

What it means to feel:
A parallel-corpus study of ‘feel’-verbs in six European languages

1. Introduction

In linguistics, the domain of perception is typically represented in the form of a paradigm (Viberg 1984; 2001), as shown in Table 1 for English.

Table 1. English perception verbs paradigm

	Experiencer-based		Phenomenon-based
	Experience	Activity	
Sight	<i>see</i>	<i>look</i>	<i>look</i>
Hearing	<i>listen</i>	<i>hear</i>	<i>sound</i>
Touch	<i>feel</i>	<i>feel</i>	<i>feel</i>
Taste	<i>taste</i>	<i>taste</i>	<i>taste</i>
Smell	<i>smell</i>	<i>smell</i>	<i>smell</i>

The rows of this paradigm correspond to the five traditionally identified sense modalities. The columns distinguish between three predicate types. Experiencer-based predicates take experiencer as their subject. They are subdivided into experience and activity predicates, which differ in whether perception takes place without or with control of the experiencer, respectively. The third type, phenomenon predicates, take the stimulus of perception as their subject.

In many languages, the verb used for some of the sense modalities can also describe emotional states and bodily sensations. In particular, this is the case of the English verb *feel*, which is the basic verb for all types of verbs expressing touch (Viberg 1984, 139). An areal pattern observed by Viberg (1984: 143) and further discussed by Raatikainen (2021) is that the experience verb used for touch is also used for smell and taste across a number of European languages, including Swedish, Estonian, Finnish, Hungarian, and at least some Slavic languages. At least in Slavic, the same verbs also describe experiential events outside of the perception paradigm, cf., the Russian verb *chuvstvovat* or the Croatian verb *osjetiti*. The reflexes of the Latin verb *sentire* ‘perceive’ continue to be used in the general sense of perception across Romance; as discussed in Viberg (1984, 149), in some languages, this verb has become the major or one of the major ways of describing hearing, but can also be used for touch, taste, and smell.

The focus on the meanings from the paradigm of perception results in the fact that the full semantic range of the verbs meaning ‘feel’ has not been subject to a detailed investigation,

cf. Wälchli's (2016) critical assessment of the paradigmatic approach to the domain of perception and the senses left beyond the traditional "five senses folk model" in Winter (2009: 11-13), see also Norcliff, Majid (2024: 84). The fact that 'feel'-verbs lie at the intersection of different experiential domains recognized in linguistics has been demonstrated in studies of colexification patterns. For instance, the Database of Cross-Linguistic Colexifications (CLICS, <https://clics.clld.org/>, Rzymiski et al. 2019) distinguishes between two concepts, FEEL 'to be in some emotional state' and FEEL (TACTUALLY) 'to sense something by touching its surface'. Both have semantic connections to sense modalities, such as hearing, as well as to the domain of cognition (UNDERSTAND, KNOW). These connections are further scrutinized by Georgakopoulos et al. (2021: 21), who note that the concept FEEL "covers different types of perception (ranging from emotional sensation and particular state of mind to examination by touching)".

The goal of the present study is to explore the semantic range of the major 'feel'-verbs in six European languages and to determine which meanings are more central and which are more peripheral to these verbs. I also analyse the grammatical constructions used with these verbs and their associations with semantic classes, and lexical distinctions drawn in the domain covered by 'feel'-verbs.

The data for the study come from the parallel corpus InterCorp, which is accessible through the interface of the Czech national corpus (Rosen et al., 2022; Rosen, 2023). I used the subcorpus of fiction texts available for the following six languages: English, German (Germanic), French, Spanish (Romance), Czech, Russian (Slavic). The languages were chosen based on the size of the subcorpora available for them. Parallel corpora have been already used for the exploration of a number of semantic domains across European and other languages, such as the studies by Wälchli, Cysouw (2012) on motion verbs, Levshina (2015) on the semantic space of giving, Becker & Guzmán Naranjo (2020) on psych predicates, and Wälchli (2016) on perception verbs.

In their foundational work on lexical typological research based on parallel corpora, Cysouw & Wälchli (2012) propose an exemplar-based approach to the construction of semantic maps. Such probabilistic semantic maps show the clustering of exemplars in terms of the similarity of (lexical) expression, and the typologically relevant functions and semantic contrasts are derived from these clusters. For instance, one of the contrasts identified and discussed by Cysouw & Wälchli (2012) is that between 'come' and 'go'. To describe these contrasts and function, researchers often use labels based on English lexical items, which are often (viewed as) self-explanatory. For the semantic domain covered by 'feel'-verbs, this

approach is problematic because there are no well-established semantic contrasts in this domain that are recognized in the literature or can be easily captured by the verbs of a specific language. For this reason, to explore the semantic structure of the ‘feel’-domain, I employed a semantic classification that emerged from the analysis of individual exemplars during data annotation, rather than from an analysis of similarity between exemplars. This semantic classification was then mapped onto the grouping based on grammatical and lexical similarity among exemplars in order to determine which semantic classes are supported by the data.

The paper is structured as follows. In section 2, I discuss data retrieval and sample annotation and present the ‘feel’-verbs analysed in the study. In section 3, I propose the semantic classification which was created based on the sample and used to explore the semantic range of ‘feel’-verbs under analysis and discuss the frequency of these classes in the sample. In section 4, I examine the major constructions featuring ‘feel’-verbs and their distribution with respect to semantic classes. In section 5, I visualize the structure of the domain covered by ‘feel’-verbs based on the constructional and lexical features of the parallel sentences from the sample and use this visualization to examine lexical distinctions drawn in the languages under analysis. In section 6, I summarize the findings of the study assessing their contribution to the research on the domain of perception.

2. Data and annotation

2.1. Corpus queries and initial data processing

The InterCorp corpus (Rosen et al., 2022; Rosen, 2023), which was used as the source of data for this study, is developed as a part of the Czech National Corpus. In the InterCorp, different subsets and types of parallel texts are available for different pairs of languages. From the start, I decided to restrict the study to fiction and exclude subtitles, which constitute a large part of the corpus. Although subtitles were shown to be similar to spontaneous oral speech (Levshina 2017), which cannot be expected of fiction, they were not used in this study for the following reasons. First, subtitles provide less context to determine the semantic class of the situation described by the clause, which is of major concern for the present study. Second, in the InterCorp corpus, the language of the original is less consistently annotated for subtitles as compared to fiction. It was one of the research questions of the study to check whether the presence of a ‘feel’-lexeme in the original makes it more likely that such lexemes are also used in translations, therefore this parameter of the text had to be controlled for.

To create a sample of parallel sentences, I used a subcorpus of texts available in all the six languages. Then I queried the major ‘feel’-verbs for each of the languages, as shown in Table 2.

Table 2. Queried ‘feel’-verbs

Language	Queried ‘feel’-verbs
English	<i>feel, sense</i>
German	<i>fühlen, spüren, empfinden, verspüren</i>
Czech	<i>cítit, pociťit, ucítit, vycítit, pociťovat</i>
Russian	<i>čuvstvovat', počuvstvovat', oščučšat', oščutit'</i>
French	<i>sentir, ressentir</i>
Spanish	<i>sentir</i>

The verbs were queried for one language at a time, i.e., only for the language which was set as primary, whereas the fields for all the other languages were empty. This means that queried ‘feel’-verbs of one language do not necessarily correspond to queried ‘feel’-verbs in the other languages. The query results for each language were downloaded, merged together, and the exemplars that were found more than once were removed.

As some of the verbs in Table 2 can be used as (or, under an alternative analysis, are homonymous with) verbs that take the perceived object as the subject, as in the sentence *The house felt empty*, I additionally ensured that all of the query results include only experiencer-subject verbs, so that this parameter is kept constant for all the languages.

The final sample contains 597 exemplars balanced in terms of the language in which the query was made and the author of the text. The sample includes 100 exemplars per each of the query languages and 24 exemplars per author.¹ Each exemplar is represented by six parallel sentences, of which at least the sentence in the language of the query necessarily contains one of the ‘feel’-verbs from Table 2.

¹ There is one exception: for Czech, only one instance of queried ‘feel’-verbs was found in Carroll’s “Alice’s Adventures in Wonderland”, therefore only for 97 exemplars Czech is the language of the query and there are 21 exemplars authored by Carroll.

2.2. Sample annotation

The first step of sample annotation was to determine, for each exemplar, how the event described by a ‘feel’-verb in the query language is expressed in the parallel sentences in each of the other languages. The most straightforward case is when a ‘feel’-verb is used, as in (1).

(1) *suddenly I **felt** the tears coming again*

*Aber auf einmal **fühlte** ich wieder die Tränen aufsteigen*

*Ale najednou jsem zase **ucítila**, jak se mi derou slzy do očí*

*No vdrug opjat' **počuvstvovala**, čto sejčas rasplačus'*

*j' ai **sentí** les larmes monter de nouveau*

*de repente **sentí** que me venían las lágrimas*

Additionally, further verbs and constructions, not included in the queries, were identified as expressing the meaning ‘feel’, for instance possessive constructions with the noun ‘feeling’,² or verbs frequently used as translational equivalents of the queried ‘feel’-verbs of other languages. The list of ‘feel’-verbs identified for each language as a result of data annotation is given in Table 3, including the verbs already given in Table 2 above. This table also shows the frequencies of each of the ‘feel’-verbs.³

Table 3. ‘Feel’-verbs in the languages under analysis: queried and identified as a result of data annotation

Language	Queried ‘feel’-verbs	Further ‘feel’-verbs
English	<i>feel</i> (246), <i>sense</i> (7)	<i>experience</i> (4), <i>have a feeling</i> (5), <i>have a sense</i> (2)
German	<i>fühlen</i> (120), <i>spüren</i> (45), <i>empfinden</i> (38), <i>verspüren</i> (5)	<i>Gefühl haben</i> (9), <i>sich vorkommen</i> (6)
Czech	<i>cítit</i> (132), <i>pocítit</i> (31), <i>ucítit</i> (19), <i>pociťovat</i> (14), <i>vycítit</i> (7)	<i>připadat si</i> (13), <i>mít pocit</i> (10), <i>vycítovat</i> (2), <i>zakoušet</i> (2)

² In all the languages except Russian these constructions feature the verb ‘have’. In Russian, the noun ‘feeling’ in the possessive construction is the subject of the verb ‘be’. As these constructions predicate the fact of feeling, they were regarded as parallel to ‘feel’-verbs proper and for simplicity will be further included in ‘feel’-verbs. Nominal constructions with nouns meaning ‘feeling’ other than possessive ones, e.g., ‘with a feeling’, as well as verbal constructions where the experiencer is not the subject were excluded from the analysis.

³ The frequency distribution of ‘feel’-verbs in the sample only approximately reflects their relative frequencies in the respective languages, because different number of verbs was included in the queries for the languages.

Russian	<i>čuvstvat'</i> (98), <i>počuvstvat'</i> (64), <i>oščuščat'</i> (20), <i>oščutit'</i> (7)	<i>ispytyvat'</i> (14), <i>počujat'</i> (3), <i>čujat'</i> (3), <i>ispytat'</i> (3), <i>učujat'</i> (2), <i>byt'</i> <i>čuvstvo</i> (2), <i>byt' oščuščenje</i> (1)
French	<i>sentir</i> (198), <i>ressentir</i> (30)	<i>éprouver</i> (22), <i>avoir impression</i> (6), <i>avoir sentiment</i> (1)
Spanish	<i>sentir</i> (239)	<i>encontrarse</i> (13), <i>experimentar</i> (5), <i>haber sensación</i> (3)

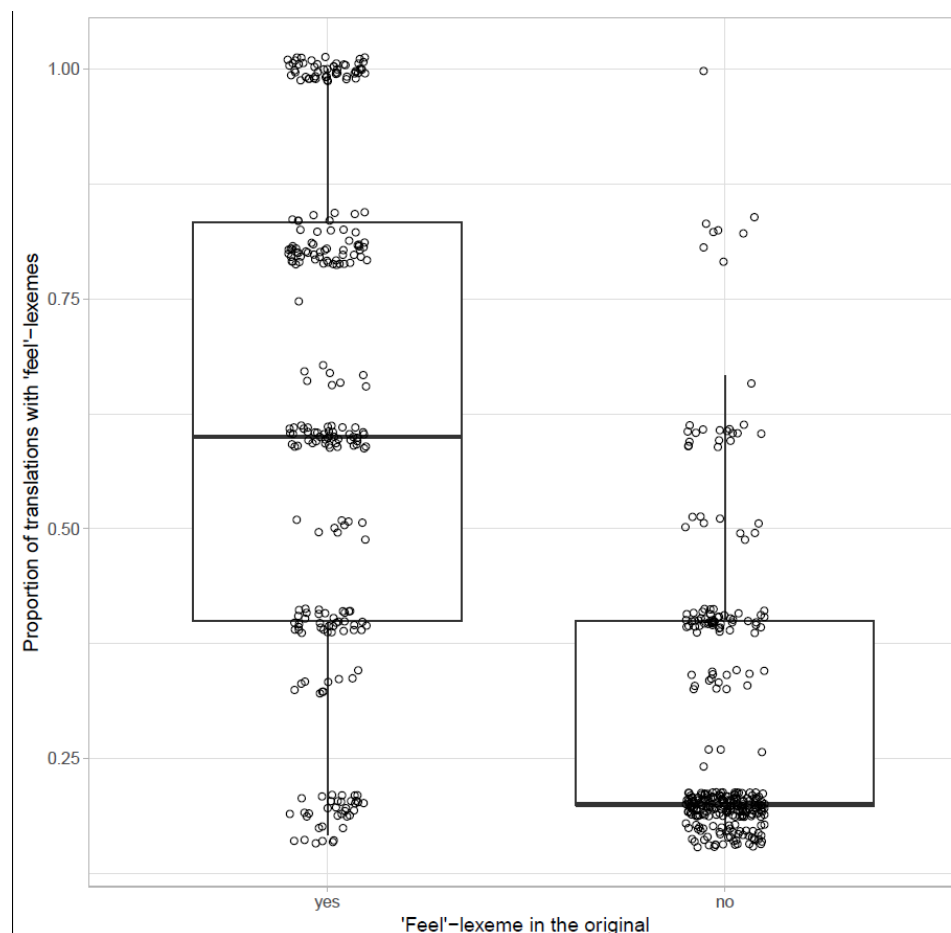
Sentences with ‘feel’-verbs in each of the languages were also annotated for the grammatical properties of construction with the ‘feel’-verb and the expression of the perceived object, see section 4.1.

For each exemplar, I also annotated the language of the original and included the original sentence in the dataset. The sample is heavily imbalanced in terms of the language of the original, with almost a half (261 out of 597) of the exemplars written originally in English. 72 exemplars are originally written in Czech. French, German, and Italian are the source languages of 48 exemplars each, and Dutch, Portuguese, Polish, Russian, and Norwegian are the source languages of 24 exemplars each.

The original sentences were annotated in terms of whether they include a ‘feel’-verb. To examine the influence of the original sentence, for each exemplar I calculated the proportion of translations in the sample where a ‘feel’-verb is attested to the total number of available translations. The total number depended on what the language of the original is and whether there are any gaps in the data (e.g., when the relevant part of the work was simply omitted by the translator). The distribution of proportions depending on whether a ‘feel’-verb is present or absent in the original is shown in Figure 1.⁴

⁴ In some cases, the only language in the sample containing a ‘feel’-verb is the original, resulting in the proportion 0 across translations. Such cases are not shown in Figure 1.

Figure 1. The proportion of ‘feel’-verbs in translations for exemplars with and without a ‘feel’-verb in the original



As Figure 1 shows, if a ‘feel’-verb is used in the language of the original, it is also much more likely to be used across translations (Wilcoxon rank sum test, $p < 0.001$). Since the semantic range of ‘feel’-verbs can differ across languages, the presence of a ‘feel’-verb in the original can bias the distribution of ‘feel’-verbs in translations. For this reason, the presence of a ‘feel’-verb in the original sentence will be controlled in the analysis of the semantic classes expressed by ‘feel’-verbs, presented in section 3.

3. Semantic classes covered by ‘feel’-verbs

Each exemplar of the sample was annotated with respect to the semantic class of the situation it describes. In case when different translations represented the situation expressed by a ‘feel’-verb in the queried sentence somewhat differently, the annotation of semantic class was based primarily on the original sentence and on the sentence retrieved as a result of the query.

Semantic classes of situations covered by ‘feel’-verbs are the central parameter considered in this study and at the same time the least tangible aspect of the undertaken analysis. The annotation of semantic class didn’t follow any pre-established classification and was carried out bottom-up, with the idea to capture the semantic classes relevant for ‘feel’-verbs. For this reason, along with traditionally distinguished semantic classes, such as emotions, cognition and bodily sensations, the situation types included several less conventional classes. The identification of some of these classes and their relevance for the domain covered by ‘feel’-verbs can be regarded as a result of the study in its own right.

As a result of data annotation, 19 semantic classes were identified, which testifies to a broad semantic range of ‘feel’-verbs under analysis. This might look like too many classes, but there seem to be no natural higher-level semantic groupings that some of these classes could be lumped into, although some semantic connections between them can be suggested. In sections 4 and 5, such connections will be established as a result of the data analysis. Table 3 shows the list of these classes and their frequency in the sample. Below, I give definitions of the classes, also used as guidelines for the annotation.⁵ The order of the classes in the description is based on their presumed semantic similarity.

Table 3. Frequency of semantic classes in the sample

Semantic class	N of exemplars	Semantic class	N of exemplars
emotions	150	necessity	16
body	78	wish	16
cognition	62	smell	15
certainty	49	agentive touch	9
touch	44	movement	7
well-being	41	gaze	6
attitude	29	affect	5
self-assessment	26	hearing	5
another’s emotion	19	sight	2
environment	18		

⁵ The subsample of 300 exemplars was additionally annotated for cross-validation. The judgements on the semantic class diverged in 69 out of 300 mismatches. The evaluation of inter-annotator agreement using Cohen’s *kappa* (as implemented in the R package *psych* (Revelle 2025)) showed substantial agreement between the two annotators ($\kappa = 0.74$).

Table 3 shows that the largest class in the sample encompasses exemplars describing **emotions**, which were taken to include only relatively temporary states resulting from reactions to events or participants, such as anger, happiness, sadness, worry, etc.

(2) *I **feel** sad at the thought of the journey.*

*Ich **empfinde** Wehmut vor der weiten Reise.*

***Siento** una gran tristeza antes del viaje.*

Emotions were distinguished from **attitudes**, which are more stable than emotions and typically directed towards other people rather than events, such as respect, love, contempt.

(3) *'Yes, I am,' replied Ivan glumly and for the first time he suddenly **felt** an inexplicable revulsion to poetry.*

*"Ja, Lyriker", antwortete Iwan finster und **spürte** zum erstenmal einen unerklärlichen Widerwillen gegen alle Lyrik*

*"Jsem," odpověděl pochmurně Bezprizorný a poprvé náhle **pocítil** nevysvětlitelný odpor k poezii.*

*Poèt, - mračno otvetil Ivan i vpervye vdrug **počuvstvoval** kakoe-to neob"jasnimoe otvraščenie k poèzii*

*Poète, oui, répondit sombrement Ivan, qui **ressentit** tout à coup, pour la première fois de sa vie, un inexplicable dégoût pour la poésie.*

Another class close to emotions but involving a more rational assessment has to do with experienter's confidence and their **certainty** in their own security.

(4) *Alice began to **feel** very uneasy.*

*Alice fing an sich sehr unbehaglich zu **fühlen**.*

*Alice commençait à se **sentir** très inquiète.*

*Alicia empezó a **sentirse** incómoda.*

Self-assessment is related to experiencer's view of how they look like or are perceived by others. An important difference from the other categories is that this perception doesn't have to hold in reality, and is different from emotions and certainty in this respect.⁶

(5) *You make me **feel** uncivilized, Daisy.*

*Vedle tebe **si připadám** jako člověk bez civilisace, Daisy.*

*Dèzi, rjedom s toboj ja perestaju **čuvstvovat'** sebja civilizovannym čelovekom*

*Tú me haces **sentir** poco civilizado, Daisy.*

Exemplars describing **well-being** are concerned with the general state of the experiencer, often indeterminate between emotions and bodily sensations, e.g., 'feel well', 'feel fine', etc.

(6) *'He **felt** better at once,' said Gandalf.*

*"Er **fühlte** sich sofort besser", sagte Gandalf.*

*On **počuvstoval** sebja lučše, - skazal Gèndalf.*

*Il s' est aussitôt **sentit** mieux , dit Gandalf .*

*- Se **sintió** mejor inmediatamente - contestó Gandalf.*

Bodily sensations include physiological processes and reactions, such as pain, hunger, ache, heartbeat, etc.

(7) *And Kuzmin **felt** a chill run up his spine.*

*A **cítil**, že mu na čele vyskakuje studený pot.*

*on **počuvstoval**, kak u nego poxolodel zatylok*

*Et Kouzmine **sentit** un frisson courir dans son dos.*

*y Kusmín **sintió** frío en la nuca.*

Perception of **touch** is close to bodily sensations but the former has to do with the contact to the skin perceived by the experiencer.

(8) *"Fawkes?" Harry breathed, and he **felt** the bird 's golden claws squeeze his shoulder gently.*

⁶ For instance, if one feels stupid this doesn't mean that one is stupid, whereas if one feels happy it is possible to say that one is happy.

"To jsi ty Fawkesi?" vydechl Harry a **cítil**, jak mu pták svými zlatými drápy jemně stiskl rameno.

- Fouks? – prošeptal Garri i **počuvstvoval**, kak zolotye kogti nežno szali emu plečo

- Fumseck? murmura Harry. Il **sentit** les serres de l' oiseau presser doucement son épaule.

- ¿Fawkes? - musitó Harry, **sintiendo** la suave presión de las garras doradas.

A separate class was established for rare exemplars describing **agentive touch**, e.g., situations where the experiencer consciously directs their body part to explore something. In all the nine exemplars of this class in the sample, the verb *feel* is used in English, and two of them feature the verb *fühlen* in German. In all other languages, other verbs are used to render the situations of this class.

(9) "**Feel** mine then and tell me if I have a fever."

"Dann **fühle** doch bitte, ob ich vielleicht Fieber habe!"

A small yet semantically distinct class encompasses exemplars where the experiencer perceives the **movement** of their own body, which is typically uncontrolled.

(10) Once again I **felt** myself frozen by the sense of something irreparable.

De nouveau je me **sentis** glacé par le sentiment de l' irréparable.

A number of exemplars were identified as describing **environment**, i.e., the perception of something not directly through skin but in the surrounding space, e.g., presence of somebody or something.

(11) To his right, the third animal in line raised his head, **sensing** our presence, and whinnied.

Kůň napravo od něho, třetí v řadě, zvedl hlavu a zaržál, asi **vycítil** naši přítomnost.

tretij v tom že rjadu kon' podnjál golovu, **učujav** nas s Vilgel'mom, i zaržal

le troisième animal de la rangée leva la tête comme il **sentait** notre présence, et il hennit.

Although **smells** can also be regarded as a part of environment, they were treated separately, because they are included in the perception paradigm, and it was interesting to assess the contribution to the traditionally identified sense modalities to the overall range of use of 'feel'-verbs.

(12) *Sotva mne **ucítíl**, zaštěkal.*

***Počujav** menja, sobaka otryvisto zalajala*

*Al **sentir** mi olor, el perro lanzó un ladrido.*

‘At the scent of me he barked shortly.’

Hearing (13) and **sight** were also included in the classification, although, as Table 2 shows, they are much less frequent than touch and smell.

(13) *Wenn die Menschen Krieg führen, dann bekommt auch die Weltenseele die Kampfrufe zu **spüren**.*

*Když muži válčí, i Duše světa **cítí** křik bitev.*

*Cuando los hombres están en guerra, el Alma del Mundo también **siente** los gritos de combate.*

‘When men are at war with one another, the Soul of the World can hear the screams of battle.’

Two distinct classes include exemplars describing perception of someone’s **gaze** directed at the experiencer and **another person’s emotion**.

(14) *Ich **fühlte**, dass er sich mehr und mehr erwärmte:*

***Cítíl** jsem, jak se mu krev pomalu vrací do žil.*

*Je **sentais** qu’ il se réchauffait peu à peu:*

***Sentía** que volvía a entrar en calor poco a poco:*

‘I could see that he was reviving little by little.’

In exemplars that belong to the domain of **cognition** ‘feel’-verbs describe experiencer’s inference based on perception. The inferred situation doesn’t have to be simultaneous with the perception event.

(15) *Ich **fühlte** wohl, dass etwas Außergewöhnliches vorging.*

*Dobře jsem **cítíl**, že se děje něco neobyčejného.*

*Je **sentais** bien qu’ il se passait quelque chose d’extraordinaire.*

‘I realized clearly that something extraordinary was happening.’

Two semantic classes have to do with modality, specifically with the perception of experiencer's own desires (**wish**) or obligations (**necessity**), as in (16).

(16) *'Do you **feel** any need to leave the Shire now - now that your wish to see them has come true already?' he asked.*

"Cítíš ještě potřebu odcházet z Kraje - teď když se ti splnilo přání vidět je?" zeptal se.

*- ¿**Sientes** aún la necesidad de abandonar la Comarca , ahora que cumpliste tu deseo de ver a los elfos? - le preguntó.*

Finally, in the exemplars included in the class “**affect**” the ‘feel’-verb describes the fact of impact or perception of something, without specifying how, through which means or in what way the experiencer will be affected.

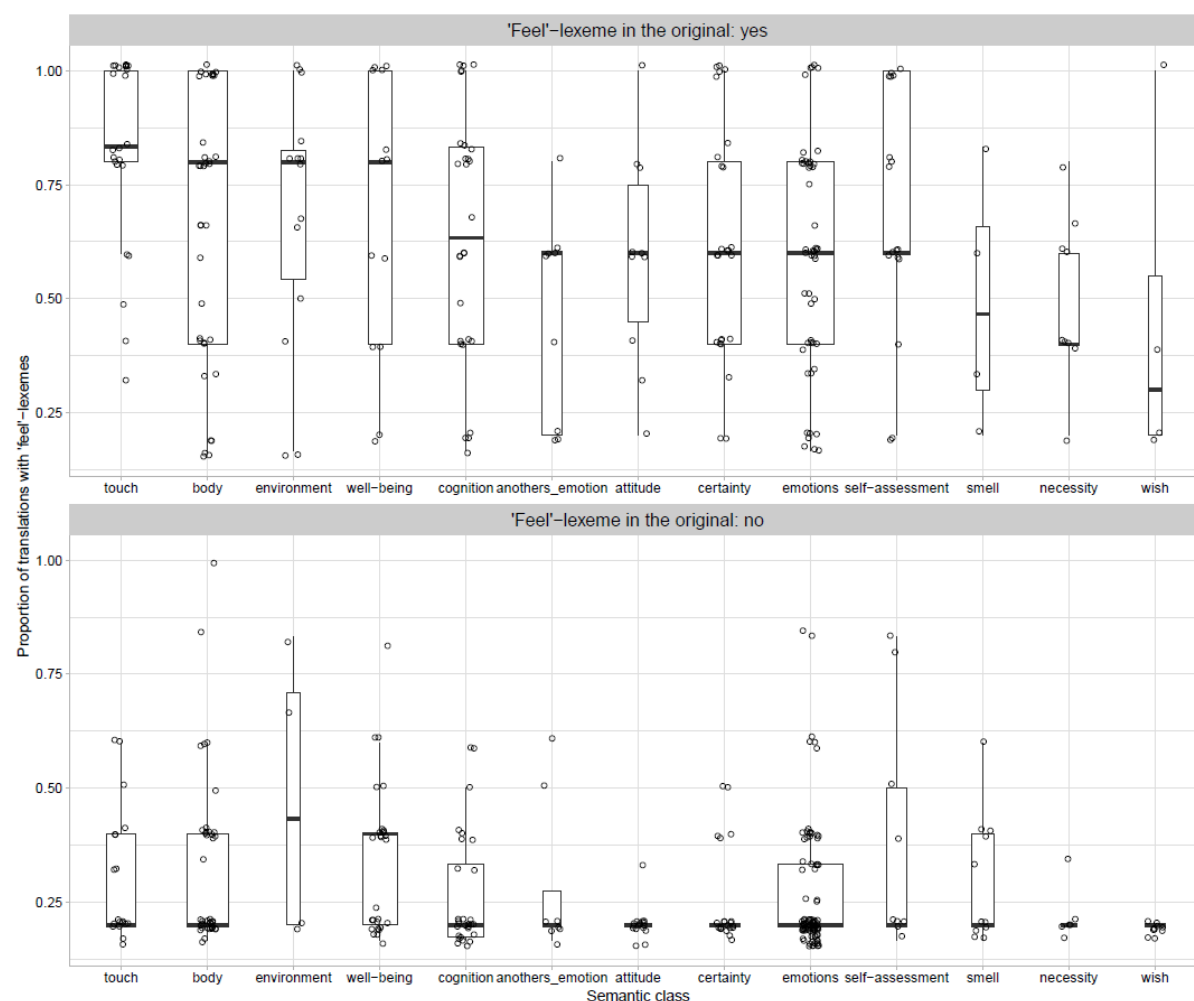
(17) *Celé hejno může reagovat i na podněty, které vyvolávají "pozitivní taxi", dokonce i když je **pocítí** jen jedno individuum.*

‘Stimulus situations which attract the fish can be responded to by a whole shoal, even when only one individual has received the stimuli.’

To sum up, emotions are by far the most frequently attested semantic class in the sample. The other widely attested classes are bodily sensations, cognition, certainty and well-being. These classes can be expected to constitute the largest proportions of token frequency of ‘feel’-verbs across the languages under analysis and probably other related European languages.

However, an exemplar where a ‘feel’-verb is used in the sentence in one language does not necessarily contain ‘feel’-verbs in the other languages. Semantic classes differ with respect to how consistently ‘feel’-verbs are used to describe them across languages. Additionally, as shown in section 2 above, if a ‘feel’-verb is present in the original sentence, it is likely to be used in translations, so this parameter should be controlled for. Figure 2 shows the distribution of the proportions of translations with ‘feel’-verbs for each of the semantic classes separately for the subsets of exemplars with and without a ‘feel’-verb in the original. The semantic classes are ordered by the median proportion for the first of the two subsets, as the differences in this subset are more pronounced. The width of the boxes reflects the number of observations. Only classes containing at least 15 exemplars are displayed.

Figure 2. Proportion of translations with ‘feel’-verbs depending on the semantic class and the presence of a ‘feel’-verb in the original



The upper part of Figure 2 shows that when a ‘feel’-verb is used in the original sentence, in case of touch, bodily sensations, perception related to environment, and general well-being, we observe ‘feel’-verbs across a larger number of translations than in case of other semantic classes. Specifically, emotions are not consistently rendered by ‘feel’-verbs across translations. The semantic classes that are least consistently rendered by ‘feel’-verbs are smells, necessity, and wish. In the lower part of Figure 2, which shows the cases where there is no ‘feel’-verb in the original, the semantic classes are less differentiated. Still, the median values for exemplars describing environment and well-being are higher than for the other classes.

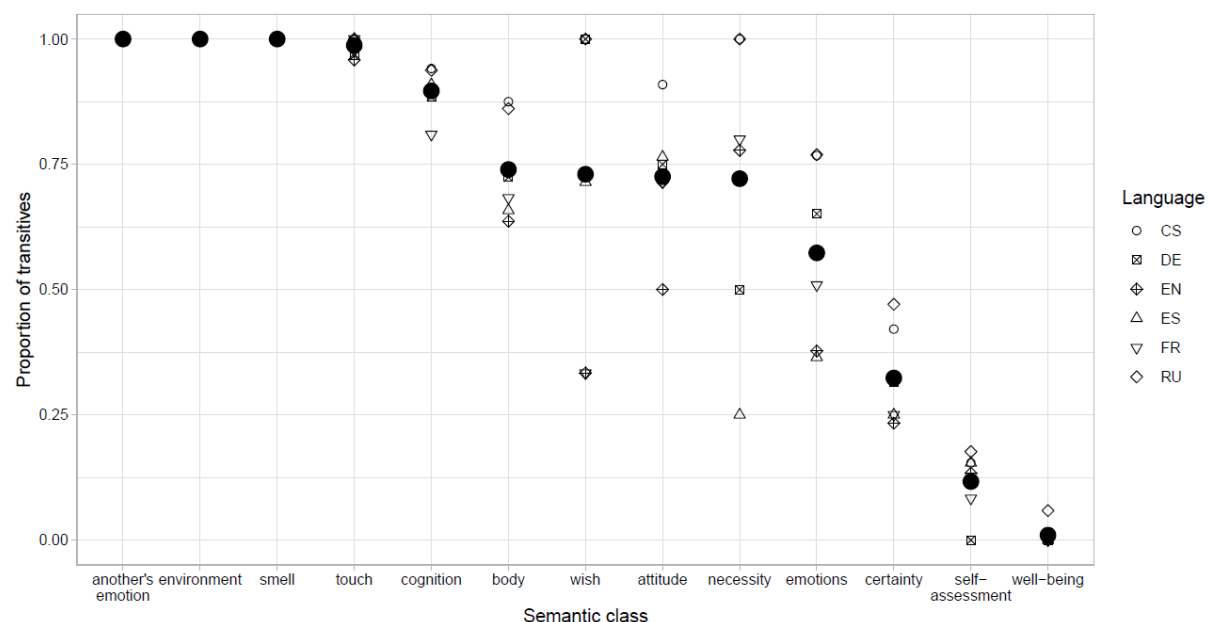
4. Constructions with ‘feel’-verbs and semantic classes

Constructions featuring ‘feel’-verbs can be classified and annotated with a different degree of granularity. Both too broad and too fine-grained classes can make the grouping of exemplars in terms of grammatical makeup and establishing their associations with semantic classes

problematic. As a result, two-layer annotation was implemented, where the first layer captures the general makeup of the construction and the second annotates the expression of the stimulus participant in a more detailed way, where possible.

In the first layer of annotation, I distinguished between transitive and intransitive verbal constructions, and possessive constructions featuring nouns with the meaning ‘feeling’. The two most frequent types are transitive and intransitive verbal constructions. Transitive constructions encompass both constructions with nominal and pronominal direct objects and constructions with complement clauses. Intransitive verbal constructions include reflexives in all the languages except English, as well as rare absolutive uses and uses with prepositional objects, as in *feel for the knife*. Figure 3 shows the proportions of transitive uses to all verbal (i.e., transitive and intransitive) constructions for the 13 most frequent semantic classes. The classes are ordered by the mean proportion of transitive constructions, indicated by solid circles. Points without fill indicate proportions for individual languages.

Figure 3. Proportion of transitive constructions among all verbal constructions for semantic classes



The highest mean proportions of transitive constructions on the left side of Figure 3 are observed for clauses describing perception of emotions of other people, environment, smells, touch, and cognitive inference based on the perceived. All these semantic classes have to do with stimuli of perception that are external to the experiencer. On the opposite side of Figure 3, the semantic classes predominantly expressed by intransitive constructions are related to the

experiencer's safety and self-confidence (certainty), to the assessment of one's own status and appearance and to general well-being. These semantic classes concern the experiencer's perception of oneself. Finally, in between we find the semantic classes of bodily sensations, wish, attitudes, necessity, and emotions. These semantic classes also show higher variation in the proportion of transitives across languages. Semantically, they are internal to experiencer but at the same time can be viewed as independent from and affecting experiencer, similarly to external stimuli discussed above. Of these intermediate semantic classes, emotions are somewhat closer to the classes on the right, related to the experiencer's perception of oneself. Thus, somewhat simplifying, emotions can be viewed either as a state of the experiencer as a whole and thus expressed as an adjective in an intransitive construction or presented as an independent entity perceived by the experiencer and expressed as a direct object. In (18), the former strategy is observed in German, French, and Spanish, and the latter, in the Czech and Russian translations ('feel'-verbs are in bold).

(18) *Aber anstatt traurig zu sein, **fühlte** er **sich** glücklich.*

*Ale místo smutku **pocítil** radost.*

*no vmesto privyčnoj uže grusti **ispytal** priliv štastja*

*Au lieu d' en éprouver de la tristesse, il se **sentit** heureux.*

*En vez de sentirse triste, se **sintió** feliz.*

'But instead of being saddened, he was happy.' (Coelho, "The Alchemist")

As mentioned above, within the broad classes of transitive and intransitive constructions further subtypes can be distinguished. Specifically, transitive constructions encompass not only the uses with ordinary direct objects, but also cases where the stimulus of perception is expressed by a complement clause. Complement clauses fall into several further subtypes, and they also show some discernible associations with semantic classes. These subtypes include a) clauses with the complementizers English *that*, German *dass*, Czech *že*, Russian *что*, French and Spanish *que*, b) clauses with the complementizers German *wie*, Czech *jak*, and Russian *kak*, c) raising constructions, and d) asyndetic constructions, which are mostly attested in English (19). In (20), the English, Czech, Russian, and Spanish sentences belong to the same type, featuring the complementizers *that*, *že*, *что*, and *que*, respectively, the German sentence with the complementizer *wie* belongs to the type b, and the French construction was classified as raising.

(19) *Arthur blinked at the screens and **felt** he was missing something important.*

(20) *EN he **felt** that the forest was awake below him – awake and alive and crowded.*

*DE er **fühlte**, wie der Wald unter ihm voll drängenden Lebens war.*

*CS a **cítíl**, že dole pod ním se les probudil – probudil a naplnil životem.*

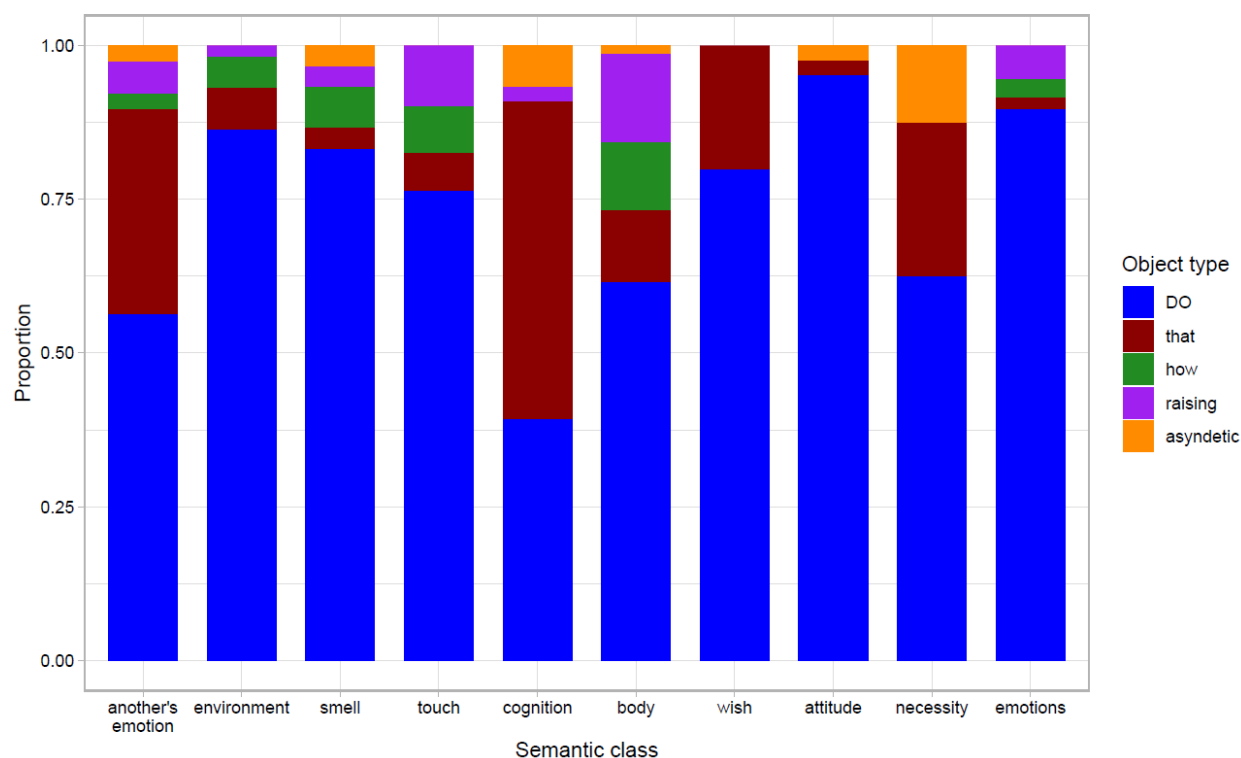
*RU **počuvstvoval**, čto vnizu, pod nim, les prosnulsja, ožil i napolnilsja kakimi-to suščestvami*

*FR et il **sentit** la forêt éveillée au-dessous de lui, éveillée, vivante et pleine d' êtres.*

*ES y **sintió** que la selva, a sus pies, estaba despierta, despierta, viva y habitada.*

Table 3 shows the frequencies of the subtypes of transitive constructions for the ten semantic classes with a higher overall proportion of transitives, in the same order as in Figure 3. Since the particular types of complement clauses are not very frequent, in this case I use aggregate frequencies, i.e., the frequencies of the construction types are summed up across languages (the distributions in individual languages are given in the supplementary materials).

Figure 3. Distribution of object types for the semantic classes associated with transitive constructions



The semantic class for which constructions with complement clauses make up the largest proportion is unsurprisingly cognition. The predominant type of stimulus expression for this

semantic class is a complement clause with ‘that’-type complementizer; asyndetic constructions are also attested as the second most frequent clausal complementation strategy for this semantic class. The other semantic classes with a considerable proportion of clauses with ‘that’-type complementizers are another’s emotion, necessity, and wish. Presumably, this grouping reflects a more rational nature of experiences belonging to these four semantic classes. The largest proportions of constructions with ‘how’-type complementizers and raising constructions are observed for the semantic classes that have to do with bodily sensations and touch. Both classes include immediately perceived sensations concerned with experiencer’s body. The other semantic classes are less frequently described by constructions with complement clauses.

Among intransitive constructions, the most frequent type is when the state of the experiencer is expressed as an adjective, as in (21). This construction type was distinguished from those cases when the state of the experiencer is described by an adverb, i.e., the item that is used for verbal rather than nominal modification (22). Finally, another type attested across languages is when the state of the experiencer is expressed by a noun, as in (23).⁷

(21) *"Yes," said Arthur, glaring at him and trying not to **feel** foolish.*

»Ja«, sagte Arthur, starrte ihn an und versuchte, **sich** nicht dämlich **vorzukommen**.

— Oui, dit Arthur, l’œil enflammé, en essayant de ne pas se **sentir** ridicule.

- Sí - dijo Arthur, lanzándole una mirada iracunda y tratando de no **sentirse** ridículo.

(22) *It never occurred to me that it was a hallucination because I **felt** perfectly well.*

*An eine Halluzination habe ich nicht gedacht, denn ich **fühlte** mich völlig wohl.*

*Halucinace to nebyly, protože jsem se **cítil** úplně dobře.*

*O gallucinacii ja ne dumal, tak kak **čuvstvoval** sebja sovsem xorošo.*

*Je n’ai pas pensé qu’il s’agissait d’une hallucination, parce que je me **sentais** très bien.*

(23) *The embankment lights came on and I **felt** like the audience in a theatre*

*Am Ufer gingen die Laternen an und ich **hatte das Gefühl** eines Zuschauers*

*Na nábreží se rozsvítily lampy a já jsem **měl pocit** diváka*

*Na naberežnoj zažglis' fonari, i ja **čuvstvoval** sebja zritelem v teatre*

⁷ English shows a number of intransitive constructions unattested in the other languages, such as the constructions *feel for* and *feel at*, used for agentive touch, for which the use of ‘feel’-verbs is generally mostly restricted to English. The construction *feel like* with the gerund is also a way to describe wish specific to English.

*Sur le quai, les lampadaires se sont allumés et j'ai **eu le sentiment** d'être un spectateur*
*Se encendieron las farolas del muelle y yo me **sentía** como un espectador*

Figure 4 shows the proportions of these construction types based on aggregate frequencies across languages, similarly to Figure 3. Only the four semantic classes with the highest proportions of intransitive uses are included, see Figure 2. Note that emotions are shown both in Figure 3 and Figure 4.

Figure 4. Distribution of intransitive constructions for the semantic classes associated with them

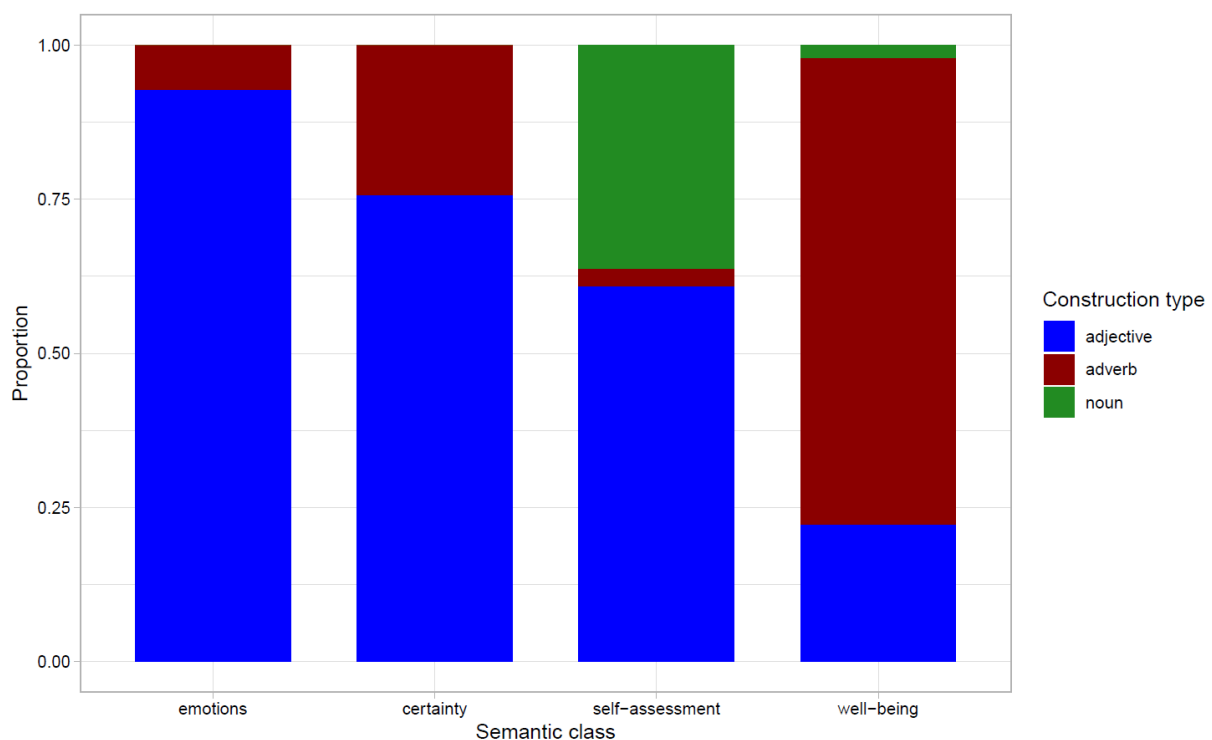


Figure 4 shows that in intransitive constructions describing emotions and certainty the state of the experiencer is mostly described by adjectives. Intransitive constructions with nouns are mostly attested in clauses describing self-assessment, and adverbs are strongly associated with the semantic class of well-being.

One final word should be said about possessive constructions with the noun ‘feeling’ (24). Interestingly, these constructions are clearly associated with the semantic class of cognition. Almost a half (0.4) of the total number of these constructions are used in clauses describing inference on the part of the experiencer, as in (24), whereas in general this semantic class makes up about one tenth of the sample.

(24) *Dnes večer **měl** Silas konečně **pocít**, že začal splácet svůj dluh.*

*Cette nuit Silas **avait** enfin **le sentiment** de commencer à rembourser sa dette.*

‘Tonight, at last, Silas felt he had begun to repay his debt.’

This section showed that there is a considerable diversity of constructions used with ‘feel’-verbs. At least for the more frequent semantic classes, differences in constructional profiles, or the frequency distributions of constructions, can be established. In the next section, I discuss how lexical distinctions drawn in individual languages map onto the semantic classes and construction types.

5. The structure of the ‘feel’-domain: constructions and verbs

In this section, I explore the structure of the ‘feel’-domain combining information on constructions featuring ‘feel’-verbs with lexical dimension. The goal is to examine the overall structure of the domain and to explore how the lexical distinctions drawn in the languages under analysis carve up the domain and whether there are any cross-linguistic regularities in its structuring.

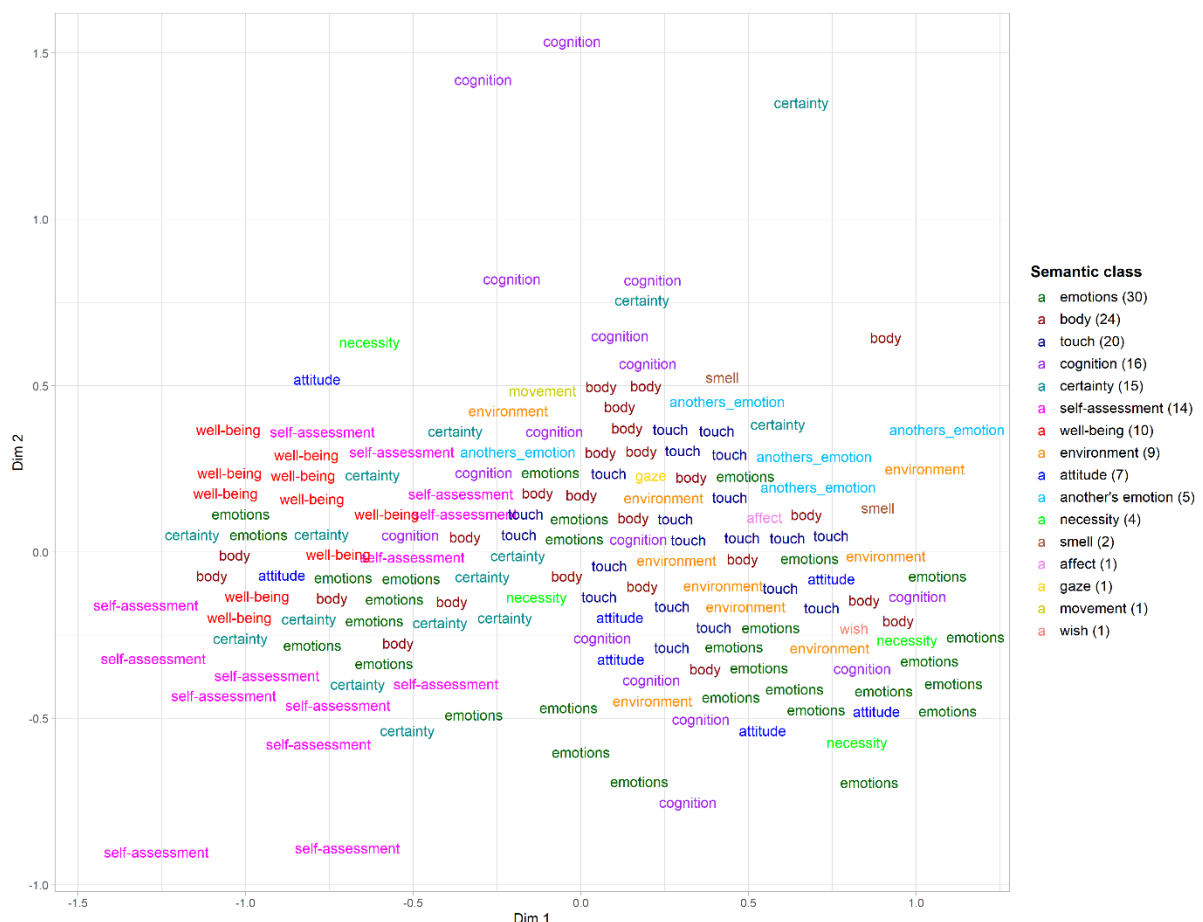
The analysis undertaken in this section is similar to other studies on lexical typology based on parallel corpora in the choice of methods. These include calculating distances between exemplars based on the similarity of linguistic expression, visualizing these distances to find groups in the data and exploring the semantic motivation for these groups, see Cysouw, Wälchli (2012), Levshina (2015). In these studies, the distances used to investigate lexical distinctions are calculated based on the similarity of lexical expression of exemplars, and they are interpreted using a metalanguage, typically based on English lexical distinctions. In contrast to these studies, I use both lexical and grammatical features to calculate distances and explore the semantic motivation for the grouping of exemplars mainly by comparing them to the set of semantic classes discussed in the previous sections.

Technically, the data were analysed as follows. First, I created a subset of exemplars in which ‘feel’-verbs were used in at least four of the six parallel sentences. This was done to ensure a more reliable measure of similarity between exemplars. This subset includes 160 exemplars. Each of the sentences with ‘feel’-verbs is annotated for the verb itself, and the two layers of annotation for construction types, as discussed in section 4. When a sentence in language doesn’t contain a ‘feel’-verb, all three annotation cells have NA’s in them. Distances between two exemplars were calculated based on the proportion of (mis)matches in the annotation values among all the valid (non-NA) cells. The match in lexical expression was

counted as 1, whereas the match in each of the values pertaining to constructions was counted as 0.5. This way the contributions of lexical and constructional annotations were made equal.

The distances were visualized using Multidimensional scaling algorithm, which is a widely used method of visualizing distances in a low-dimensionality space, aiming at their minimal distortion. In this case, I used a non-metric version of MDS, which takes into account the ranking of distances between objects rather than the similarity values as such. The 2-dimensional MDS solution, which was created using the function *smacofSym* of the package *smacof* (Mair et al. 2022), has a stress value of 0.16, which is a tolerable degree of distortion. Figure 5 shows the resulting visualization, where individual exemplars are plotted using the labels of the semantic classes they belong to, and the semantic classes are distinguished by colours, so as to better show their groupings on the plot. The legend gives the list of semantic classes and their frequencies in this subset of data.

Figure 5. MDS-visualization of the distances between exemplars based on verbs and constructions



...napomnila Sofi. Očividno, ona **oščutila**, kak naprjažěn Lèngdon.

Another area of the plots where non-basic ‘feel’-verbs are found in a number of the languages is the bottom right part with exemplars describing emotions and attitude. In English, the three exemplars featuring the verb *experience* is found in this area, in German this area is covered by the verb *empfinden*, in Czech there is a variety of verbs populating this area, in Russian, the majority of sentences with the verb *ispytyvat'* are found in this area, and in French, the verb *éprouver*.

German *sich vorkommen* and Czech *připadat si* are the verbs mostly used for self-assessment. Finally, as already discussed in section 4, possessive constructions with the verb ‘have’ and the noun ‘feeling’, although generally infrequent in the sample, are associated with cognition. Specifically, several exemplars of this semantic class found in the top parts of the plots are expressed by the expressions *have a feeling* in English, *mít pocit* in Czech, and *avoir impression/sentiment* in French.

Thus, although the associations between ‘feel’-verbs and semantic classes are not perfect, there may be some recurrent patterns in the lexical structuring of the ‘feel’-domain. One of such patterns is the existence of the verb dedicated to the expression of external stimuli, such as environment and touch.

In general, lexical distinctions appear to be more language-specific and consistent across languages than the distribution of constructions. Specifically, the degree of average mutual predictability between the distributions of lexical expressions in the pairs of languages is much lower than that for the distribution of constructions, either of the first or of the second, more detailed, layer of annotation. This was measured using normalized mutual information (NMI) with the help of the R package *aricode* (Chiquet et al. 2023). The average NMI for the distributions of lexical expressions is 0.15, whereas for the distributions of constructions it is 0.45 for the first layer of annotation, capturing broader distinctions between transitives and intransitives, and 0.48 for the second layer of annotation, capturing the expression of the stimulus. Thus, constructions are more uniform across languages than lexical distinctions.

The association with semantic classes, again measured using NMI, is also more straightforward for constructions than for the lexical means, although in this case there are some differences between the languages and the two layers of annotation for constructions. The average values of NMI are 0.11 for lexical items, 0.12 for the annotation of constructions with respect to broader types, and 0.25 for the annotation of constructions with respect to the more specific subtypes of stimulus expression, such as ‘that’-complement clauses for

transitives or adjectival secondary predicates for intransitives. The fact that the latter layer of annotation indeed shows strong associations with semantic classes was already hinted at in section 4. The NMI values for the mutual predictability between semantic classes and lexical distinctions are higher for German, Czech, and Russian (0.13, 0.18, and 0.17, respectively) than for the other three languages. Presumably, the reason for this is that in the former three languages, the frequency distribution of basic vs. other ‘feel’-verbs is more balanced and the non-basic ‘feel’-verbs are not so marginal. Thus, when non-basic ‘feel’-verbs are prominent enough, they show some correlation with the semantic classes identified in this study, albeit not a very strong one.

6. Summary and conclusions

For the traditional paradigmatic view on the domain of perception, ‘feel’-verbs were relevant as long as they appeared to be the basic way of expression for one or more of the five sense modalities included in the paradigm. These modalities usually include touch and smell and sometimes hearing. The polysemy of verbs with the meaning ‘feel’ was to a larger extent recognized in the studies on colexification, where their connections to the domains of emotions and cognition were established in an empirical way. The present study aimed at a more comprehensive exploration of the semantic range of ‘feel’-verbs, although limited to a number of well-known European languages.

Based on the analysis of the sample of parallel sentences, I proposed a detailed semantic classification of situations covered by ‘feel’-verbs. Along with the classes already mentioned in connection with ‘feel’-verbs in the literature, such as emotions, touch, and cognition, some semantic classes were identified that are relevant for the ‘feel’-domain but have not been discussed in the literature. These classes include well-being, which has to do with a general state of the experiencer, intermediate between emotions and bodily sensations, self-assessment, which is related to experiencer’s assessment of how they may be perceived from the outside, and certainty, which concerns experiencer’s self-confidence and safety. It is noteworthy that many of the sensation types not included in the traditional five-sense model are found among the situation types described by ‘feel’-verbs. These sensations stem from “physiological systems for sensing pain, temperature, balance, and our body’s location and movements (proprioception)” (Norcliffe, Majid 2024, 84, see also Winter 2019, 12-14). In this study, these types of perception were included in the semantic classes of bodily sensations, environment, and movement.

Emotions are by far the most frequent semantic class in the sample, followed by bodily sensations and cognition. At the same time, emotions and cognition are not among the classes which are consistently expressed by ‘feel’-verbs across the languages of the sample. This means that although the situations of these classes may be expressed by ‘feel’-verbs, there are other, probably more central means of describing them, such as dedicated emotion verbs and adjectives and cognition verbs. Semantic classes for which ‘feel’-verbs are more consistently employed across languages are touch, environment, bodily sensations, and well-being.

The distribution of constructions featuring ‘feel’-verbs suggests a broad division of semantic classes into two opposed poles and an intermediate zone in between. Semantic classes that are associated with transitive constructions semantically have to do with perception of an external stimulus, such as another’s emotion, environment, touch, and cognition. The semantic classes concerned with the experiencer’s perception of oneself, such as well-being and certainty, are more often described by intransitive constructions. Intermediate classes include bodily sensations, perception of necessity and wish, attitude and emotions. On a more fine-grained level, the classes within these groups are further distinguished by how the stimulus is expressed, so that the semantic classes exhibit a high degree of differentiation with respect to constructions.

Lexical distinctions are less systematic across languages, and their associations with semantic classes are far from clear-cut. What the examined lexical distributions highlight is that there is a basic ‘feel’-verb that can be used to describe situations belonging to all of the identified semantic classes, cutting across the domains of perception, emotion, cognition, and other experiential types that are typically kept separate in linguistic descriptions.

One of the more consistent semantic divisions suggested by lexical distinctions in some of the languages under analysis is that there is a verb mostly associated with external stimuli and not employed for the semantic classes concerns with experiencer’s perception of oneself. This is how the German verb *spüren*, the Czech verbs *ucítit* and *vycítit*, and the Russian verbs *oščuščat’* and *oščutit’* are opposed to the basic ‘feel’-verbs (a similar distinction can also be found in some other Slavic languages, see Ovsjannikova, in press). Although this may partially be an areal effect, Evans and Wilkins (2000, 554) report that in many Australian languages, there is a lexical contrast between the verb used for internal bodily sensations and the verb used for external touch. Thus, it may be hypothesized that perception of external stimuli is the domain where lexical renewal in the ‘feel’-domain is more likely to start, whereby the basic ‘feel’-verb is gradually replaced by a verb dedicated to expressing external stimuli.

References

- Becker, L., & Guzmán Naranjo, M. (2020). Psych predicates in European languages: A parallel corpus study. *STUF – Language Typology and Universals*, 73(4), 483-523. <https://doi.org/10.1515/stuf-2020-1017>.
- Chiquet J., Rigai G., Sundqvist M. (2023). *_aricode: Efficient Computations of Standard Clustering Comparison Measures_*. doi:10.32614/CRAN.package.aricode <<https://doi.org/10.32614/CRAN.package.aricode>>, R package version 1.0.3, <<https://CRAN.R-project.org/package=aricode>>.
- Cysouw, Michael, and Bernhard Wälchli. 2012. Lexical typology through similarity semantics: Toward a semantic map of motion verbs. *Linguistics*, 50(3), 671–710. DOI 10.1515/ling-2012-0021.
- Dickey, S. M. (2000). *Parameters of Slavic Aspect. A cognitive approach*. Stanford, CACSLI Publications.
- Evans, Nicholas & David Wilkins. 2000. In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76.3, 546–592. <https://doi.org/10.2307/417135>.
- Georgakopoulos, Th., Grossman, E., Nikolaev, D., Polis, S. (2021). Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain, *Linguistic Typology*, 26(2), 439–487. <https://doi.org/10.1515/lingty-2021-2088>
- Levshina, Natalia. 2015. How Europeans GIVE: A two-layered semantic typology based on two parallel corpora. In: B. Nolan, G. Rawoens, & E. Diedrichsen (Eds.). *Causation, Permission, and Transfer: Argument realisation in GET, TAKE, PUT, GIVE and LET verbs*, 147–175. John Benjamins. <https://doi.org/10.1075/slcs.167.05lev>.
- Levshina, Natalia. 2017. Online film subtitles as a corpus: An n-gram approach. *Corpora*, 12(3), 311–338. DOI: 10.3366/cor.2017.0123.
- Mair P., Groenen P., de Leeuw J. (2022). Multidimensional Scaling Using Majorization: SMACOF in R. *Journal of Statistical Software*, 102. doi:10.18637/jss.v102.i10 <<https://doi.org/10.18637/jss.v102.i10>>.
- Norcliffe, Elisabeth, Asifa Majid. 2024. Verbs of perception: A quantitative typological study. *Language*, Vol. 100(1), 81-123. <https://doi.org/10.1353/lan.2024.a922000>.
- Raatikainen, Oona. 2021. Colexification patterns of perception verbs in the Circum-Baltic Area. Master's Thesis, University of Helsinki.

- Revelle, W. (2025). *_psych: Procedures for Psychological, Psychometric, and Personality Research_*. Northwestern University, Evanston, Illinois. R package version 2.5.6, <<https://CRAN.R-project.org/package=psych>>.
- 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.
- Rzyski, Christoph and Tresoldi, Tiago et al. 2019. The Database of Cross-Linguistic Colexifications, reproducible analysis of cross- linguistic polysemies. DOI: [10.1038/s41597-019-0341-x](https://doi.org/10.1038/s41597-019-0341-x)
- Viberg, Å. (1984). The verbs of perception: a typological study, *Linguistics*, 21.1, 123-162.
- Viberg, Å. (2001). Verbs of perception. In: M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language Typology and Language Universals: An International Handbook*, pp. 1294–1309. Berlin: Walter de Gruyter.
- Wälchli, B. (2016). Non-specific, specific and obscured perception verbs in Baltic languages, *Baltic Linguistics*, 7, 53–135.
- Winter, Bodo. 2019. *Sensory Linguistics: Language, Perception and Metaphor*. Amsterdam: John Benjamins Publishing Company.