MARIA OVSJANNIKOVA
University of Potsdam, maria.ovsjannikova@uni-potsdam.de
ORCID 0000-0002-8313-0374

WHAT IT MEANS TO FEEL IN SLAVIC:
AN EXPLORATION INTO THE SEMANTICS OF SLAVIC 'FEEL'-VERBS

Abstract

The study examines verbs with the meaning 'feel' across Slavic. Based on parallel corpus data from seven Slavic languages, I explore the semantic classes described by these verbs, lexical distinctions drawn in individual languages, and similarity relations between Slavic 'feel'-verbs. In a number of Slavic languages, the lexical distinction is drawn between verbs that are able, among other classes, to describe well-being and self-assessment, as opposed to verbs that typically describe situations featuring external stimuli, such as environment, touch, and smell.

Keywords: perception verbs, parallel corpus, stimulus, lexical typology

1. Introduction

Verbs with the meaning 'feel' can describe a wide range of experiential situations, including touch, smell, emotional states, and bodily sensations (Viberg 1983, pp. 139, 143), as shown in (1)-(2) by parallel examples from Slavic[1].

(1)BG *Като се обърнах,* ***усетих*** *до бедрото си джобен магнитофон.*

CR *U jednome trenutku, kada sam se pomaknuo,* ***osjetio*** *sam na bedru plosnati oblik magnetofona.*

CZ *Jednu chvíli, když jsem se pohnul,* ***ucítil*** *jsem plochý tvar magnetofonu, jak mě tlačí do boku.*

PL *W pewnej chwili, kiedy się poruszyłem,* ***poczułem*** *przylegający do biodra płaski kształt magnetofonu.*

RU *Пошевелившись, я* ***почувствовал*** *прижатую к бедру плоскую коробку магнитофона.*

UK *Поворухнувшись, я* ***відчув*** *притиснуту до стегна пласку коробку магнітофона.*

[1] All examples in the paper are taken from the InterCorp parallel corpus (Rosen et al., 2022; Rosen, 2023). The English context from the corpus is used as translation, unless indicated otherwise.

‘Shifting my position, I felt the flat shape of the tape-recorder against my hip.’

(2)BG *Докато пиша,* ***чувствам*** *същата умора, която* ***усещах*** *онази нощ или по-точно - онази сутрин.*

CR *Dok ovo pišem,* ***osjećam*** *se umornim kao što sam se* ***osjećao*** *one noći, to jest onoga jutra.*

CZ *Jak tu píši, přepadá mě únava, jakou jsem* ***cítil*** *tenkrát v noci, vlastně ráno.*

PL *Kiedy piszę te słowa,* ***czuję*** *się znużony jak owej nocy, a właściwie owego ranka.*

RU *Вот я пишу эти строки и снова* ***испытываю*** *мрачную, невыносимую усталость, которая была во мне той ночью, вернее тем утром.*

UK *Пишучи це, я* ***почуваюся*** *таким самим зомореним, як тієї ночі, себто того ранку.*

‘In setting down these words, I feel weary, as I felt that night – or, rather, that morning.’

Across languages, ‘feel’-verbs can differ to a greater or lesser extent in their semantic range. Moreover, the semantic range covered by just one verb in one language can be divided, with more or less considerable overlap, between several verbs in another language, such as *чувствам* and *усещам* in Bulgarian and *відчути* and *почуватися* in Ukrainian in examples (1)-(2). The investigation of the inventories of ‘feel’-verbs and their semantic ranges across Slavic is the main goal of this paper.

The study consisted of two parts, both of which were based on parallel corpus data from the InterCorp parallel corpus (Rosen et al., 2022; Rosen, 2023). In the first part, based on a sample of parallel contexts, I explore the semantic classes of situations expressed by ‘feel’-verbs in seven Slavic languages: Bulgarian, Croatian, Slovenian (South Slavic); Czech, Polish (West Slavic); Russian, Ukrainian (East Slavic). I identify lexical means which are used to express these situations and zoom in on three languages to examine lexical distinctions drawn in them and their associations with semantic classes.

In the second part of the study, I take a wider range of ‘feel’-verbs and languages into account. Using the data on the frequency of correspondences between these verbs in parallel texts, I explore the structure of the semantic domain covered by ‘feel’-verbs in Slavic and their groupings.

The paper is structured as follows. Section 2 presents the first part of the study. I start by describing data retrieval and annotation (2.1), then move on to the semantic classes I distinguished on the basis of the sample (2.2), discuss the associations between these semantic classes and aspect (2.3), and present an analysis of the major ‘feel’-verbs in Bulgarian, Czech, and Ukrainian. In section 3, I look at the frequencies of ‘feel’-verbs (3.1) and correspondences between them (3.2) in a larger corpus. Section 4 summarizes the main findings of the study.

2. Semantic classes and verbs

2.1. Data

In this part of the study, I analyse the semantic range of situations described by 'feel'-verbs in Slavic and examine whether there are associations between semantic classes and lexical distinctions in individual languages.

First, for each of the languages, I identified the major 'feel'-verbs using dictionaries and parallel texts. These verbs given in the central column of Table 1. These verbs were included in the queries which were used to create the sample for the study of semantic classes covered by Slavic 'feel'-verbs.

Table 1. 'Feel'-verbs included in the corpus queries and identified based on the sample

Language	Verbs included in the query[2]	'Feel'-lexemes identified in the sample
BG	*чувствам*, *почувствам*, *усещам*, *усетя*, *чувствувам*, *почувствувам*	*изпитвам*, *изпитам*, *обзема*, *чувство*, *усещане*, *имам чувство*, *имам усещане*
CR	*osjećati*, *osjetiti*	*osjećaj*, *obuzimati*, *obuzeti*, *ćutjeti*, *imati osjećaj*
CZ	*cítit*, *pocítit*, *ucítit*, *vycítit*, *pociťovat*	*mít pocit*, *připadat si*, *pocit*, *cit*, *zakoušet*, *zakusit*, *vyciťovat*
PL	*czuć*, *poczuć*, *wyczuć*, *odczuć*, *odczuwać*, *wyczuwać*	*uczucie*, *mieć uczucie*, *ogarniać*, *ogarnąć*, *poczucie*, *wyczuwanie*, *odczucie*
RU	*чувствовать*, *почувствовать*, *ощущать*, *ощутить*	*испытывать*, *испытать*, *чуять*, *почуять*, *чувство*, *ощущение*, *охватить*
UK	*відчувати*, *відчути*, *почувати*(*ся*)[3]	*почути*, *чути*, *відчуття*, *почуття*, *вчуватися*, *мати відчуття*, *охопити*
SL	*čutiti*, *občutiti*, *počutiti*, *začutiti*	*imeti občutek*

[2] For Czech, Polish, Russian, and Slovenian, the search query included all lexemes with the stems *-cítit*, *-czuć* and *-czuwać*, *-чувствовать*, and *-čutiti*, respectively. As a result of data analysis, the verbs given in Table 1 were identified.

[3] The Ukrainian stem *почувати* is typically used with a reflexive marker *-ся*, but it can also be used without it, with a reflexive pronoun *себе* and in other constructions. The distribution of constructions remains beyond the scope of this paper, and *почуватися* and *почувати* are regarded as the same verb.

To create the sample, I performed seven searches in the InterCorp corpus restricting the search to fiction. Importantly, I queried verbs only for one language at a time, whereas the fields for the other languages remained empty. In this way, I could find both the contexts where 'feel'-verbs in one language correspond to 'feel'-verbs in all or some of the other languages and the contexts where a 'feel'-verb is used only in the language in which the query was performed. Then the downloaded search results were combined in one Excel table, repetitions were removed, and the sample for further annotation was created. This sample contains 515 contexts. It is close to being balanced in terms of the language of the query and the author of the text. There are naturally some gaps in the data, where a translation is not available for a language, and they are especially frequent for Slovenian.

For each context in each language, I annotated the lexical item that corresponds to the 'feel'-verb in the query language and whether this lexical item belongs to 'feel'-verbs or not. At this stage, further lexical items, not included in the original queries, were identified as 'feel'-lexemes. These lexical items are given in the rightmost column of Table 1. Along with verbs, they include the noun(s) 'feeling' and in a number of languages, the combination of the verb 'have' and the noun 'feeling'. This expression was treated separately from other uses of the noun 'feeling', since the former is syntactically parallel to other 'feel'-verbs analysed in the study and typically more frequent than other constructions with the noun 'feeling'. An example of a possible correspondence between the sentences is given in (3).

(3)BG *Несъмнено почтеният младеж се* ***почувства*** *много нажален*

CR *Mladić se* ***osjećao*** *kukavno*

CZ *Pravda, hodný hoch* ***měl pocit*** *ubohosti*

PL *Nasz ambitny młodzian* ***poczuł się*** *dość żałośnie*

RU *Конечно, честный малый* ***чувствовал*** *себя довольно уныло*

UK *Звичайно, бідолашний парубок* ***почувався*** *досить кепсько*

'The poor fellow really felt sad.'

The contexts were annotated with respect to the semantic class of the situation denoted by the 'feel'-verb(s). If the renderings of the context in different languages didn't naturally fall into the same semantic class, the original context was used to annotate the semantic class. The resulting classification and the criteria used to distinguish between semantic classes are discussed in section 2.2.

2.2. Semantic classes

Semantic classification is always arbitrary to some extent, both because the borderline between semantic classes is not always clear-cut and because one can introduce more fine-grained or

more generalized semantic classes. The classification adopted in this paper incorporates the semantic domains of experience typically mentioned in the literature on 'feel'-verbs, as mentioned at the beginning of section 1. At the same time, as a result of the analysis of the contexts included in the sample, many more semantic classes had to be established to capture the semantic diversity of situations covered by 'feel'-verbs. The semantic classes are given below, ordered from more to less frequent in the sample.

Emotions constitute the most frequent semantic class in the sample. This class includes both temporary emotional states, as in (3) above, and more permanent descriptions of attitude.

Cognition in case of 'feel'-verbs has to do with the mental processing of perceived evidence and with the inference based on it, as in (4). This class encompasses situations that cannot be classified as any kind of sensory or emotional experience by the perceiver. At the same time, the inference made by the perceiver in such contexts often has to do either with the conditions imposed by the circumstances, i.e., with modal meanings, or with intentions or emotions of another person. The latter two semantic classes are also distinguished in the study and they can be considered as semantically close to cognition.

(4)BG - *И аз не мога - обади се Артър, **почувствал**, че е крайно време да се намеси по-решително.*

CR *"Ni ja ne znam", reče Arthur, koji je **osjetio** da je došlo vrijeme da i on uzme nekog udjela u svemu tome.*

CZ *"Já to taky neumím," řekl Arthur, který **ucítil**, že je načase, aby se i on začal prosazovat.*

PL - *Ja też nie - dodał Artur, który **poczuł**, że nadszedł czas, aby zaznaczyć swoją obecność.*

RU - *Я тоже не умею, - сказал Артур, **почувствовавший**, что настало время заявить о себе.*

SL *"Tudi jaz ne znam," je rekel Arthur, ki je **začutil**, da si mora utrditi položaj.*

Bodily sensations, exemplified in (2) above, include the perception of internal physiological processes, most frequently pain, weariness, hunger, heartbeat.

One of the semantic classes that arose over the course of data annotation is **well-being** (5). The contexts included in this class are largely indeterminate between physiological and emotional state and describe a general state of the experiencer.

(5)BG *А когато хвърляли някого само в душегубката, той се **чувствал** като отново роден.*

CR *A kad bi ga samo zatvorili u kulu, taj bi se čovek **osećal** ko preporođeni.*

SL *Če pa so ga pustili samo stradati, se je **počutil**, kot da je znova rojen.*

CZ *A když ho hodili jenom do lidomorny, to se takovej člověk **cítil** jako znovuzrozenej.*

PL *A jeśli wtrącili tylko takiego do wieży na śmierć głodową, to* ***czuł*** *się z uciechy jako nowo narodzony.*

RU *Если же преступника просто бросали в подземелье, на голодную смерть, то такой счастливчик* ***чувствовал*** *себя как бы заново родившимся.*

UK *А якщо злочинця кидали лише до підземелля, на голодну смерть, то він* ***почував*** *себе так, немовби знову на світ народився.*

'When they decided to only throw him into the hunger pit, such a man felt like he was being given a new lease on life.'

Another class which, to my knowledge, has not been discussed in relation to 'feel'-verbs so far is referred to as "**self-assessment**" (6). The situations of this class concern the way the experiencer perceives their own appearance, status or behaviour, as though stepping outside themselves.

(6)BG *"От това "господин Лангдън "се* ***чувствам*** *стар!"*

CR *Kad me se oslovljava s 'gospodine Langdon'* ***osjećam*** *se staro.*

CZ *Při panu Langdonovi si* ***připadám*** *starý.*

PL *Pan Langdon sprawia, że* ***czuję*** *się staro.*

RU *Когда я слышу "мистер Лэнгдон", я* ***ощущаю*** *себя стариком.*

UK *Від містера Ленгдона я* ***почуваюся*** *старим.*

'Mr. Langdon makes me feel old.'

Among skin sensations, I distinguish between those related to **environment** at large (most frequently, air quality, temperature, and movement) and **touch** proper, see example (1), although the two classes are both semantically and, as the study shows, linguistically very close.

In a number of contexts in the sample, 'feel'-verbs are used to express the awareness of desire or necessity, i.e., **modal meanings**, as in (7). While desire is closer to emotional or bodily sensations, (external) necessity is similar to cognition in that it typically involves a component of mental assessment.

(7)BG *и мене ме* ***обзе*** *желание да опиша всичко така, както си е*

CR *i* ***osjetio*** *sam želju da sve onako kako se dogodilo zapišem*

SL ***začutil*** *sem željo, da bi vse to zapisal*

CZ *a já jsem* ***pocítil*** *touhu všechno, jak to bylo, napsat*

PL ***Poczułem*** *wtedy potrzebę spisania tego, co się wydarzyło*

RU *и я* ***почувствовал*** *желание описать все как было*

UK *і я* ***відчув*** *бажання описати все як було*

'And I longed to write everything down just as it was.'

Several minor classes are close to environment but still are fairly well semantically defined. These are **smells**, an **emotion of another person**, and the **gaze of another person**.

A class that is interesting in terms of the semantic range of 'feel'-verbs, although marginal in terms of frequency encompasses uses in which 'feel'-verbs convey the idea of **affecting** the experiencer as such, without specifying in what way this affect will be carried about, see (8).

(8)BG *Ще го **изпитате** на собствената си кожа.*

CR ***Osjetit** ćete ga već na svojoj koži.*

SL *Le počakajte, dobro jo boste še **občutili**.*

CZ *Však ten zákon **pocítíte**.*

PL *Pan je jeszcze na sobie **odczuje***

RU *Вы его **почувствуете** на себе.*

'You'll find out when it affects you.'

Finally, several contexts in the sample describe situations related to hearing, sight, and taste.

The frequency of these semantic classes in the sample is given in Table 2. The semantic classes with the frequency of less than 5 contexts are not included.

Table 2. Frequency of semantic classes in the sample

Semantic class	N	%	Semantic class	N	%
Emotions	113	22%	Modal meanings	27	5%
Cognition	86	17%	Smell	16	3%
Bodily sensations	70	14%	Emotion of another person	15	3%
Well-being	56	11%	Affect	12	2%
Self-assessment	39	8%	Gaze of another person	8	2%
Environment	33	6%	Hearing	7	1%
Touch	28	5%			

Given such a diversity of semantic classes described by 'feel'-verbs, one could undertake a deeper semantic analysis of the relations between them and address an intriguing question of the possible diachronic development of 'feel'-verbs, see on the semantic range of the Slovenian verb *čutiti* in (Będkowska-Kopczyk, 2015). However, in this study I focus on the empirical analysis of the synchronic distribution of 'feel'-verbs across contexts.

In the next section, I examine the associations between these semantic classes and aspect, and in section 2.4, I turn to lexical distinctions drawn in individual languages. Due to the

limitations of the sample, only more frequent semantic classes and verbs will be considered. At the same time, the sample ensures the common ground for the analysis of verbs of different languages. I will focus only on the subset of the languages, showcasing the most frequent 'feel'-lexemes in Bulgarian, Czech, and Ukrainian. In these languages, the association between semantic classes and lexical distinctions appears to be more pronounced than in the others, at least as far as the sample data is concerned.

2.3. Semantic classes and aspect

Before turning to lexical distinctions in individual languages, I will briefly discuss the differences between semantic classes with respect to aspect. Although there is no perfect match between imperfectives and perfectives across Slavic, both in general (Dickey, 2000, Knjazev, 2007) and in the use of 'feel'-verbs, there are recurrent associations between semantic classes and aspect. To represent these associations, for the most frequent semantic classes, I calculated the mean proportion of imperfective and perfective verbs across the Slavic languages under analysis. These proportions are given in Table 3; the semantic classes are ordered from more to less associated with imperfectives.

Table 3. Mean proportions of imperfective and perfective 'feel'-verbs across Slavic for the semantic classes

Semantic class	Imperfectives	Perfectives
Well-being	0.90	0.10
Self-assessment	0.88	0.12
Emotions	0.68	0.32
Bodily sensations	0.67	0.33
Environment	0.53	0.47
Cognition	0.53	0.47
Modal meanings	0.53	0.48
Touch	0.40	0.60

Table 3 shows that well-being and self-assessment, exemplified by (5)-(6) above, are by far more frequently described by imperfective 'feel'-verbs than the other semantic classes. Although imperfectives generally are more frequent than perfectives, the highest proportions of perfectives are observed for touch and environment, where they highlight the moment of contact between the perceiver and the perceived object, as well as cognition and modal meanings, which are related to mental processing and inference, see (4) and (7) above.

2.4. Semantic classes and lexical distinctions in individual languages

2.4.1. 'Feel'-verbs in Bulgarian

Due to the data retrieval procedure, as described in section 2, only a subset of the 515 contexts included in the sample contain 'feel'-lexemes in sentences in a particular language. For example, for Bulgarian, the 'feel'-lexemes given in Table 1 above are used in 217 of 515 cases. There are several pairs of 'feel'-verbs in Bulgarian, each containing an imperfective and a perfective verb. The most frequent verbs are *чувствам – почувствам*. They have orthographic variants *чувствувам – почувствувам*, which are characteristic of older written works and are not widely used anymore (Petya Osenova, p.c.). Still, the two pairs are treated separately, as they are regarded as separate lemmas in the corpus. The other two pairs are *усещам –усетя* and *изпитвам –изпитам*. The distribution of uses of these verbs in the sample with respect to semantic classes is given in Table 4.

Table 4. Bulgarian 'feel'-lexemes and semantic classes

Semantic class	***чувствам почувствам***	***чувствувам почувствувам***	***усещам усетя***	***изпитвам изпитам***
Emotions	15		3	15
Cognition	10	7	12	1
Bodily sensations	10	6	10	6
Well-being	19	6		
Self-assessment	12	4		1
Environment		2	18	1
Touch	7	1	7	
Modal meanings	4			3
Smell			2	
Total	**80**	**27**	**60**	**32**

To analyse Table 4, we can look both at the columns, i.e., the semantic classes with the highest frequencies for a particular pair of verbs, and at the rows, i.e., the verb pairs preferably used to describe situations of a particular class. Table 4 shows that all the pairs of verbs are relatively heterogenous in terms of the semantic classes they cover. Still, some asymmetries in the distributions can be identified. The verbs *чувствам – почувствам* show a stronger association with well-being and self-assessment. They are also more frequently used to describe emotions than the next two pairs. Interestingly, *чувствувам – почувствувам* do not fully mirror the

distribution of the first pair, as they are less frequently used in contexts related to emotions and more frequently, in those related to cognition. The verbs *усещам – усетя* are primarily associated with the description of environment, and the verbs *изпитвам – изпитам*, with emotions and bodily sensations.

Russian exhibits similar lexical distinctions but the associations with semantic classes are less pronounced than in Bulgarian. Well-being and assessment are expressed almost exclusively by the verbs of the most frequent pair *чувствовать – почувствовать*, which are also used for other semantic classes. Against their background, the pair *ощущать – ощутить* appears to be associated with environment. The verbs *испытывать – испытать* are mainly used to describe emotions and bodily sensations.

2.4.2. 'Feel'-verbs in Czech

The most frequent Czech 'feel'-lexemes are *cítit, pocítit, ucítit, mít pocit,* and the verb *připadat* with the dative reflexive particle *si*. The distribution of these verbs in the sample with respect to the semantic classes is shown in Table 5 (*pocítit* is given together with the secondary imperfective *pociťovat*). Their total frequency in the sample (including less frequent semantic classes) is given in the bottom line.

Table 5. Czech 'feel'-lexemes and semantic classes

Semantic class	***cítit***	***pocítit pocit'ovat***	***ucítit***	***mít pocit***	***připadat si***
Emotions	21	16	1	1	1
Cognition	20	6	2	6	
Bodily sensations	21	3	2	2	1
Well-being	13				
Self-assessment	14			3	8
Environment	12	1	7	1	
Touch	13	3	3	1	
Modal meanings	4	3		2	
Smell	6		3		
Emotion of another person	4				
Total	**113**	**37**	**21**	**16**	**10**

Table 5 shows that the imperfective *cítit* is by far more frequent than the other verbs and that semantically, it is very general, as it is the most frequent verb for all the semantic classes. The perfective verbs *pocítit* and *ucítit* have a somewhat different distribution. The verb *pocítit* is more frequently used to describe emotions, whereas *ucítit* is associated with environment, touch, and smell. For the expression *mít pocit*, the most frequent semantic class is cognition, and *připadat si* is used almost exclusively for self-assessment.

2.4.3. 'Feel'-verbs in Ukrainian

In Ukrainian, similarly to Czech (as well as Polish and Slovenian, as suggested by Table 1), lexical differentiation in the domain of 'feel'-verbs primarily involves prefixes rather than different stems, as in Bulgarian and Russian. The most frequent 'feel'-verbs are *відчувати* and *відчути*, followed by the imperfective *почувати*(*ся*). The distribution of these three verbs with respect to semantic classes is given in Table 6.

Table 6. Ukrainian 'feel'-lexemes and semantic classes

Semantic class	*відчувати*	*відчути*	*почувати*(*ся*)
Emotions	14	11	13
Cognition	14	19	4
Bodily sensations	17	15	6
Well-being	2	1	14
Self-assessment	5	5	14
Environment	9	9	
Touch	10	6	
Modal meanings	3	4	
Smell		3	
Emotion of another person	5	2	
Total	**87**	**80**	**53**

Table 6 shows that the verbs *відчувати* and *відчути* are attested in contexts of all or almost all semantic classes, but in contexts describing well-being and self-assessment, they are clearly less favoured than *почувати*(*ся*). In addition to these two semantic classes, the latter verb is also frequently used to describe emotions.

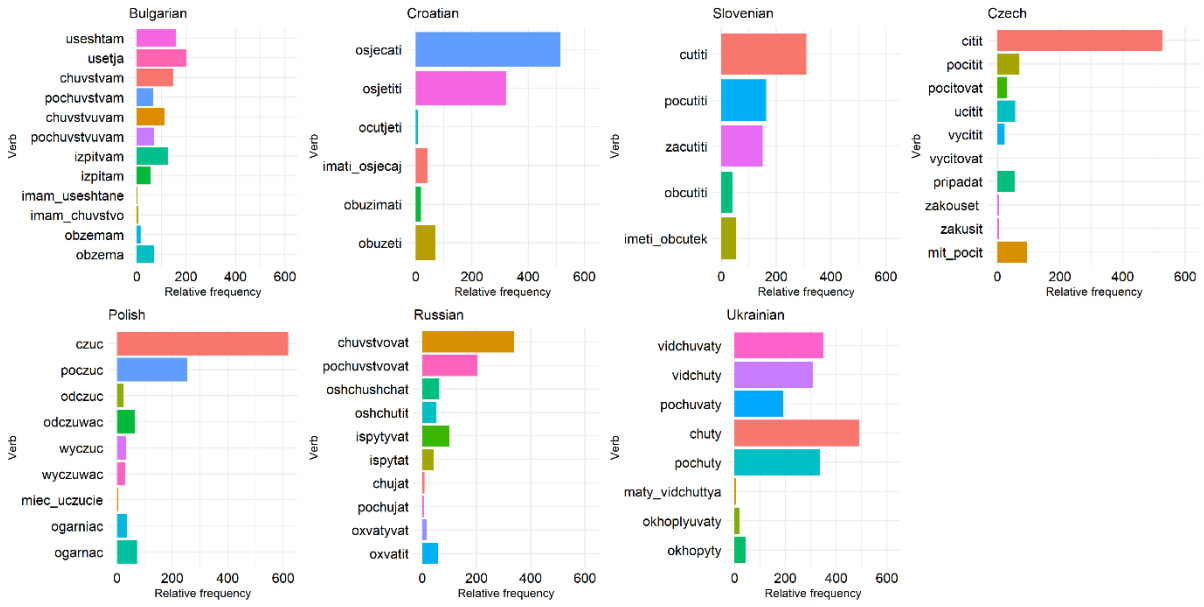
3. 'Feel'-verbs in Slavic: frequency and similarity

3.1. Setting the stage: frequency of Slavic 'feel'-verbs

The second part of the study was based on the frequencies retrieved from the parallel fiction texts of the InterCorp. The use of a larger text corpus made it possible to get more precise data on the frequencies of 'feel'-verbs and their correspondence, including less frequent members of the class. At the same time, these data do not lend themselves to direct examination of semantic distinctions which were discussed in the previous section.

Before analysing the correspondences between verbs and their semantic similarity, I look at the frequencies of 'feel'-verbs in individual languages. These frequencies were retrieved from monolingual subcorpora of the InterCorp, which were restricted to fiction. Although these texts are mostly translations, this corpus was chosen for a greater comparability of frequency data across languages. As the subcorpora are very different in size, the relative frequencies expressed in ipm (instances per million) were used. Bar plots in Figure 1 show the frequencies of the 'feel'-verbs identified for the seven Slavic languages[4].

Figure 1. Relative frequency of 'feel'-verbs in Slavic languages



Even though 'feel'-verbs across Slavic are largely etymologically close and structured by similar aspectual distinctions, the inventories of verbs and their frequency distributions appear to be relatively diverse. In the majority of Slavic languages, the most frequent 'feel'-verb is an imperfective, either morphologically basic with respect to its perfective counterparts, such as

[4] The multiword expressions of the type 'have' + 'feeling' were searched with the possibility of no or one intervening element between the two lemmas. The query results for the Czech verb *připadat* were filtered for the presence of the dative reflexive particle *si*. For Ukrainian, *почувати* includes the frequency of both *почувати* and *почуватися*.

čutiti in Slovenian, *cítit* in Czech, *czuć* in Polish, *чувствовать* in Russian, or derived, such as *osjećati* in Croatian and *відчувати* in Ukrainian. In some languages, most clearly in Bulgarian and Russian, the verbs can be organized in aspectual pairs. In other languages, the aspectual relations are less straightforward, with several prefixed perfectives derived from the basic imperfective verb, as in Czech, Polish, and Slovenian.

The expressions featuring the verb 'have' and the noun 'feeling' are relatively infrequent in Slavic, except for Czech. For the transitive verbs with the perceived object in the subject position, such as Croatian *obuzeti* or Ukrainian *охопити*, the perfective counterpart is more frequent in all the languages where such verbs were included in the search queries.

The plot for Ukrainian includes the verbs *чути* and *почути*, which are primarily used to denote hearing[5]. Still, they can also be used to describe other sensation types, typically expressed by 'feel'-verbs in other languages, for instance touch (9) and smell (10). Figure 1 represents their overall frequencies, regardless of the meaning.

(9)BG *Докосвам я по рамото, но* ***усещам****, че някой ме хваща за ръката.*

CR *Samo što joj nisam dotaknuo rame kad* ***osjetim*** *kako me netko hvata za ruku.*

UK *Я збираюсь торкнутися її плеча, як* ***чую****, що хтось бере мене під руку.*

'I was going to touch her shoulder, when I felt someone taking me by the hand.' (translation is mine)

(10) CR *I tada je ponovno* ***osjetio*** *miris.*

CZ *A právě tehdy* ***ucítil*** *ten pach.*

RU *Вот тогда-то он и* ***почуял*** *этот запах.*

UK *Тоді ж бо він і* ***почув*** *цей запах.*

'And then he felt this smell again.' (translation is mine)

In the next section, I discuss how the correspondences between 'feel'-verbs were retrieved and analysed in order to explore their semantic similarity.

3.2. Correspondences between 'feel'-verbs and their semantic similarity

To examine to what extent Slavic 'feel'-verbs are similar in terms of contexts of use and, therefore, semantics, I collected and analysed the frequency of correspondences between these verbs in parallel texts, see a methodologically similar study on sight verbs in Bulgarian, Polish, and Russian in Author (under review).

[5] The development of this root to denote hearing, as a result from the narrowing or shift from the more general meaning of perception is also observed in South Slavic (Preveden, 1932, p. 148; Grković-Mejdžor, 2011).

For this part of the study, I took 48 ‘feel’-verbs, including the majority of verbs shown in Figure 1. Excluded were the verbs with the perceived object in the subject position, such as Bulgarian *обзема* or Russian *охватить*, as they are relatively infrequent and lack the syntactic parallelism with the other verbs selected for the analysis. I also excluded the Czech verbs *zakoušet* and *zakusit*, as there were too few correspondences between them and the ‘feel’-verbs of the other languages.

For each pair of languages, I searched for all pairwise correspondences between ‘feel’-verbs in a bilingual parallel subcorpus restricted to fiction. As the texts are aligned with respect to sentences rather than words, some of the searches contained sentences where the queried verbs did not refer to the same situation. For this reason, the search results were looked through and the frequencies of correspondences were corrected based on the proportion of correct matches in the first 20 occurrences. As an example, Table 7 shows correspondence frequencies for the Czech and Polish ‘feel’-verbs.

Table 7. Correspondence frequencies for pairs of Czech and Polish ‘feel’-verbs

	cítit	*pocítit*	*pociťovat*	*ucítit*	*vycítit*	*vyciťovat*	*připadat si*	*mít pocit*
czuć	485	15	16	16	4	1	44	39
poczuć	115	45	2	61	2	0	12	12
odczuć	4	8	2	1	1	0	1	1
odczuwać	31	3	15	0	0	0	1	0
wyczuć	8	1	0	5	11	0	0	1
wyczuwać	21	0	1	1	6	1	0	1
mieć uczucie	1	0	0	0	0	0	0	4

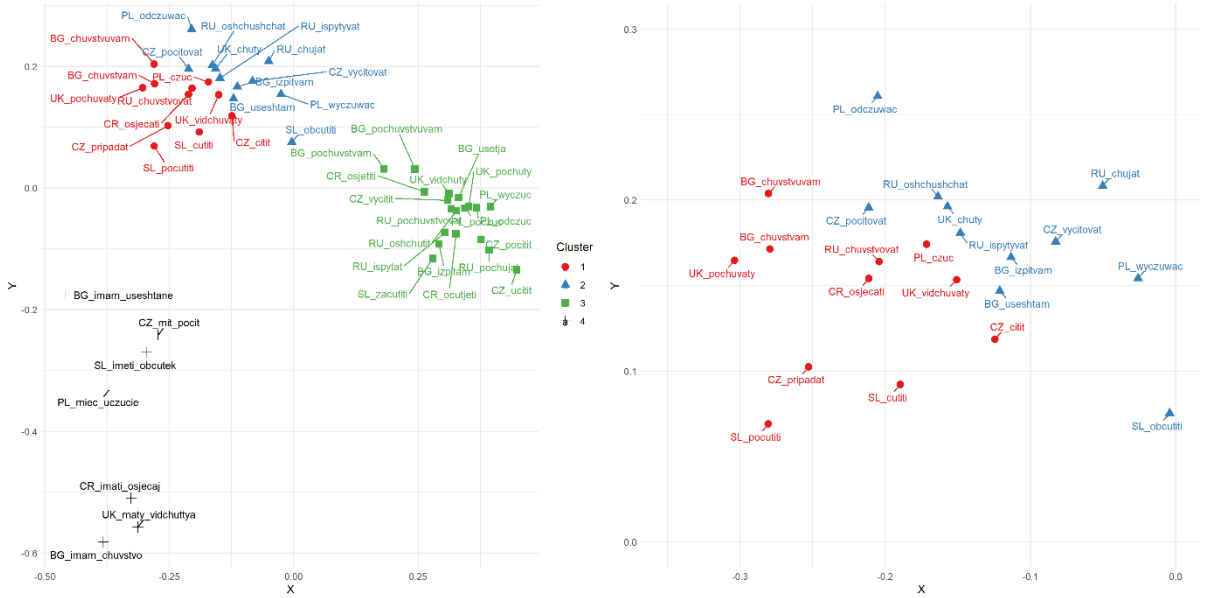
The bilingual subcorpora of different pairs of languages can be very different in size. To make the data on correspondence frequencies comparable across pairs of languages, I divided each table by the sum of all correspondences and multiplied by one thousand. Thus, the resulting tables of correspondences for each pair represented the distribution of one thousand contexts with ‘feel’-verbs.

The next step was to calculate pairwise distances between the verbs based on these correspondences. Both for the two verbs of the same language and for the two verbs of different languages, I compared the distribution of their correspondences to the verbs in all the remaining

languages. Similarity between the two distributions was measured using the cosine similarity, which is often used to compare documents on the basis of words they contain (Singhal, 2001); the cosine similarity was calculated using the function cosine() of the R package *lsa* (Wild, 2022). Then, cosine similarity was turned into distances between verbs but subtracting it from one. As a result, I got a distance matrix containing all the pairwise distances between the verbs.

To analyze the resulting distance matrix, I used the fuzzy clustering algorithm, as implemented in the R package *cluster* (Maechler et al., 2024). I opted for a clustering solution with four clusters. Thus, for each verb the algorithm determined to what extent it belongs to each of the four clusters, and the cluster with the highest value was used to divide the verbs into groups. Then the distances were visualized using the multidimensional scaling (MDS) algorithm, which is used to represent distances in a low number of dimensions aiming at the minimal distortion of the original distances. I used the R package *smacof* (Mair et al., 2022) to create an MDS model, and the packages *ggplot2* (Wickham, 2016), *ggrepel* (Slowikowski 2023), and *RColorBrewer* (Neuwirth, 2022) to create the plot. Figure 2 shows the two-dimensional MDS visualization for the distances between the verbs, with the clusters plotted by different symbols and colors. The plot on the left shows the grouping of verbs into the four clusters and the plot on the right zooms in on the two clusters located in the upper left corner of the first plot.

Figure 2. MDS plot of Slavic ‘feel’-verbs and their clusters



In general, the plot on the left shows three spatial groupings of verbs, which are also identified if the three-cluster solution is opted for. The less tight cluster in the bottom left corner encompasses all the constructions featuring the verb 'have' and the noun 'feeling'. This suggests that across Slavic, they are semantically closer to each other than to the 'feel'-verbs proper. As discussed in section 2.4.2, in Czech, where this construction is more frequent than in the other languages, it is associated with cognition, i.e., situations where the perceiver makes a mental conclusion on the basis of the perceived circumstances, as in example (11). It is likely that the same association exists in other Slavic languages, although in them this expression is less widely used.

(11) BG *Тази нощ най-после* ***усещаше****, че е започнал да изплаща дълга си.*
CR *Večeras, konačno, Silas* ***je osjećao*** *da je počeo vraćati svoj dug.*
CZ *Dnes večer* ***měl*** *Silas konečně* ***pocit****, že začal splácet svůj dluh.*
PL *Dziś wieczór Sylas nareszcie* ***poczuł****, że zaczął spłacać swój dług.*
RU *По крайней мере сегодня Сайлас* ***чувствовал****, что начал оплачивать долги.*
UK *Сьогодні Сайлас* ***відчув****, що нарешті почав віддавати борг.*
SL *Nocoj* ***je začutil****, da je končno začel odplačevati dolg.*
'Tonight, at last, Silas felt he had begun to repay his debt.'

The other two groupings mostly correspond to imperfectives (upper left corner) and perfectives (righthand part of the plot). The exceptions to this pattern are Slovenian verbs *počutiti* and *občutiti*, which are grouped together with imperfectives; still, the latter can function both as a perfective and an imperfective verb, see Będkowska-Kopczyk (2016) on the Slovenian verbs *občutiti* and *začutiti*.

In the four-cluster solution, opted for in this study and visualized on the plot, the grouping of verbs in the upper left corner is further subdivided into two clusters shown on the separate plot on the right. As these clusters are not clearly separated from each other, verbs lying at the border of the two clusters are likely to show more similarity to the verbs of the other cluster than those lying at the outer edges of the two clusters. To interpret the semantic relations between these verbs, I will use the associations established in section 2.4 as reference points.

In general, the subdivision of imperfectives into two clusters largely mirrors the distinction between verbs used to describe well-being and self-assessment, on the one hand, and verbs which tend to be used with external stimuli, such as environment, touch and smell, on the other hand. This interpretation is supported by the fact that the Bulgarian verbs *чувствам* and *усещам*, which are associated with this distinction, appear in two different clusters. The same applies to the Russian verbs *чувствовать* and *ощущать*, which show a similar semantic

contrast, albeit in a weaker form. The Ukrainian verb *почувати*(*ся*) is also mostly used for well-being and self-assessment, as confirmed by its location next to *чувствам*. However, in Ukrainian the semantic contrast between *почувати*(*ся*) and the other frequent imperfective verb *відчувати* must be less sharp than between *чувствам* and *усещам* in Bulgarian, since both Ukrainian verbs are found in the same cluster. The Czech verb *připadat si* is strongly associated with self-assessment. The fact that Slovenian verbs *počutiti* and *čutiti* are located close to this Czech verb may suggest that they are also often used for situations of this semantic class. The other verbs belonging to this cluster are the basic imperfective 'feel'-verbs: Croatian *osjećati*, Czech *cítit*, and Polish *czuć*. They are found closer to the other cluster, suggesting their more general semantic profile.

The common feature of the verbs included in the second cluster is that they do not occur in the contexts describing well-being and self-assessment. As already mentioned, some verbs of this cluster are associated with environment and other external stimuli: Bulgarian *усещам* and Russian *ощущать*. As the sample data suggest, the group of verbs in the rightmost part of the plot, which includes Russian *чуять*, Czech *vyciťovat*, and Polish *wyczuwać* are also used for external stimuli, specifically for the perception of another person's emotions, as in (12). The Polish verb is the most frequent of the three, as shown in Figure 1, and in parallel contexts it corresponds to the Czech perfective *vycítit* more frequently than to the rare imperfective *vyciťovat*, see Table 7.

(12) BG *Очевидно* ***усетила*** *затруднението му, Виторија се обърна към екскурзовода.*
CR *Očito* ***osjećajući*** *da se našao bezizlaznoj situaciji, Vittoria se okrenula docentu.*
CZ *Vittoria zřejmě* ***vycítila****, že je v úzkých, a obrátila se proto k průvodci.*
PL *Najwyraźniej* ***wyczuwając****, że znalazł się w trudnej sytuacji, Vittoria zwróciła się do przewodnika.*
RU ***Почувствовав*** *смятение американца, девушка повернулась к гиду.*
UK *Очевидно,* ***відчувши*** *розгубленість Ленґдона, Вітторія знову повернулась до гіда.*
'Apparently sensing Langdon's stymie, Vittoria turned to the docent.'

Other verbs in this cluster are associated with the description of emotions. These verbs include Polish *odczuwać* and Czech *pociťovat* in the upper part of the cluster, as well as Bulgarian *изпитвам* and Russian *испытывать*.

The cluster which contains perfective verbs, located at the right side of the left plot, does not show any distinct subgroupings which could be identified with semantic classes. This may be due to the fact that the semantic classes of well-being and self-assessment, which largely

contribute to the clusters within the imperfectives, are mostly associated with imperfective verbs, see section 2.3. Since perfectives are in principle rarely used to describe situations of these semantic classes, there is less contrast between the perfective verbs.

4. Conclusion

The study examined Slavic verbs with the meaning ‘feel’ analyzing their semantics and usage on the basis of parallel corpus data from the InterCorp parallel corpus. Two datasets were mainly used.

The first dataset contained a sample of parallel contexts containing a ‘feel’-verb in at least one of the seven Slavic languages under analysis. This dataset testified to the diversity of situations covered by ‘feel’-verbs. Among the most frequent semantic classes in the dataset were emotions, cognition, and bodily sensations. Somewhat less frequent semantic classes which were identified on the basis of the sample and are probably specific to ‘feel’-verbs have to do with general well-being and the assessment of one’s own status or appearance by the perceiver. A closer examination of the inventories of ‘feel’-verbs in Bulgarian, Czech, and Ukrainian suggested that these two semantic classes often lie at the basis of lexical distinctions within ‘feel’-verbs and in such cases are opposed to semantic classes concerned with external stimuli, such as environment, touch, and smell. Well-being and self-assessment also show aspectually skewed distribution, as they tend to be by far more frequently described by imperfective verbs than the situations of other semantic classes.

The second dataset contained the frequencies of correspondence between verbs in a parallel corpus of fiction. These data made it possible to investigate semantic similarity between a wider range of ‘feel’-verbs. In particular, the analysis showed that expressions featuring the verb ‘have’ and the noun ‘feeling’ group together despite differences in their frequencies across languages. The other two larger groups include (mostly) imperfectives and perfectives, confirming the overarching role of aspect in structuring verbal lexicon in Slavic. Imperfectives can be further subdivided into verbs that cover the semantic classes of well-being and self-assessment. This group includes both more semantically specific verbs, such as Bulgarian *чувствам* and Ukrainian *почувати*(*ся*), and verbs with more general meaning, such as Czech *cítit* and Slovenian *čutiti*. These verbs are opposed to ‘feel’-verbs which do not cover these two semantic classes: some of them typically describe situations with external stimuli, as Bulgarian *усещам* or Polish *wyczuwać*, and some are mostly used for emotions, such as Czech *pociťovat* and Russian *испытывать*.

Although language-specific studies, based on more data, may capture the semantic distinctions between 'feel'-verbs in a more fine-grained and precise way, the approach based on parallel corpus data makes it possible to compare the inventories and semantic ranges of 'feel'-verbs using the same of a similar set of contexts. The study highlights that there are common patterns within the 'feel'-domain drawn by lexical distinctions across Slavic. At the same time, it shows that languages with similar inventories of verbs (in terms of the general makeup or even cognate relations), such as Bulgarian and Russian or Czech and Polish, can diverge in terms of their frequency distribution and in the strength of their associations with semantic classes. Further research in this domain will show whether these patterns are specific to Slavic or they concern a wider range of languages.

References

Będkowska-Kopczyk, A. (2015). Glagol *čutiti* v slovenščini: med slovarjem, skladnjo in rabo. In *Slovnica in slovar - aktualni jezikovni opis* (93-100). Ljubljana. https://repozitorij.uni-lj.si/IzpisGradiva.php?lang=slv&id=166467

Będkowska-Kopczyk, A. (2016). *Začutiti* in *občutiti*: Kognitivna analiza pomensko-skladenjskih lastnosti dveh predponskih tvorjenk iz glagola *čutiti.* In *Obdobja 35*. Ljubljana: Center za slovenščino kot drugi in tuji jezik. https://centerslo.si/wp-content/uploads/2016/11/Bedkowska.pdf.

Dickey, S. M. (2000). *Parameters of Slavic Aspect. A cognitive approach*. Stanford, CACSLI Publications.

Grković-Mejdžor, J. (2011). Semantika glagola čulne percepcije u staroslovenskom jeziku, *Slavistična revija*, 59(1), 21–30. https://srl.si/ojs/srl/article/view/98d0b8ac4df74c2dbfb1d12f11eb5345

Knjazev, Y. P. (2007). *Grammatičeskaja semantika: Russkij jazyk v tipologičeskoj perspective.* Moscow.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. (2024). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.8.

Neuwirth, E. (2022). _RColorBrewer: ColorBrewer Palettes_. R package version 1.1-3, <https://CRAN.R-project.org/package=RColorBrewer>.

Preveden, F. R. (1932). Some Balto-Slavic Terms of Acoustic Perceptions, *Language*, 8(2), 145-151. https://doi.org/10.2307/409311. https://www.jstor.org/stable/409311

Rosen, A., Vavřín, M., Zasina, A. J. (2022). InterCorp, Release 15 of 11 November 2022. Institute of the Czech National Corpus, Charles University. Available at: http://www.korpus.cz

Rosen, A. (2023). The InterCorp Parallel Corpus with a Uniform Annotation for All Languages, *Journal of Linguistics/Jazykovedný časopis*, 74(1), 254-267.

Singhal, A. (2001). Modern Information Retrieval: A Brief Overview, *IEEE Data Engineering Bulletin*, 24(4), 35-43.

Slowikowski, K. (2023). _ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'_. R package version 0.9.3, <https://CRAN.R-project.org/package=ggrepel>.

Viberg, Å. (1984). The verbs of perception: a typological study, *Linguistics*, 21.1, 123-162.

Wickham H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag, New York.

Wild, F. (2022). lsa: Latent Semantic Analysis. R package version 0.73.3, <https://CRAN.R-project.org/package=lsa>.

Submitted on the 11th of February, 2025