **must** pay taxes.*
- (3) *John **must** have arrived.* (van der Auwera *et al.* 1998: 81)

According to the definition given above, core modals are polyfunctional: that is, they can express at least two of the three types of modality – dynamic, deontic, or epistemic. English *must*, as examples (2) and (3) show, fulfil this requirement, in that it can express deontic and epistemic modality.

A further semantic distinction (which does not bear any consequences as to the classification of element as a modal or not) is the one concerning the degree of intensity of the expressed obligation, which may be weak or strong. In German, the modals *sollen* and *müssen* are usually distinguished in terms of intensity: *sollen* expresses weak obligation and *müssen* strong obligation.

From the syntactic point of view, modals behave like auxiliaries: their only open argument position must be filled by a predicate. In (4), the Russian modal *dolžen* ‘must’ governs the lexical stem *posmotret* ‘watch’:

- (4) *Ėto klassik, kotoruju chotja by raz v žizni dolžen posmotret’ každyj* (NKRJa).  
 ‘This is a classic, which everyone should watch at least once in their lifetime’.

Modals themselves do not select any nominal arguments; however, they influence the encoding of the nominal arguments of the verb. In examples (1) through (4), the highest-ranking (that is, the most subject-like; see Wiemer, Bjarnadottir 2014) argument of the predicate is in the nominative case. In (5), again an example from Russian, conversely, it is in the dative case.

- (5) *Ivanu.DAT nado bylo rabotat’* (Besters-Dilger *et al.* 2009: 175).  
 ‘Ivan had to work’.

Structurally, Slavic modals do not form a homogeneous class. Three main types (which can co-exist in one language) may be distinguished, depending on the word class modals have derived from (Besters-Dilger *et al.* 2009): de-verbal modals, such as Slovenian *morati* ‘must’ and Polish *mieć* ‘have (to)’; de-adjectival modals, such as Russian *dolžen* and Ukrainian *povynen* ‘must’; and de-adverbial modals, such as Russian *nado* and Belarusian *trëba* ‘be necessary’. For a detailed review of the morphosyntax of Slavic modals, I refer the reader to Besters-Dilger *et al.* (2009).

### 3. *Modals of Necessity in Belarusian: Overview*

#### 3.1. *Previous Research*

So far, only a small number of publications dedicated to the investigation of the morphosyntactic means of expression of modality (thus including modals) in Belarusian have appeared: among others, Pitsch (2010), an analysis of the syntactic behaviour of Belarusian deverbal and deadverbial modals from a formal point of view; Palásti (2009), an investigation of the Russian, Ukrainian and also Belarusian impersonal modal constructions (thus

excluding de-verbal modals such as *music* ‘must’, which selects the nominative case of the highest-ranking argument of the main predicate); and Besters-Dilger *et al.* (2009), who, in their overview of Slavic modals, list only two core modals of necessity for Belarusian: *music* and *pavinen*.

The only book-length study of the linguistic expression of modality (in the broadest possible sense) in Belarusian is Paŭloŭskaja (2001), who discusses what she labels as *madal’nyja adzinki* ‘modal units’, that is, all linguistic means – lexical, morphological or syntactic – that are used to code modal meanings. However, she does not provide any data about their frequency of use, and discusses the semantic differences between competing forms only very briefly. Furthermore, Paŭloŭskaja does not limit her investigation to modals, but she also discusses the use of correlated linguistic elements such as non-indicative moods (imperative, conditional) and conditional-requiring complementizers (*jak by* ‘as if’, *kab* ‘in order that’, *niby* ‘as though’), as well as lexicalized expressions such as *mec’ mahčymasc’* ‘to have the possibility’. In fact, the concept of ‘modal’ (in the sense described above) as a linguistic category is absent in her analysis: among the syntactic means to express modality, she lists *prėdykatyvy* ‘predicatives’ (de-adverbial modals), *prėdykatyŭnyja prymetniki* ‘predicative adjectives’ and *madal’nyja dzejaslovy* ‘modal verbs’, without recognizing them as members of the same category, that of modals.

Among the expressions of modality discussed by Paŭloŭskaja, those which comply with the semantic and morphosyntactic requirements for inclusion in the category of modals as defined above are *varta* ‘(be) worth; (be) necessary’, *neabchodna* ‘inevitable; necessary’, *trėba, patrėbna* ‘(be) necessary’, *pavinen, abavjazany* ‘obliged’, *mec’* ‘have’, *music* ‘be under obligation; be compelled to’<sup>1</sup>. These will be the focus of the following sections.

### 3.2. Morphosyntactic Characteristics

The modals listed by Paŭloŭskaja (2001) are instances of different construction types, according to the typology suggested in Besters-Dilger *et al.* (2009). The de-adjectival modals *pavinen* and *abavjazany* and the deverbal modals *mec’* and *music* belong to Besters-Dilger *et al.* (2009) construction Type 1: they select the nominative case for the highest-ranking argument of the infinitival verb, see examples (6a) and (6b).

- (6) a. *Hėta zasługa medrabotnikaŭ, jakija svaimi silami zrabili bal’nicu takoj, jakoj jana pavinna byc’.*

‘This is a merit of the medical staff, who with their own forces have made the hospital such as it must be.’

<sup>1</sup> Paŭloŭskaja (2001: 109ff) lists also a number of other verbs, which have non-modal primary meanings but which can be used to express necessitive modality: *nakanavac’* ‘(pre)designate’, *naležyc’* ‘belong’, *nadarycca* ‘happen’, *davadzicca* ‘happen’, *zastavacca* ‘remain’, *pryjscisja* ‘happen. I do not take them into consideration here.

- b. *Tut, maŭljaŭ, chočaš nja chočaš, a ŭs'michacca musiš!*

'Here, please, want it or not, you must smile!'

*Pavinen* and *abavjazany* can also be used adverbially, as in example (7)<sup>2</sup>. In this case, they represent an instance of the construction Type 5 in Besters-Dilger *et al.* (2009), admitting only zero or genitive-coded privileged arguments.

- (7) *Palityka budue časam bar"ery tam, dze ich ne pavinna byc'.*

'Politics sometimes builds barriers where there shouldn't be any.'

The remaining modals – *varta*, *nebachodna*, *nepatrěbna*, and *trěba* – are representatives of the construction Type 3: the highest ranking argument of the dependent verb is either coded as a dative (8a), or as zero (8b), usually with a non-generic reference.

- (8) a. *Ne, ja stamiŭsja. Mne trěba adpačyc'.*

'No, I am tired. I need to rest.'

- b. *Našy bac'ki trykali haspadarku i my trymaem. My pryzvyčalisja, što trěba pracavac'. I ne možam bez pracy.*

'Our parents kept a farm, and we are keeping it, too. We got used to the fact that we need to work. And we cannot do without working.'

#### 4. Modals in the Belacorpus

##### 4.1. Distribution

The absolute number of occurrences of Belarusian modals of necessity in the *Belacorpus* is shown in TABLE 2; as the table shows, the most used modals are by far *trěba* and *pavinen*, closely followed by *music'*.

It is worth remarking that, of all these forms, only *pavinen* and *music'* always behave as auxiliaries. The other ones – *abavjazany* 'obligated; owe', *mec'* 'have', *nepatrěbna* 'unnecessarily', *patrěbna* 'necessarily', and *nebachodna* 'unavoidably' can be used in other constructions, with a non-modal semantics; see (9), where *abavjazany* is used in its sense of 'owe; be indebted (to someone for something)', and governs two pronominal expressions (*vy* '2PL' and *usě* 'everything').

- (9) *Pavažanyja vetěranŭ! Vam my abavjazanyja ŭsim – žyccěm, svabodaj, ščasćem našych dzjacej.*

'Honoured veterans! We owe you everything – our life, freedom, the happiness of our children.'

<sup>2</sup> Because of the orthographical rendition of the *akan'e*, the neuter singular ending of adjectives and adverbials and the feminine singular nominative ending of adjectives are both <-a>: *pavinn-a* 'must-NOM.F.SG/-N.SG'.

TABLE 2.  
Modals of necessity in the *Belacorpus*

Form	Occurrences
<i>nepatrëbna</i>	1
<i>patrëbna</i>	26
<i>mec'</i>	77
<i>abavjazany</i> (including feminine and plural forms)	81
<i>neabchodna</i>	118
<i>varta</i>	350
<i>music'</i> (including all present and past forms)	573
<i>pavinen</i> (including feminine and plural forms)	880
<i>trëba</i>	estim. 1381

Only occurrences where these forms behave syntactically like modals (that is, governing an infinitive) are included in the table. Because the *Belacorpus* is not tagged for morphosyntactic categories, automatic disambiguation of homonymy is not possible: I manually checked all occurrences of the searched lexemes. However, I made an exception for *trëba*, whose total number of occurrences in the corpus is very high (1749). I decided to analyze only 200 of them, out of which 158 (79%) are modal. Projecting this percentage to the total number of occurrences of *trëba* in the corpus, the result is a likely number of 1381 occurrences, where *trëba* is used as modal.

#### 4.2. Types of Expressed Modality

In order to determine the types of modality expressed by the modals listed in TABLE 2 above, I manually analyzed all occurrences of *(ne)patrëbna*, *mec'*, *abavjazany*, *neabchodna*, and *varta*; because of the high number of occurrences of *music'*, *pavinen*, and *trëba*, I analyzed a randomly selected subset of 200 occurrences per each one of them.

The corpus analysis confirms what already observed in Besters-Dilger *et al.* (2009): as TABLE 3 shows, only *music'* and *pavinen* can express three types of modality, thus fulfilling the polyfunctionality criterion and qualifying as core members of the category of modals. Other modals, conversely, are specialized in the expression of only one type of modality, and are thus peripheral. In terms of intensity *abavjazany*, *music'* and *pavinen* alike express strong obligation; weak obligation is expressed by *mec'* 'have'.

In the following sections, I examine more in detail the two groups of modals, core and peripheral.



TABLE 3.  
Types of expressed modality

	Types of modality
<i>nepatrěbna</i>	dynamic
<i>patrěbna</i>	dynamic
<i>neabchodna</i>	dynamic
<i>varta</i>	dynamic
<i>trěba</i>	dynamic
<i>mec'</i>	deontic (weak obligation)
<i>abavjazany</i>	deontic (strong obligation)
<i>music'</i>	dynamic, deontic (strong obligation), epistemic
<i>pavinen</i>	dynamic, deontic (strong obligation), epistemic

#### 4.3. Core Modals: *music'* and *pavinen*

*Music'* is a de-verbal modal: as Hansen (2000: 77) observes, *music'* is ultimately a borrowing from German *müssen*, through Polish *musieć*. As mentioned above, *music'* expresses dynamic (10), deontic (11) and epistemic (12) modality.

- (10) *Nohi 'harac', i tady ustaeš, tamu ja ne mahu naskroz' caluju noč, mušu ũstavac' nekal'ki razoŭ.*  
'[Because of osteocondrosis] Your legs 'burn' and then you get up, so I can't sleep through the night, I **must get up** many times.'
- (11) *Zhodna z damovaj kožny bok music' harantavac' dostup rybavaleckim sudam inšaha boku.*  
'According to the agreement, each party **must guarantee** the access to the fishing vessels of the other party.'
- (12) *Papa Urban VIII musiŭ mec' tonki mastacki hust kab zaiŭvažyc' u Bernini dyj Baramini henial'nych architektaraŭ.*  
'Pope Urban VIII **must have had** a fine artistic taste to recognize in Bernini and Borromini two brilliant architects.'

The epistemic use of *music'* is extremely rare: only two out of the 200 analyzed occurrences of *music'* have an epistemic meaning. However, two lexicalized expressions based on *music'* exist, namely *music'*'must.PRS.3SG' and *music'byc'* (< *music'byc'* 'must be'), which are used to mean 'probably; highly likely', see Hansen (2000: 83) and examples (13) and (14).

- (13) *Pašli da maich – šapnula Anatoliju. A to, music', užo spac' lehli.*  
'Let's go to my parents' place – she whispered to Anatol'. They **must have gone** to bed already.'



- (14) *Oj, music', ja pamiraju? Haru ūsja...*  
 'Oh, **must be**, I am dying? I'm all burning...'

As shown in **TABLE 2** above, *pavinen* is slightly more frequent than *music'*. It can also express dynamic (15), deontic (16) and epistemic (17) modality.

- (15) *Kol'ki jaščė pabromaŭ pavinna adbycca na prastorach Rasei kab hramadztva adčula, što fašyzm ūžo nja tol'ki pabroza?*  
 'How many more pogroms **must take place** in Russia before people feel that fascism is no longer merely a threat?'
- (16) *Na pracjahu 15 dzėn sud pavinen abo vyznačycca z dataj razboru, abo nakiravac' spravu na das'ledavan'ne.*  
 'Within 15 days, the court **must either decide** on a date for the trial, or **send** the case to be investigated.'
- (17) *U lese ljažau sneh, i jany doŭha išli da paljany, dze, jak scvjardžaŭ navučenc, snehu ne pavinna byc'.*  
 'In the forest there was snow, and they walked for a long time to the clearing, where, like the student claimed, there **should be** no snow.'

As it was the case with *music'*, the use of *pavinen* as an epistemic marker is quite rare, too: out of 200 occurrences, only one has an epistemic meaning. Moreover, *pavinen* is not used to express inferential epistemicity: it is always used in reported speech clauses, where the narrator (or the speaker) does not take responsibility for the veridicality of what they report; see also (18), another example I found in the corpus (not included in the analyzed 200 occurrences).

- (18) *Vy ž čuli padanne pra knjazja Vitaŭta i Darynu? Dyk tut nedze pavinen znachodzicca jae maėntak, tam, dze rasce dzikaja mal'va, u lese.*  
 'Have you heard the legend of Prince Vytautas and Daryna? Actually, her estate **should be** here somewhere, there, where the wild mallow grows, in the forest.'

There seem to be no difference in their distribution across different textual genres. Both *music'* and *pavinen* are widely used in all kinds of texts, with one exception: neither of them is ever used to express deontic modality in legislative texts, where *abavjazany* is used instead (see § 4.4).

#### 4.4. *Peripheral Modals of Necessity*: (ne)patrėbna, neabchodna, varta, trėba, and abavjazany

As mentioned in § 4.2, most Belarusian modals do not qualify as core members of the category, because they are specialized in the expression of only one type of modality. (Ne)patrėbna, neabchodna, varta, and trėba encode exclusively dynamic modality (19; 20; 21; 22), and abavjazany exclusively deontic modality (23).

- (19) *Kožnaja kapusta ŭ svoj čas smačnaja. Tamu patrěbna brac' i rannjuju, kab u lipeni byla, i poznjuju, kab u žniŭni jaščė esci možna bylo.*  
 'Every cabbage is tasty at its time. Therefore, it is necessary to take both the early one, so that you will have it in July, and the late one, so that you can still eat it in August.'
- (20) *Dzelja vyratavannja belaruskaj movy, kul'tury ab'ektyŭna neabchodna jak maha chutčėj razburyc' pabudavany silaj nad usej Belarussju rasejskamoŭny pancyr.*  
 'In order to save the Belarusian language, culture, it is objectively necessary to destroy the Russian-speaking shell built by force over the whole of Belarus as soon as possible.'
- (21) *Ėŭrope varta bylo b na chvilinky spynicca i zadumacca, ci ne njase jana ŭ hėtaj sytuacyi chacaj b častku adkaznašci za pabrozu ŭlasnaj ěnerhetyčnaj bjaspecy.*  
 'Europe should stop for a moment and think whether it does not bear at least part of the responsibility for the threat to its own energy security in this situation.'
- (22) *Ne, ja stamiŭsja. Mne trěba adpačyc'.*  
 'No, I am tired. I need to rest.'
- (23) *Artykul 52. Kožny, chto znachodzicca na terytoryi Rėspubliki Belarus', abavjazany vykonvac' jae Kanstytucyju, zakony i pavažac' nacyjanal'nyja tradycyi.*  
 '[Constitution of the Republic of Belarus] Article 52. Everyone who find themselves on the territory of the Republic of Belarus is obliged to comply with its Constitution, its laws and respect its national traditions.'

In terms of distribution, (*ne*)*patrěbna*, *varta* and *trěba* are used in all kinds of textual genres. *Neabchodna*, conversely, is mostly limited to journalistic and scientific prose; similarly, *abavjazany* is almost exclusively found in legislative texts: deontic modality is expressed in non-formal texts by either *music'* or *pavinen*.

#### 4.5. *Modal mec'*

The de-verbal *mec'* 'have' deserves a more detailed discussion. A comprehensive investigation of the morphosyntax and semantics of *mec'* in contemporary Belarusian is beyond the scope of the present study: I refer the reader to two previous publications (Mazzitelli 2011; 2015: 176-183), of which I present here the major points of interest.

In contemporary standard Belarusian, *mec'* is primarily used to express possession, as exemplified in (24).

- (24) *Ja maju svaju kavtėru.*  
 'I have my own apartment.'

*Mec'* only plays a secondary role in the expression of modality in Belarusian. First of all, its use as a modal is quite rare (only 77 occurrences in the *Belacorpus*), and it has a severely restricted paradigm: most occurrences are in the third person, and in the past tense;

its use in future tense is not attested in the corpus at all. Semantically, *mec'* is primarily used to express post-modal meanings, namely scheduled future (25), counterfactual future in the past (26), purpose (27) (*mec'* describes the goal of an artefact or of an event, according to the intentions of their creators or organisers), and 'fatalistic future' (28).

- (25) *23 červenja ŭ Homeli mae adbycca koncert Za Belarus'.*  
 'On June 23<sup>rd</sup> in Homel' the concert *Za Belarus'* is **expected to take place**/is scheduled'.
- (26) *Impréza mela adbvcca 13 ljutaha, ale administracyja ŭstanovy ŭ seradu admovila ŭ jaho pravjadyen'ni pa 'téchničnych pryčynach'.*  
 'The party **should have taken place** on February 13th, but on Wednesday the administration refused permission for 'technical reasons'.
- (27) *Na dumku Gejtsa [...] Vista mae spras'cic' zachoŭvan'ne mnoštva fatazdymkaŭ i fil'maŭ.*  
 'According to Gates [...] Vista is **supposed/expected to simplify** the process of saving many photos and movies'.
- (28) *Hětyja peš'ni zaiŭždy buduc' aktual'nyja i zapatrěbavanyja, a sam dysk mae vytrymac' ne adno peravydan'ne.*  
 'These songs will always be modern and popular, and this album is **destined to be released** some more times'.

*Mec'* is relatively rarely used express deontic obligation (in 17 occurrences out of 77), and it always involves an element of 'reference to an external source' (typical of its semantics): the action indicated by the infinitive governed by *mec'* has to be accomplished because a source, mentioned in the text or implied (a person or, much more often, a law or a regulation) wants or requires it. The use of *mec'* often implies that the law (or the requirements of another source of authority) are likely not to be complied with: in (29), to the law (the source), legal proceedings should be conducted in Belarusian or in Russian, but they might not (i.e. usually, they are conducted only in Russian, and the Belarusian option is never considered). Such counterfactual interpretation would be absent with either *music'* or *pavinen*.

- (29) *Zhodna z art. 14, sudavodstva ŭ Respublicy Belarus' mae ves'cisja 'na belaruskaj abo rasijskaj move'.*  
 'According to art. 14 [of the Constitution], legal proceedings in the Republic of Belarus **should be conducted** 'in Belarusian or Russian'.

The use of *mec'* is decidedly higher in the *taraškevica* standard than in the *narkamaŭka* one. Out of the 77 occurrences of modal *mec'*, only 28 are found in texts written in *narkamaŭka*. As I observed in a previous publication, "the example of the newspaper *Naša Niva* provides striking evidence of that: The newspaper was edited in *taraškevica* until 2008, when it switched to *narkamaŭka*. Before [2008], the constructions with [modal] *mec'* are relatively very frequent (51); after the switch, only two occurrences of [it] were found". (Mazzitelli 2011: 181).

## 5. Conclusions

This study has confirmed the statement identification by Besters-Dilger *et al.* (2009) of *music'* and *pavinen* as the only two core Belarusian modals of necessity: only them fulfil the polyfunctionality criterion, being able to express more than one type of modality. All other expressions of modality, identified in previous studies such as Paŭloŭskaja (2001), are to be classified as peripheral modals, because their semantic range is limited to only one type of modality.

The data from the *Belacorpus*, however, do not only confirm the validity of Besters-Dilger *et al.* (2009) classification, but also shed light on previously unnoticed semantic and discourse properties of Belarusian modals. First of all, the corpus analysis shows that *music'* and *pavinen* are largely limited to the expression of dynamic and deontic modality; epistemic uses are attested, but are extremely rare. Moreover, the corpus study has individuated a semantic difference between the two: *music'* is used to express inferential certainty as well as suppositions, whereas *pavinen* is only used to express the latter.

Furthermore, my study has shown that peripheral modals are unevenly distributed across textual genres. *Neabchodna* is typically used in a restricted and well-defined number of textual genres, namely journalistic and non-fiction prose. Similarly, the use of *abavjazany*, a marker of deontic modality, is limited to legal texts, where it is predominant; in all other types of texts, either *music'* or *pavinen* are used. (Interestingly, the Belarusian situation differs dramatically from the one found in Ukrainian, where the most common expressions of deontic modality in legislative texts are *maty* 'have' and *povinen* 'must', clear cognates of Belarusian *mec'* and *pavinen*; Goletiani 2015). As for the distribution of modals in the two standards in which Belarusian is codified, *taraškevica* and *narkamaŭka*, my analysis shows that there is no noticeable difference between the two, with the notable exception of modal *mec'*, which is predominantly used in the *taraškevica* standard.

Even though admittedly brief, my analysis has shown that only corpus-based quantitative investigations can unearth patterns of use and pragmatic-semantic properties which would otherwise remain hidden; more such studies are sorely needed for Belarusian, a language which has only rarely been the target of corpus-based investigations. An urgent research *desideratum* would be to have a corpus of spoken Belarusian, following the example of the Oldenburg corpus of *trasjanka* (Hentschel *et al.* 2014), which would allow to extend the research to the spoken language, too.

## Literature

## A. CORPORA AND SOFTWARES

- Anthony 2022: L. Anthony, *AntConc (Version 4.1.2)* [Computer Software]. Tokyo 2022; <<https://www.laurenceanthony.net/software>> (accessed on: 19.09.2022).
- Corpus Albaruthenicum*: <<http://grid.bntu.by/corpus>> (accessed on: 19.09.2022).
- Mazzitelli 2021: L.F. Mazzitelli, *Belacorpora. A corpus of written Standard Belarusian texts*; <<https://github.com/Belarusian-Corpus>> (accessed on: 19.09.2022).
- NKRJa: *Nacional'nyj korpus russkogo jazyka*. 2003-2022; <[ruscorporu.ru](http://ruscorporu.ru)> (accessed on: 19.09.2022).

## B. STUDIES

- Aikhenvald 2004: A. Aikhenvald, *Evidentiality*, Oxford 2004.
- Bazhutkina 2020: A. Bazhutkina, *Belarussische Standardsprache(n) im Diskurs*. München 2020.
- Besters-Dilger et al. 2009: J. Besters-Dilger, A. Drobnjaković, B. Hansen, *Modals in the Slavonic Languages*, in: B. Hansen, F. de Haan (eds.), *Modals in the Languages of Europe. A Reference Work*, Berlin 2009, pp. 167-198.
- Brüggemann 2010: M. Brüggemann, *Koloniales sprachliches Erbe: Sprache und Nation in Lukašenkas Belarus*, "OSTEUROPA", 2010, 12, pp. 69-80.
- Bybee et al. 1994: J. Bybee, R. Perkins, W. Pagliuca, *The Evolution of Grammar. Tense, aspect, and modality in the languages of the world*, Chicago 1994.
- Goletiani 2015: L. Goletiani, *Zur Übersetzung deontischer Modalmarker ins Ukrainische: eine korpusgestützte Untersuchung anhand von EU-Richtlinien*. "Zeitschrift für Slawistik", LX, 2015, 2, pp. 269-293.
- Hansen 2000: B. Hansen, *The German Modal 'müssen' and the Slavonic Languages – Reconstruction of a success story*, "Scando Slavica", XLVI, 2000, pp. 77-93.
- Hansen 2001: B. Hansen, *Das slavische Modalauxiliar – Semantik und Grammatikalisierung im Russischen, Polnischen, Serbischen/Kroatischen und Altkirchenslavischen*, München 2001.
- Hansen 2003: B. Hansen, *The Grammaticalization (and Degrammaticalization?) of Modals in Slavonic*, in: T. Berger, K. Gutschmidt (eds.) *Funktionale Beschreibung slavischer Sprachen. Beiträge zum 13. Internationalen Slavistenkongress Ljubljana 2003*, München 2003, pp. 97-115.

- Hentschel, Kittel 2011: G. Hentschel, B. Kittel. *Weissrussische Dreisprachigkeit? Zur sprachlichen Situation in Weissrussland auf der Basis von Urteilen von Weissrussen über die Verbreitung "ihrer Sprachen" im Lande*, "Wiener Slavistischer Almanach", LXVII, 2011, pp. 107-135.
- Hentschel *et al.* 2014: G. Hentschel, J. Zeller, S. Tesch, *Das Oldenburger Korpus zur weißrussisch-russischen gemischten Rede: OK-WRGR*. Oldenburg 2014, <<http://diglib.bis.uni-oldenburg.de/bis-verlag/ok-wrgr/>>.
- Hentschel *et al.* 2015: G. Hentschel, Georg, J. Zeller, H. Geiger, M. Brüggemann, *The Linguistic and Political Orientation of Young Belarusian Adults Between East and West or Russian and Belarusian*, "International Journal of the Sociology of Language", 2015, 236, pp. 133-154.
- Mazzitelli 2011: L. F. Mazzitelli, *Possession, Modality and Beyond: The case of mec' and mecca in Belarusian*, in: M. Nomachi (ed.), *Grammaticalization in Slavic Languages: From Areal and Typological Perspectives (Revised and Enlarged Edition)*. Sapporo 2011, pp. 179-202.
- Mazzitelli 2015: L. F. Mazzitelli, *The Expression of Predicative Possession. A comparative study of Belarusian and Lithuanian*, Berlin-Munich-Boston 2015.
- Palásti 2009: K. Palásti *Bezličnye predloženiya, vyražajuščie modal'nye značeniya, v russkom, belorusskom u ukrainskom jazykach*, "Studia Russica", XXIII, pp. 280-288.
- Paŭloŭskaja 2001: N.Ju. Paŭloŭskaja, *Katėhoryja madal'nasci ŭ sučasnaj belaruskaj move*. Minsk 2001.
- Pitch 2010: H. Pitch, *Ličnye modal'nye glagoly i modal'nye predikativy v belorusskom jazyke*, in: S.M. Zaprudski, H.A. Cychun. *Novae slova ŭ movaznaŭstve: matėryjaly V Mižnarodnaha kanhrėsa belaruskistaŭ, Minsk, 20-21 maja 2010 h.*, Minsk 2010, pp. 193-201.
- van der Auwera *et al.* 1998: J. Van der Auwera, V. A. Plungian. 1998. *Modality's semantic map*, "Linguistic Typology", 1998, 2, pp. 79-124.
- Wiemer, Bjarnadóttir 2014: B. Wiemer, V. Bjarnadóttir, *On the Non-Canonical Marking of the Highest-Ranking Argument in Lithuanian and Icelandic*, in: A. Holvoet, N. Nau, *Grammatical Relations and Their Non-Canonical Encoding in Baltic*, Amsterdam 2014, pp. 301-361.
- Zeller, Levikin 2016: J. Zeller, D. Levikin, *Die Muttersprachen junger Weißrussen. Ihr symbolischer Gehalt und ihr Zusammenhang mit sozialen Faktoren und dem Sprachgebrauch in der Familie*, "Wiener Slavistisches Jahrbuch" (Neue Folge), IV, 2016, pp. 114-144.

*Abstract*

Lidia Federica Mazzitelli

*Belarusian Modals of Necessity. A Corpus-Based Analysis*

The paper presents an analysis of standard Belarusian modals of necessity, based on the *Belacorpus*, a corpus of contemporary written Belarusian I built in 2010. I investigate the modals' semantics (which types of modality they express), their frequency in the corpus and their distribution across textual genres. My study confirms what was already observed in previous accounts of Belarusian modals. Namely, only two of them, *music'* and *pavinen*, are polyfunctional (that is, they can express more than one type of modality – dynamic, deontic and epistemic) and thus prototypical (core) members of the modal category. All other modal expressions are dedicated to one type of modality and are thus peripheral. The corpus analysis has also revealed a number of previously unnoticed properties of Belarusian modals, such as: the extreme rarity of the epistemic use of polyfunctional *music'* and *pavinen*; the uneven distribution of peripheral modals across different types of texts; and the fact that the two standards in which Belarusian is codified, *taraškevica* and *narkamaŭka*, present no significant difference as far as the use of modals is concerned, with the notable exception of modal *mec'* 'have', which is much more frequent in *taraškevica* texts.

*Keywords*

Modality; Belarusian; Corpus Studies.