



# How morphologically related synonyms come to make up a paradigm

## Как морфологически связанные синонимы формируют парадигму

Björn Wiemer<sup>1</sup> · Joanna Wrzesień-Kwiatkowska<sup>1</sup> ·  
Piotr Wyroślak<sup>2</sup>

Accepted: 15 June 2020 / Published online: 29 October 2020  
© The Author(s) 2020

**Abstract** The Slavic perfective (pfv): imperfective (ipfv) opposition is based on stem derivation. It creates a complex network of functions for finite and non-finite forms, which largely applies regardless of aspectual pairedness (and actionality groups), since this opposition has classificatory properties. However, can derivationally related stems claimed to represent identical lexical concepts be treated as representatives of one or of two paradigms? The issue becomes especially intricate with aspect triplets in which two ipfv stems correspond to one pfv stem, as though combining two productive patterns of aspect derivation. On this background, we test some core assumptions of the morphology-lexicon interface on one typical aspect triplet from Polish and Czech, the cognate ipfv Pol. *dzielić – rozdzielać*, Cz. *dělit – rozdělovat* ‘divide, separate’. We provide their token-based analysis for the period 1750–2017. The two ipfv stems show preferences for different basic functions associated with the ipfv aspect, the coding of marginal arguments and adjuncts also yields clear biases of choice. These preferences prove stable over time, distinctions in form typically associated with inflection turn out to be altogether irrelevant. Our findings, as well as a revision of theoretical positions, support a notion of paradigm in which typical inflectional distinctions are brought into an equilibrium with functional inventories and collocational constraints.

**Аннотация** Славянская оппозиция совершенного (СВ): несовершенного (НСВ) вида строится на противопоставлении основ, одна из которых образована от другой. Так создается сложная система функций личных и неличных форм, которая в целом не

---

We gratefully acknowledge the funding of our research by the German Science Foundation (WI 1286/19-1) and the Polish Science Foundation (2016/23/G/HS2/00922), programme *Beethoven II*.

---

✉ B. Wiemer  
[wiemerb@uni-mainz.de](mailto:wiemerb@uni-mainz.de)

J. Wrzesień-Kwiatkowska  
[kwiatkow@uni-mainz.de](mailto:kwiatkow@uni-mainz.de)

P. Wyroślak  
[pi.wyroslak@gmail.com](mailto:pi.wyroslak@gmail.com)

<sup>1</sup> Johannes-Gutenberg-Universität, Mainz, Germany

<sup>2</sup> Uniwersytet im. Adama Mickiewicza, Poznań, Poland

зависит от видовой парности (и акциональных групп), поскольку ей присущи слово-классифицирующие свойства. Здесь возникает вопрос: представителями одной парадигмы или же двух следует считать морфологически связанные основы, выражающие идентичные лексические концепты? С особой остротой этот вопрос возникает при обращении к видовым тройкам, в которых с одной основой СВ соотнесены две основы НСВ, причем как бы переплетаются два продуктивных способа видообразования. В данном контексте мы проверяем ряд ключевых предположений на стыке морфологии и словаря, исследуя НСВ одной типичной видовой тройки в польском и чешском языках, а именно родственные пол. *dzielić – rozdzielać* и чеш. *dělit – rozdělovat* ‘делить, разделять’. Анализ корпусных вхождений в период с 1750 по 2017 г. указывает на заметное влияние различных канонических функций, приписываемых основам НСВ, а также разных периферийных аргументов и адьюнктов, на превалирование той или иной из двух основ НСВ. Это влияние стабильно во времени. Напротив, различие между формами, связанное со словоизменением, оказывается нерелевантным. Наши наблюдения вкупе с пересмотром ряда теоретических положений свидетельствуют в пользу такого понимания парадигмы, в рамках которого классические словоизменительные различия находятся в равновесии с функциональными инвентарями и сочетаемостными ограничениями.

## 1 Introduction

The opposition of perfective (pfv): imperfective (ipfv) aspect can be considered a ‘proprietary label’ of Slavic languages as a whole, inasmuch as in all contemporary and historically attested Slavic languages the core of their aspect system is organized in an identical manner, both concerning basic morphological patterns and basic functional distinctions. This does not imply that the system across Slavic languages is identical in all of its parts, but we can say that its fundamental architecture is the same for all Slavic languages. In contrast to many other languages with grammatical aspect, the distinctions borne by pfv vs ipfv aspect are based on patterns of stem derivation marked by prefixes and suffixes. One consequence is that aspect can be determined not only for finite, but also for non-finite forms (infinitives, participles, etc.); thus, choice of aspect, and its interactions with other morphological categories, with clausal syntax and semantics as well as with pragmatic distinctions like illocutionary force, cannot be avoided in practically any occurrence of a verb (stem); see Sect. 2. Another consequence of the stem-derivational character of Slavic aspect is a notorious problem with determining which morphologically related stems can be qualified as representing the same lexical concept (and be considered an aspect pair) and in which cases stem derivation leads to modified lexical meanings (which blocks aspectual pairedness). In other words: when does stem derivation lead to synonyms which differ only in their grammatical behaviour, and when does it create items which should be considered as coding different lexical concepts? Jointly with this, the question arises whether related stems with different grammatical behaviour should be organized into one paradigm, or whether each of them has its own paradigm, which however might be somehow united, or connected (see Sect. 3).

Therefore, Slavic aspect presents us with the problem of how morphologically related verb stems are organized into near synonyms, how these synonyms are distributed over grammatical contexts (which is why they are divided into pfv and ipfv stems) and how their paradigmatic structure can be adequately described. We want to show how these issues are interrelated and which insights about interfaces between morphology and lexicon are gained by a joint treatment. Apart from this general issue we present a usage-based case study on two

cognate verbs in Czech and Polish, Cz. *dělit* and Pol. *dzielić* ‘divide, separate; share’, which participate in so-called aspect triplets. This case study accounts for a large variety of factors, including a sequence of diachronic periods, as we take the time period from 1750 to 2018 into account. The study has been performed within the *DiAsPol250*-project<sup>1</sup> carried out in Mainz and Warsaw.

We proceed as follows. After some basics about the Slavic PFV : IPFV opposition (Sect. 2), problems connected with morphological theory and the lexicon will be discussed (Sect. 3). Subsequently, we will present the case study (Sect. 4). The final section (Sect. 5) contains conclusions and an outlook.

If not indicated otherwise, examples are from our samples (see Sect. 4), which were collected from the respective national corpora (NKJP for Polish, ČNK for Czech). Glosses are reduced to a minimum, namely to indicating the aspect of the relevant verbs.

## 2 The Slavic aspect system (PFV : IPFV opposition)

In Slavic languages, pfv and ipfv aspect are not indicated by unequivocal, monofunctional morphemes. Instead, aspect is built on the basis of the functional reinterpretation of patterns of stem derivation, for which both prefixes and suffixes are used (cf. Breu 2000; Wiemer 2008; Wiemer and Seržant 2017, among many others). That is, it is not forms of one lexeme that distinguish aspect, but lexemes (roughly, stems with an identical lexical concept) related to each other via patterns of morphological derivation employing prefixes and suffixes (see Fig. 1). These affixes can be distinguished from morphology typically associated with inflection and finiteness (person-number agreement, tense).<sup>2</sup> Since aspect is distinguished for stems, it also applies to non-finite forms. Compare the following Polish example in which the stem is taken in curly brackets. We can also see some of the systematic morphonological alternations at the end of the stem; these do not affect the paradigmatic unity:

(1)	Polish <i>przekazać</i> <sup>pfv</sup> – <i>przekazywać</i> <sup>ipfv</sup> ‘convey’		
	infinitive	{ <i>przekaz-a</i> }-ć	{ <i>przekaz-ywa</i> }-ć
	non-past, 1sg	{ <i>przekaż</i> }-ę	{ <i>przekaz-uj</i> }-ę
	past, F.2sg	{ <i>przekaz-a</i> }-t-a-ś	{ <i>przekaz-ywa</i> }-t-a-ś
	imperative	{ <i>przekaż</i> }-∅	{ <i>przekaz-uj</i> }-∅
	converb	{ <i>przekaz-a</i> }-wszy	{ <i>przekaz-uj</i> }-ąc
	agent-demoting participle,		
	NOM.SG.M	{ <i>przekaz-a</i> }-n-y	{ <i>przekaz-ywa</i> }-n-y
	action noun(neuter)	{ <i>przekaz-a</i> }-ni-e	{ <i>przekaz-ywa</i> }-ni-e

The stems jointly build a complex network of functions, starting from basic distinctions of actionality (event vs process / state) and event-external pluractionality (single vs repeated situation),<sup>3</sup> over functions related to modality, see (2), up to illocutionary force and presupposition management (3):

- (2) a. Nie mogłem *otworzyć*<sup>pfv</sup> drzwi (bo klamka była zepsuta).  
 ‘I couldn’t *open* the door (because the handle was broken).’

<sup>1</sup><http://www.diaspol.uw.edu.pl/eng/>.

<sup>2</sup>Stems can thus be defined as those parts of (often morphologically complex) word forms which exist prior to the application of person-number and tense morphology.

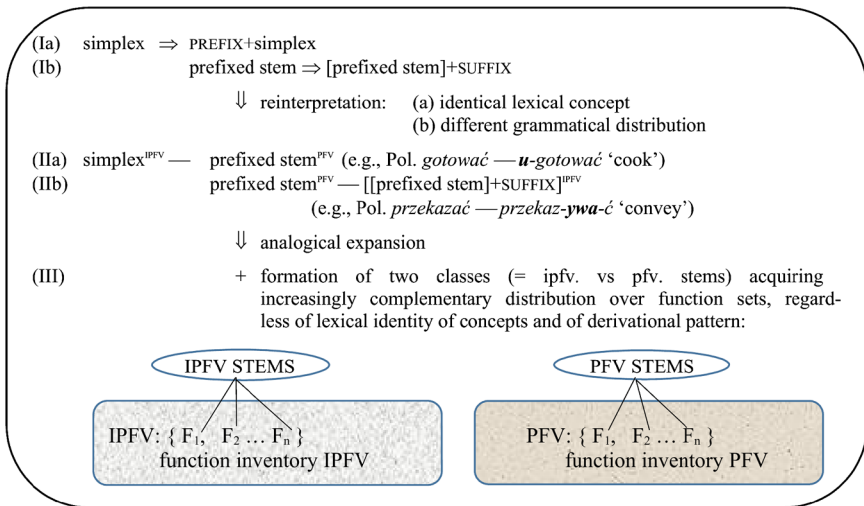
<sup>3</sup>For a classification of pluractionality types cf. Šluinskij (2006), Wood (2007), and Mattioli (2019).

- b. Nie mogłem *otwierać*<sup>iipfv</sup> drzwi (bo nie było wolno).  
 ‘I couldn’t *open* the door (because it was forbidden).’
- (3) a. *Opowiedz*<sup>ipfv</sup> (Jeśli chcesz).  
 ‘Tell us! (If you want.)’
- b. *Opowiadaj*<sup>iipfv</sup> (Bo czekamy i pękamy od ciekawości).  
 ‘Tell us! (We are waiting and very curious.)’

That is, the conditions that influence the choice of a verb stem result from the interaction with grammatical categories (e.g., interaction with tense, the imperative with or without negation, modals), from semantic oppositions (e.g., single situation vs unlimited repetition, or restrictions in the interpretation of converbs), or from pragmatic features (oppositions).<sup>4</sup> This becomes particularly obvious for those stems, which, related via salient morphological patterns, share their lexical meaning. Contrary to ordinary synonyms, such related stems are not randomly distributed over functions from these domains, but their use is constrained by categorial distinctions and functions that are admissible for either the pfv or ipfv stems; under particular conditions ipfv and pfv stems can (or even must) replace each other in order to convey an identical lexical concept (see Sect. 3.1). Correspondingly, the aforementioned conditions can be ordered by morphosyntactic formats, from word-form and predicational level (e.g., complex predicates with phasal or modal verbs) to clause level, clause-combining and discourse level. The strength of the particular restrictions varies, from obligatory choices (see trivial functions in Sect. 3.1) to more probabilistic distributions providing some leeway of interpretation.

From a diachronic viewpoint, we may assume that these conditions, and the functional domains behind them, successively became prominent as factors influencing the choice of pfv vs ipfv stems (cf. Wiemer 2008, pp. 393–406 for a rough outline). Since here we are not dealing with the long-term diachronic development of the aspect system, only one central remark has to be made. The aforementioned conditions on aspect choice are of a very heterogeneous nature, so that conflicts between particular conditions are inevitable. For instance, pfv verbs by definition mark a situation as limited, either because of its inherent character (with telic verbs) or because a temporal limit is indicated (with atelic verbs). On the other hand, ipfv verbs are usually the preferred choice when situations are presented as repeated an unlimited number of times. A conflict arises for situations that are repeated (without limits), but the boundaries of each particular instance may matter as well (as in *John goes to cinema every Friday after lunch* or *Harriet used to play table-tennis once a week*). Similar examples could be given for other and more complex factor constellations, but, in any case, the resolution of such conflicts depends on which factor predominates. Slavic languages differ as to which factor, on average, predominates, and this is the main reason for differences in aspect choice between Slavic languages. Some of them are well known from Dickey’s work on the East-West Theory of Slavic aspect (Dickey 2000 and subsequent publications), for others see Benacchio (2004), Wiemer (2008, 2015, with further references). However, whatever these conflict resolutions look like, overall it is the growing number of contexts with complementary distribution that has made the choice between morphologically related stems grammatical. Even in the extreme case of lexical synonymy (i.e. with aspect pairs), stems belonging to opposite aspects cannot replace each other *ad libitum*, instead the distribution of pfv vs ipfv stems—and their paradigmatic relations—have become more predictable and reliable. No less important, the distinction of pfv vs ipfv stems, and their derivational relations, are not restricted by telicity. These tenets are summarized schematically in Fig. 1.

<sup>4</sup>Alternatively, we might say that aspect choice is dependent on contexts defined by these conditions. In this sense, conditions and contexts are just two sides of the same coin.



**Fig. 1** Most productive patterns of stem derivation and opposite functional inventories

Part (III) of Fig. 1 together with the boxes at its bottom are indicative of another crucial point. The tendency toward complementary distribution over opposite sets of heterogeneous functions applies to pfv vs ipfv stems as two generalized classes of verb stems, not to particular stems. Moreover, this applies largely regardless of the specific pattern of derivation (see (IIa) vs (IIb)) and of whether stems involved in such derivation can be considered aspect pairs or not; see Sect. 3. That is, stems that do not enter into such pairs, belong to the pfv or ipfv class depending on the functions that can be assigned to them; conversely, not all functions assigned to the ipfv aspect are ‘available’ to every ipfv stem, but the functions which an ipfv stem does have, all belong to the ‘ipfv set’.<sup>5</sup> The same applies, *mutatis mutandis*, to pfv stems. For this reason the PFV : IPFV opposition can be considered a classificatory category (Wiemer 2006; Wiemer and Seržant 2017). Such a category can be defined using the following quote from Plungjan (2000, p. 125):

[...] *opredelennoe množestvo leksem dannogo jazyka bez ostatka razbivaetsja na nepere-sekajuščiesja podklassy, každyj iz kotoryx karakterizuetjsja svoim značenim nekotoroj grammatičeskoj kategorii. Ėta kategorija pripisyvaetsja, takim obrazom, leksemam, a ne slovoformam i zadaet grammatičeskuju klassifikaciju leksiki; poëtomu ona i nazywaetsja klassificirujuščej.*

‘a given amount of lexemes of a language distributes over non-overlapping subclasses without any remainder; every subclass is characterized by its meaning for a certain grammatical category. Therefore, this category is assigned not to word forms, but to lexemes and determines the grammatical *classification* of the [given section of the] lexicon. This is why it is called classificatory.’ (emphasis original; see also Plungjan 2011, p. 53f.)

<sup>5</sup>In principle, the same holds true if, instead of ‘classical’ aspect pairs, we accept a ‘looser’ treatment of lexical relatedness between stems, like ‘aspectual partnership’ (Lehmann 1988, 1999; Mende et al. 2011), or ‘actionality groups’ (Tatevosov 2016, pp. 312–358), or even ‘aspectual clusters’ (Janda 2007). For a brief discussion cf. Wiemer (2019a, pp. 122–124).

### 3 Synonyms, grammatical distribution and paradigms

We now need to understand how paradigmatic relations can be modelled for distinct, but related stems which differ in aspect, but each of which have their own inventory of inflectional forms for finite and non-finite forms. This issue is particularly relevant for aspect pairs, i.e. those stems that are morphologically related, belong to opposite aspects (= grammatical classes) and, in addition, share a lexical meaning. Conversely, one might ask which kind of lexical unit underlies complex paradigms of such verb stems: is it one lexical entry, or lexeme, or more than one? The quote from Plungian at the end of Sect. 2 implies that we are dealing with different lexemes. This creates the two following questions: (a) How can two lexemes constitute one paradigm, given that each of them has its own entry in the lexicon? (b) Does it make sense to have a model in which two lexemes make up one paradigm? Clearly, in order for any decision to be made relating to the above-listed questions, independent, or superordinate, criteria are required. But, whatever we decide, any paradigmatic relation between the derivationally related pfv and ipfv stem must be built on complementary functions and constraints of use for the related stems.

We first deal with these constraints (Sect. 3.1), before moving on to so-called aspect triplets, which seemingly make the problem even more complicated (Sect. 3.2), and then discussing ensuing issues for the morphology-lexicon interface (Sect. 3.3).

#### 3.1 Trivial conditions of replacement and canonical functions

Aspect pairs can be considered the backbone of the PFV : IPFV opposition inasmuch as they most straightforwardly illustrate how complementary functions and collocational restrictions of form correlate with morphological patterns of stem derivation. A rigid notion of aspectual pairedness requires that two morphologically related stems code an identical lexical concept (at least in one of their meanings) and that these two stems follow reliable conditions (or factors, constraints) under which one of them replaces the other (see Sect. 2). In the strictest treatment aspectual pairedness operates on an ontology of eventualities (or situation types) since it requires that the replacement does not change the eventuality type of the situation, more precisely: the replacement must not change an event, understood as a situation with a limit, to something else (a process or a state). In practice, this means that only ipfv stems can replace pfv ones, but not vice versa, as pfv stems by default denote situations that are limited in time (either because they are telic and focus on reaching the ‘telos’, or because there is an aspect operator putting an external boundary on the situation), and ipfv verbs are able to denote such events as well, but the reverse does not hold true: pfv verbs cannot denote processes or states, they cannot therefore replace ipfv verbs denoting processes and states (single or repeated ones).<sup>6</sup> Now, conditions under which an ipfv stem replaces a pfv stem, but does neither change the latter’s lexical meaning nor even its eventuality type (i.e. denotes an event), are called ‘trivial’.

There are three tests (or types of contexts) by which trivial pairedness can be established in Russian as well as in Polish; the illustrations in (4)–(6) are from Polish (constructed):<sup>7</sup>

<sup>6</sup>This strict view on aspectual pairedness was pursued mainly by Russian aspectologists and semanticists dealing with aspect. Its roots go back to Maslov’s ground-breaking article (Maslov 2004[1948]), the theoretical considerations leading to the requirement of ontological identity are concisely laid out in Zaliznjak and Šmelev (1997, pp. 37–44), cf. also Zaliznjak, Mikoëljan, and Šmelev (2015, pp. 43–50). For a discussion cf. Wiemer (2017, p. 227f.).

<sup>7</sup>The second test might work slightly less reliably in Polish, and the first two tests together are less reliable in Czech (as a result of different strength of conflict resolutions mentioned in Sect. 2), but these differences prove irrelevant for our case study in Sect. 4.

- i. narrative past: PFV (4a) → narrative present: IPFV (4b)
- (4) a. W 1832 roku Mickiewicz *przyjechał*<sup>pfv</sup> do Paryża, a dwa lata później *ożenił*<sup>pfv</sup> się z Celiną Szymanowską. *Żona urodziła*<sup>pfv</sup> mu sześćcioro dzieci.  
'In 1832 Mickiewicz *arrived* in Paris, two years later he *married* Celina Szymanowska. His wife *gave birth* to six children.'
- b. W 1832 roku Mickiewicz *przyjeżdża*<sup>ipfv</sup> do Paryża, a dwa lata później *żeni*<sup>ipfv</sup> się z Celiną Szymanowską. *Żona rodzi*<sup>ipfv</sup> mu sześćcioro dzieci.  
'In 1832 Mickiewicz *arrives* in Paris, two years later he *marries* Celina Szymanowska. His wife *gives birth* to six children.'
- ii. single event (in the past): PFV (5a) → unlimited repetition of the same event (in the past): IPFV (5b)
- (5) a. Dzisiaj Marta *obudziła*<sup>pfv</sup> się o piątej rano.  
'Today Marta *woke up* at 5 am.'
- b. Zwykle Marta *budziła*<sup>ipfv</sup> się o piątej rano.  
'Usually Marta *woke up* at 5 am.'
- iii. directive speech act (marked with imperative or equivalent forms) for a single action: PFV (6a) → the same, but under negation (i.e. prohibition related to a single situation): IPFV (6b).
- (6) a. *Zostaw*<sup>pfv</sup> trochę zupy.  
'Leave over some soup.'
- b. *Nie zostawiaj*<sup>ipfv</sup> zupy.  
'Don't leave over soup.'

These replacements are trivial in the sense that no ontological change ensues, but they are trivial also inasmuch as they are only diagnostics to check for aspectual pairedness. Crucially, these tests work regardless of whether the related pfv and ipfv stems show semantic differences under other conditions. In fact, there are pairs of morphologically related stems of opposite aspect which are lexically identical in all of their meanings (e.g., Pol. *potknąć się* – *potykać się* 'stumble', *zauważyć* – *zauważać* 'notice', *dostać* – *dostawać* 'receive'); this entails that the ipfv stems can only denote events, as do their pfv equivalents. However, among all aspect pairs established by tests of trivial replacement those verbs certainly represent a minority. Usually, the ipfv member of an aspect pair can also be used in diverse functions in which it does not just 'copy' the event of the pfv member (and for which the pfv member is not 'available'). From this it follows that the function range for ipfv stems is by far more diversified (and often less predictable) than for pfv stems. This variability of the semantic relationship between the pfv and the ipfv member of aspect pairs brings about their various types which have been worked out in aspect theories with a primary focus on the interaction between aspect (as an operator) and actionality features inherent to (or implied by) the lexical meaning of verbs (understood as operands). These functions belong to what we can call 'canonical functions' of (pfv vs ipfv) aspect.

Canonical functions<sup>8</sup> form a subset of the function network mentioned in Sect. 2. They refer to basic actionality types (events, processes, states), to event-external plurality (single vs repeated situation) and to temporal definiteness (a.k.a. episodicity), thus to core functions of aspect (in Slavic and other languages). They are usually listed in the literature on Slavic

<sup>8</sup>The term is used here after Lehmann (1999, pp. 220–223) with some amendments.

(in particular, Russian) aspect among a standard inventory of functions. Pfv stems are basically ascribed only one canonical function, namely: ‘concrete-factual’ (i.e. reference to a single event). Ipfv stems can have a large array of functions. Without engaging into formal descriptions or dwelling on a general classification, e.g., of event-external pluractionality (for which see fn. 3), the terms used here follow the general practice of Slavic aspectology and can briefly be characterized as follows:

- Progressive (a.k.a. processual): denotation of an episodic activity; this implies that particular phases are viewed as simultaneous to subsequent reference intervals. See ex. (7)–(8) from Polish:
  - (7) Tata posmutniał jakoś i nic nie mówił, a ja *działem* pomarańczę.  
‘My father got sad and didn’t say anything and I *was dividing* the orange.’  
(S. Kowalewski. Czarne okna. 1993)
  - (8) Wiarus w kapie *rozdzielał* kielichy między pijących.  
‘An old soldier in a blanket *was distributing* cups among drinkers.’  
(St. Żeromski. Popioły. 1904)
- General-factual: practically identical to the function of the experiential perfect, known from typological studies. The speaker states that, or asks whether, a situation occurred at least once; contrary to the progressive function, the general-factual is *per se* not sensitive to actionality types. This function is well-attested for ipfv verbs in Russian, it is much more difficult to find good examples in Polish (or Czech); cf. Łaziński (2020, §4.3.3). Here are two standard examples (typically with verbs of communication), adapted after Łaziński (2020, §4.3.3), from Polish:
  - (9) *Dzwoniłem* już do niej. (Więc nie mów, że nie starałem się dojść z nią do porozumienia.)  
‘I’ve already *called* her. (So don’t say that I didn’t try to reach an agreement with her.)’
  - (10) *Ostrzegalem* ją przed nim. (Nie wiem, czy usłuchała mojej rady.)  
‘I *warned* her against him. (I don’t know if she followed my advice.)’

Sentence negation can favor a general-factual interpretation because its scope can comprise the meaning component ‘at least once’: NEG(at least once)  $\cong$  NEVER. Compare one of the very rare examples from our Czech sample for which this function could be considered:

- (11) Já jsem nikdy zastupitele *nedělila* na muže a ženy.  
‘I have never *divided* representatives into men and women.’ (D. Moravia. 2004)
- iterative: by this we mean only those repeated situations which are conceived of as occurring within a limited larger time span, but without any signs of lawfulness; this precludes the rise of modal overtones, which have typically been associated with the habitual function (see next type):
    - (12) Tymczasem Magda mówi. Jej opowieść rozrasta się, rozlewa, mówi tak, jakby chciała powetować sobie miesiące milczenia, wprost dławi się słowami, od czasu do czasu, niby przecinkiem, *rozdziela* zdania krótkim szlochem.  
‘Meanwhile Magda talks. Her story grows, spreads, she speaks as if she wanted to make up for the months of silence, she chokes the words, from time to time, she *separates* her sentences with a short sob, as if with a comma.’  
(E. Nowacka: Małe kochanie, wielka miłość. 1997)



- habitual: the situation is presented as observable at an unlimited number of occasions (regularly or irregularly), so that it can be expected to occur; this can (but need not) imply dispositional modality ('X typically does S'  $\supset$  'X is able to do S'). See the Polish examples without this implicature:
  - (13) W taki dzień przyjemność odczuwania smaku czekolady miała inny wymiar. Obaj jedli ją w różny sposób: Wojtek pożerał od razu całą, Marek *dzielił* na części.  
'On such a day, pleasure of experiencing chocolate had a different dimension. Both ate it in a different way: Wojtek devoured it all at once, Marek *divided* it into parts.'  
(M. Kamiński, W. Moskal, S. Swerpel: Nie tylko w biegu. 1996)
  - (14) Kiedyś przywileje *rozdzielał* osobiście tylko monarcha.  
'Once, only the monarch *distributed* personally the privileges.'  
(R. Kapuściński: Cesarz. 2005)
- stative: the situation is presented as unlimited in time, but without any internal division into distinct subevents (or subintervals); see the Polish examples:
  - (15) Ścianę fasady *dzielią* pilastry wielkiego porządku.  
'The pilasters of great order *divide* the facade wall.'  
(M.I. Kwiatkowska, M. Kwiatkowski, K. Wesołowski: Znane i nieznanne: rezydencje, ludzie, wydarzenia. 2001)
  - (16) Wspomniane tu lasy sierosławskie [...] *rozdzielały* skupiska osadnicze dwóch sąsiadujących opoli: woborskiego i rozpierskiego.  
'The Sierosław forests mentioned here [...] *separated* two neighboring settlements: the Wolbórz and the Rozprza settlements.'  
(K. Modzelewski: Barbarzyńska Europa. 2004)

Some more specific functions could, in principle, be added, but we have refrained from even mentioning them for reasons of space and because they proved to be irrelevant to our case study (see Sect. 4.2). As concerns the aforementioned functions, in our samples of Pol. *dzielić* – *rozdzielać* and Cz. *dělit* – *rozdělovat*, the progressive and even more so the iterative and the general-factual meaning are extremely rare.<sup>9</sup>

The principle point to make is: if pfv and ipfv stem each have their function inventories, and if some pfv and ipfv stems share a lexical meaning so that they enter into an aspect pair, why shouldn't they be regarded as representatives of a single lexical unit? That is, why couldn't we accept that the set of finite and non-finite forms of pfv and ipfv stem jointly constitute a unitary paradigm? By comparison, in English, all forms of a verb that distinguish simple and continuous / progressive forms (and the latter are regarded as 'analytical inflections') are considered members of one paradigm (= forms from one lexical entry), and we can discern a comparable array of aspectual functions as we distinguish between an ipfv and a pfv stem, see (18) in Table 1, although in English the distinctions are distributed over simple forms and combinations of auxiliary and participle forms, see (17) in Table 1.

The comparison in (17)–(18) has been adopted from Tatevosov (2016, pp. 303–356). On the basis of this comparison between the English tense-aspect system, which is assumed to be inflectional, and the Russian (Slavic) system, which he considers representative of

<sup>9</sup>Thus, token frequency distinguishes 'ordinary' iterative situations from habitual ones. Moreover, pfv verbs are used quite freely with present tense forms (even in Russian), only for the latter, but not for the former, and it is primarily pfv present tense forms which trigger the aforementioned modal implicatures. However, these issues exceed the scope of this article.

**Table 1** Functional distribution across forms of stems

English	Russian	Aspect function
(17) to plough <i>He ploughed the field.</i> <i>He was ploughing the field.</i> <i>Have you ever ploughed a field?</i> <i>He used to plough the field.</i>	(18) paxat <sup>ipfv</sup> – vs-paxat <sup>pfv</sup> <i>On vs-paxal pole.</i> <i>On paxal pole.</i> <i>Ty kogda-libo paxal pole?</i> <i>Obyčno on paxal pole.</i>	PFV: [+once], [+limit.] IPFV: [+once], [±limit.], process IPFV (past): experiential IPFV: habitual

classificatory-derivational aspect, he suggests that in the latter type of aspect system verb stems are organized in actionality groups just like the aspect grams of a system with inflectional aspect mark actionality distinctions for one and the same lexeme. This consideration is very sober and consistent, but Tatevosov does not say anything about the relationship between the stems organized in actionality groups (whose extreme would be an aspect pair) and lexical entries; in fact, he admits he remains agnostic with regard to this issue (2016, pp. 18–35 et passim).

Yet, the matter becomes even more complicated. In the illustrations of canonical functions above we used examples of Pol. *dzielić* and *rozdzielać* (taken from the sample of our case study shown in Sect. 4). Both are ipfv and mean ‘separate, divide’. The pfv counterpart for both is *rozdzielić*, i.e. either ipfv stem can ‘take claim’ to enter into an aspect pair with that pfv stem. Bearing in mind the function inventories and collocational restrictions which we employed above to justify a unitary paradigm for members of aspect pairs, we now, at first glance, face the problem of how the functions of ipfv stems might be distributed over the two ipfv stems, one without a prefix, the other one with a prefix shared with the pfv stem.

### 3.2 Natural Perfectives and triplets

With a few exceptions, prefixation of a simplex stem changes an ipfv verb into a pfv one. In most cases, this grammatical change is accompanied by a modification or change of the lexical meaning of the simplex stem. There is, however, a considerable amount of pfv stems whose prefixes do not lead to meaning modification of the simplex stem; compare, for instance, Russ. *stroit* ⇒ *po-stroit* ‘build’, Pol. *spowiadać* ⇒ *wy-spowiadać* ‘confess (a penitent)’, Cz. *psát* ⇒ *na-psat* ‘write’. Some scholars have dubbed the meaning contribution of such prefixes vaguely as ‘general-resultative’. Simultaneously, they deny that such ‘resultative’ prefixes derive aspect pairs from ipfv simplex stems, i.e. they do not accept that the derivational pattern (IIa) in Fig. 1 can yield aspect pairs, thus assuming that ‘true’ aspect pairs only emerge as a result of suffixation of already prefixed pfv stems (so-called ‘secondary imperfectivization’, see pattern (IIb) in Fig. 1); cf. most prominently Isačenko (2003[1960]), among many others. Obviously, the general assumption is that an ideal grammatical opposition between pfv and ipfv aspect must rest on a distinction which is binary in terms of its formal expression as well, and that the ‘natural’ direction of development for this opposition is to end up with a uniform way of forming pairs via secondary imperfectivization. Similarly, although in a Neo-Davidsonian semantic framework, Tatevosov (2015, pp. 272–295) argues that in cases of pattern (IIa) the prefix is not void of meaning, as it adds a resultative subevent. This does not imply that there cannot be pairs derived according to this pattern (in fact, Tatevosov is not committed to aspectual pairedness as such), but neither his analysis,

nor the vague ‘general-resultative’ function of prefixes, explains why the ipfv simplex is able to replace its pfv derivative in trivial tests of aspectual pairedness (see Sect. 3.1).<sup>10</sup>

Debates about the role of prefixes added to simplex stems were, for longer periods, reduced to the question whether semantically empty prefixes exist. However, several scholars, starting with Vey (1952) and van Schooneveld (1958) in the 1950s, have emphasized that the question about ‘empty prefixes’ is misplaced (for instance, Breu 1984, pp. 11–14). Recently, Janda and associates have convincingly argued that prefixes need not be semantically empty in order to derive pfv stems from simplex stems to yield aspect pairs which fulfil trivial conditions of replacement. Instead, the only thing required is that the meaning potential of the prefix is compatible with a meaning component which is inherent to the lexical meaning of the simplex. In other words: an overlap between the meaning of the simplex and of the prefix must exist (Janda et al. 2013).<sup>11</sup> Accordingly, pfv stems that enter into aspect pairs with ipfv simplex stems are called Natural Perfectives (NatPerfs), as the prefix yields an easily compatible (‘natural’) semantic overlap with a meaning component of the simplex. If, in the derivation of a pfv stem, the prefix causes a change (or modification) of the lexical meaning of the simplex, we are dealing with Specialized Perfectives (SpecPerfs).

Janda and associates tested their hypotheses for contemporary Russian using dictionaries and the RNC. They found that 1,429 simplex stems supply the basis of altogether 1,981 NatPerfs (i.e. there are simplex stems with more than one NatPerf, derived with different prefixes), and that 16 out of 21 verbal prefixes are used in the derivation of NatPerfs. The type frequency of these prefixes in NatPerfs differs enormously. On the token level, NatPerfs are on average about 10 times more frequent than SpecPerfs. Therefore, NatPerfs are no *quantité négligeable*.

A similar picture arises for Polish from figures provided by Łaziński (2020) on the basis of the Great Polish-English Dictionary (PWN Oxford 2004). This dictionary lists 4,630 Polish aspect pairs (of any derivational pattern), among them 1,670 (= 36%) represent NatPerfs.<sup>12</sup> 15 prefixes are attested in NatPerfs, basically all of them cognates of the Russian ones occurring in NatPerfs, here in the order of decreasing type frequency: *z-* / *ze-* / *s-* / *ś-* (47.5%), *za-* (19.5%), *wy-* (10.5%), *po-* (7.5%), *u-* (6.5%), *o-* / *ob-*, *na-*, *prze-*, *roz-*, *w-*, *od-*, *przy-*, *pod-*, *do-*, *wz-*. However, the share of simplex stems that give rise to more than one NatPerf (e.g., *malować* ‘paint’ ⇒ *namalować* / *pomalować*) is lower than in Russian, as the amount of simplex stems almost equals the amount of NatPerfs. See some good examples of aspect pairs with NatPerfs:

#### Polish

- (19) *s-/z-*    budować – zbudować ‘build’, kończyć – skończyć ‘finish’  
*na-*        pisać – napisać ‘write’  
*prze-*     czytać – przeczytać ‘read’  
*u-*         gotować – ugotować ‘cook’  
*za-*        angażować – zaangażować ‘engage, enrol’

<sup>10</sup>Actually, Tatevosov’s formal analysis of actionality and the subevent structure of eventualities target differences of meaning, whereas tests of trivial pairedness search to establish lexical and ontological identity (cf. Wiemer 2017, p. 227f.). The aims of these approaches thus inherently differ from the start.

<sup>11</sup>Actually, this consideration ultimately builds a specific variety of a very general, and widely acknowledged, regularity of lexical semantics: for two (or more) word forms to enter into a collocation, they have to share some non-trivial meaning component(s) with each other.

<sup>12</sup>For comparison, there are 2,960 suffixal pairs (64%).

This being said, we have to realize that the suffixation of prefixed (pfv) stems yielding ‘secondary imperfective’ stems (cf. pattern (IIb) in Fig. 1) is very productive and is actually the pattern with the highest type frequency in contemporary Polish and Russian (and certainly also Czech). Thus, NatPerfs provide the necessary precondition for the rise of aspect triplets, as in the following cases from Polish:

(20)	a.	IPFV1	⇒	PFV	⇒	IPFV2	
		<i>tworzyć</i>		<i>s-tworzyć</i>		<i>stwarz-a-ć</i>	‘create, produce’
	b.	<i>szacować</i>		<i>o-szacować</i>		<i>oszacow-ywa-ć</i>	‘estimate, assess’

Secondary imperfective stems (henceforth: IPFV2) arise from productive patterns of suffixation, while the respective ipfv simplex stems (= IPFV1) keep their pair relation with their pvf derivatives whose prefix does not modify the lexical meaning of the simplex. Therefore, triplets result from a coincidence of two factors:<sup>13</sup> the conditions that lead to NatPerfs and the productivity of secondary suffixation, which has become the dominant pattern in the formation of aspect pairs. As (20a–b) indicate, these processes ‘work’ with both inherited and borrowed verb roots (*szac-(owa-)* < Germ. *schätzen* ‘estimate’).

From these considerations it follows that aspect triplets form a subset of NatPerfs: any triplet entails a corresponding NatPerf, but not every aspect pair with a NatPerf forms part of a triplet. Nonetheless, as with NatPerfs, the amount of aspect triplets in Russian, as well as in Polish and Czech, is considerable. The Exploring Emptiness database, created by Janda et al. in connection with their book (2013), lists triplets for contemporary Russian.<sup>14</sup> According to the information given on the website of this database, there are between 753 and 1,583 triplets in modern Russian; the enormous difference in the counts of type frequency depends on how we account for the data sources, token frequency and some other factors. On the one hand, this difference makes one suspicious about the validity of the information; on the other hand, even if the lower number of 753 (which corresponds to the most ‘conservative’ assessment) is taken as a basis, we see that triplets are by no means a marginal phenomenon. This holds even though figures concerning aspectual triplets require particular circumspection, especially if a diachronic perspective is involved (see Sect. 5, fn. 22).

### 3.3 Repercussions for the morphology-lexicon interface

Now, how can triplets be integrated into a model of the interface between morphology and the lexicon? Theoretically, we can consider three options for contemporary Russian, which have already been discussed in Apresjan (1988, 1995). Either we exclude the simplex stems (IPFV1) and assume aspect pairs only for PFV and IPFV2; IPFV1 would remain as IPFV tantum stems:

Model A



This is the decision made by those who do not accept aspect pairs derived from simplex stems by prefixes (see Sect. 3.2). This is problematic for a variety of reasons (which are generally acknowledged), but as far as triplets are concerned the main problem is that in many

<sup>13</sup>More precisely, we are dealing here with what Zaliznjak et al. (2015, p. 58 et passim) have dubbed ‘biimperfektivnye vidovye trojki’: two related ipfv stems ‘take claim’ as the ‘proper correlate’ of the respective pvf stem. The notion of triplets may be understood more broadly (cf. Wiemer 2020, §3).

<sup>14</sup>See: [http://emptyprefixes.uit.no/triplets\\_eng.htm](http://emptyprefixes.uit.no/triplets_eng.htm).

cases IPFV2 behave less like *ipfv* ‘copies’ of ‘their’ *pfv* stems than IPFV1 do. Compare, for instance, Russ. *vit’ – s-vit’ – svi-va-t’* (*gnezdo*) ‘twist (a nest)’, *lepit’ – s-lepit’ – slepl-iva-t’* (*velikana iz snega*) ‘sculpt (a giant out of snow)’.

The second option consists in regarding the triplet as a combination of two aspect pairs which are, as it were, inserted into one another, with the *pfv* stem serving as a kind of pivot:

Model B



The problem here lies in the assumption that one and the same *pfv* stem has two different *ipfv* ‘partners’, and this multiple partnership does not correspond to any meaning alternation in the pivotal *pfv* stem. If such alternation existed, the *pfv* stem could be declared polysemous, but this obviously does not comply with the facts. Conversely, the two *ipfv* stems would have to be considered to be synonyms, but it is not clear what conditions apply when they are chosen, nor what would follow for the lexicographic account if we knew these conditions.

The third option amounts to uniting all three stems into one paradigm, i.e. to treat all of them as representatives of the same lexical entry:

Model C



This is what we already suggested at the end of Sect. 3.1, but this suggestion becomes even more radical if more than two stems of the same aspect are involved which, for some reason, vary but can be regarded as ‘equally good’ partners of the *pfv* stem (see fn. 15). In traditional accounts, both the second and the third option raise objections, because they violate acknowledged principles in the relation between lexical units and paradigms. Apresjan (1988, p. 38) called the second option ‘paradoxical’, because it would require a crisscrossing of the paradigms of the grammatical forms of two synonyms (IPFV1 and IPFV2). He believed the third option to be even worse (‘unnatural’), because it would require such a radical revision of our conception of verb paradigms and would ‘not fit into any realistic picture of the language’ (1988, p. 39; our translation).

However, Apresjan’s further argument gives us a key toward a solution for this conundrum. Namely, his strong objection against a model like Model C rests on the assumption that a language does not have the luxury of having two or more viable variants of one and the same grammatical form that are, in addition, generated by productive processes in the formation of verb stems (1988, p. 40).<sup>15</sup> Actually, why wouldn’t it be natural to expect that productive patterns create variants? More so, we would predict that high productivity causes variation that not necessarily, and not always, leads to meaningful oppositions of choices (at least not if inspected by the naked eye); only after some time, seemingly functionless variation may be differentiated, or reduced e.g. by the loss of variants and the preservation of others. After all, the productive patterns of stem derivation that we observe today as crucial for the *PFV* : *IPFV* opposition in Slavic languages (see Fig. 1) are the outcome of the continuous

<sup>15</sup>See the original formulation: “my prixodim k protivostestvennomu dopuščeniju, budto jazyk pozvoljaet sebe roskoš’ imet’ ot dvux do pjati odinakovo žiznesposobnyx variantov odnoj i toj že grammatičeskoj formy, poroždaemyx k tomu že vpolne živymi, a inogda i produktivnymi processami vido- i slovoobrazovanija”. In this context, Apresjan took into consideration also cases in which more than two stems could be considered as variants, e.g. Russ. *kopit’<sup>ipfv</sup>* (*den’gi*) – *nakopit’<sup>pfv</sup>* – *nakaplivat’<sup>ipfv</sup>* – *nakopljat’<sup>ipfv</sup>* – *skopit’<sup>pfv</sup>* – *skaplivat’<sup>ipfv</sup>* ‘save, accumulate (money)’. We are not going to consider such cases here, but cf. Wiemer (2020) for discussion.

renovation of a very ancient technique (namely, suffixation) and its combination with a more recent technique (namely, prefixation), which both have been preserved (in contrast to other IE families in Europe; cf. Wiemer and Seržant 2017). The vitality of triplets, in particular, can be explained as the coincidental meeting of two factors: the preservation of the motivating simplex (IPFV1) stems (maybe because of their higher token frequency?) and the appearance of IPFV2 stems on the basis of a pattern (imperfectivization of prefixed pfv stems) which, though much younger, has become the predominant one, at least at type level.

Two additional comments are in order here. First, in a preliminary account, among the triplets derived from Slavic roots (or old borrowings) only very few simplex stems can be regarded as the result of deprefixation of prefixed pfv stems; an exclusion are deadjectival verb stems (e.g., Pol. *pełnić* appeared later than *s-pełnić* – *s-pełni-a-ć* ‘(ful)fil’; cf. Wiemer 2017, p. 236f.). Second, the derivation of IPFV2 stems from aspect pairs of the type [simplex stem.IPFV—prefixed stem.PFV] is extremely productive (as a quick google makes more than apparent), but such IPFV2 stems also come and go. For instance, we have Pol. *napisywać* and Cz. *napisovat* in older sources, but they seem to have disappeared again; similarly, Russ. *napisyvati* is only attested 9 times in the RNC, the last attestation dating from 1920 (stage: Sept. 2019), but this IPFV2 can be found in internet sources (with the meaning ‘write’) without problems.

To sum up: a synchronic view into a language is largely a reflex of diachronic changes, and a reasoning like Apresjan’s is, to some extent, based on the belief that a language system has to be economic and avoid redundancies. Thus, as long as the interest primarily or exclusively lies in ‘synchronic stock-taking’, e.g. for the purpose of a consistent and systematic lexicographic account, we will encounter the paradoxes that Apresjan rightly pointed out. From this perspective, the aforementioned Models B and C are not viable, but Model A is inadequate as a decision to be applied in a wholesale manner. Apresjan is aware of this and concludes that, from a lexicographic perspective, each case has to be treated individually, and this yields a continuum of triplet-types for which he points out four focal points (Apresjan 1988, 1995). The criteria for this classification basically amount to applying canonical and trivial tests illustrated in Sect. 3.1.

In the following case study diachrony turns out to not be very relevant at all, although we took its existence into account. We will now focus on trying to find out what might cause the distribution of the two ipfv members of a typical aspect triplet (Sect. 4), before we once again question where a more radical revision of the conception of the verbal paradigm would take us in theory, and what is required to carry this theory into practice (Sect. 5).

#### 4 Case study: distribution of Cz. *dělit* – *rozdělovat*, Pol. *dzielić* – *rozdzielać* ‘divide; share’

As explained in the preceding sections, the distinction of pfv and ipfv verbs is based on a close to complementary distribution of functions over sets of finite and non-finite forms of stems and their combinatorial constraints. Simultaneously, two morphologically related stems of opposite aspect come closest to an ideal aspect pair if, apart from these functional distinctions, the lexical meaning remains identical. Moreover, the function inventories apply to the class of pfv and the class of ipfv stems, respectively, in their entirety; that is, many (most?) particular stems do not allow for the full inventory, but the functions they do allow for belong to the inventory of their respective class only (pfv or ipfv).

What, then, about the two ipfv stems IPFV1 (= unprefixated, simplex) and IPFV2 (= prefixed and additionally suffixated) that form, together with their PFV, a triplet? Both of them

have functions that belong to the inventory of the ipfv class, but how are these functions distributed among IPFV1 and IPFV2? Theoretically, the following constellations along a continuum from (i) to (iii) are possible:

- (i) IPFV1 and IPFV2 have an identical range of functions.
- (ii) IPFV1 and IPFV2 have overlapping ranges of functions.
- (iii) IPFV1 and IPFV2 have complementary functions, i.e. they choose non-overlapping subsets from the general set of functions characteristic of ipfv stems.

These options exist for both finite and non-finite forms of IPFV1 and IPFV2. Of course, these theoretical considerations should be applied to a larger number of tokens for both ipfv stems, since functional distributions result from observations about occurrences (= tokens) and not about types. Thus, even if, theoretically, lexicon entries of IPFV1 and IPFV2 gave support to claims that these stems have non-overlapping aspect functions from the canonical inventory of ipfv stems, it would hardly be realistic to assume a 100% complementary distribution for the sum of their tokens in real texts. In turn, option (i), i.e. identical ranges of functions, would not necessarily result in a random distribution (in a random sample drawn from a sufficiently representative corpus), just like functions listed in dictionaries (and characteristic for IPFV1 and IPFV2) need not be distributed evenly for both ipfv stems in natural discourse. That is, the distribution of the two morphologically related, synonymous stems—both belonging to the grammatical class of ipfv stems—will probably be skewed anyway.

The empirical question that follows is how large this skewedness is, that is, whether the specific usage characteristics associated with the two morphologically related, imperfective and synonymous stems are significantly different. To address this problem we focused on one exemplary case: the Czech IPFV1 *dělit* and IPFV2 *rozdělovat* ‘divide, separate; share’ and their Polish cognates *dzielić* (IPFV1) and *rozdzielać* (IPFV2). We adopted a profile-based approach, which included the following steps: (i) extracting authentic usage examples from corpora, (ii) data annotation for multiple features, and (iii) statistical modelling. Both binomial logistic regression and tree-based techniques were used during the last step.

Profile-based approaches to lexical semantics have been proposed since the early nineties (cf. Geeraerts, Grondelaers, and Bakema 1994), cf. also Divjak and Gries (2006) and Glynn (2010). Such approaches have also been applied to grammatical features (cf., for instance, Klavan 2014). In our case, both the particular forms and the particular functions that are dominant for each of the two ipfv stems constitute their specific profiles. This is similar to the idea of grammatical profiles established in Janda and Lyashevskaya (2011) for the ipfv and pfv members of aspect pairs. The basis for differentiation of grammatical profiles in that study were dominant morphological patterns in contemporary Russian, namely the two represented in (IIa–b) of Fig. 1. These patterns are the same in Polish and Czech. However, Janda and Lyashevskaya’s profiles account only for grammatical forms (on a medium level of granularity), whereas we combine an account of forms and functions. Moreover, we are concerned with the distribution of two ipfv stems either of which has the ‘right’ to be accepted as an ipfv member in a pair with the pfv verb. Finally, we are dealing with two distinct Slavic languages, and we ask the question whether the distribution that exists between two ipfv stems reveals any considerable changes that have taken place during the past 250+ years.

#### 4.1 The database and the choice of the IPFV1 – IPFV2 pairing

As a part of the *DiAsPol250*-project, we created a database of aspect triplets in Czech, Polish and Russian. This database is at a very advanced stage, facilitating selection from a large

number of lexical items and various characteristics.<sup>16</sup> We decided on Cz. *dělit – rozdělovat*, Pol. *dzielić – rozdzielać* ‘divide, separate; share’. Before we comment on this choice (see Sect. 4.1.2), we will provide information on the corpus data.

#### 4.1.1 Composition and provenance of the data

We divided the period 1750–2018 into three subperiods: 1750–1917, 1918–1989, 1990–2018. Their uneven range resulted from two main considerations: on the one hand, cuts in the periodization should reflect cornerstones in the external history of Polish (the language which occupies center stage in the project); most of them coincide with important dates in Czech and Russian history (see, first of all, the end of World War I and the year 1989, when the communist regimes in Eastern Central Europe and Russia began to disintegrate). On the other hand, we needed large enough corpora to be able to obtain sufficient amounts of tokens for the chosen cognate stems. As a consequence, the first subperiod comprises a very large interval (which starts in 1750, i.e. before the division of the Polish-Lithuanian Commonwealth), for which the amount of data is much smaller anyway than for the respective last subperiod, which represents the smallest interval (see Table 2). The differences in size between the Polish and Czech subcorpora for the particular subperiods are due to different strategies in their composition. All Czech data were taken from subcorpora of the Czech National Corpus, *syn\_v7* (synchronic) and *diakorp6* (diachronic). The Polish data were taken from different corpora: the Polish National Corpus, which comprises the 20th century until 2010, the newly available Corpus of 17th and 18th century Polish (data until 1772, so-called *KorBa*), a Polish corpus comprising the period 1830–1918 and created for research on inflectional morphology,<sup>17</sup> and a corpus with texts dating from the end of the 18th century until 1918, created for purposes relating to *DiAsPol250* and not yet publicly available. Additional items to cover the 18th and 19th centuries were collected for Polish from texts in *Wikisource* and *Wolne Lektury* (see Sources), which were examined with AntConc.

Due to the different sizes of the Polish and Czech subcorpora (see Table 2), non-random samples were used for two of the predefined periods (1750–1917, 1918–1989) in Czech and for the first period in Polish, for which all available relevant observations were included. Random samples were provided for the period 1990–2018 in both languages and for the second and third period in Polish, as the relevant corpora were sufficiently large.<sup>18</sup>

**Table 2** Distribution of tokens over periods (IPFV1+IPFV2 = sum)

	1750–1917	1918–1989	1990–2018	sum total
Cz. <i>dělit –</i>	27 + 11 = 38	114 + 169 = 283	300 + 300 = 600	
( <i>rozdělit –</i> )	tokens in corpus:	tokens in corpus:	tokens in corpus:	921
<i>rozdělovat</i>	2.3 mln	25.5 mln	5 bln (random)	
Pol. <i>dzielić –</i>	115 + 121 = 236	101 + 100 = 201	100 + 100 = 200	
( <i>rozdzielić –</i> )	tokens in corpus:	tokens in corpus:	tokens in corpus:	637
<i>rozdzielać</i>	14.2 mln	60 mln (random)	240 mln (random)	

<sup>16</sup>On a preliminary account we can say that for the period 1750–2017 we have 654 triplets in Czech, approx. 1016 in Polish and approx. 1270 in Russian.

<sup>17</sup><https://szukajwslownikach.uw.edu.pl/f19/>.

<sup>18</sup>We gratefully acknowledge the help by Dorota Górnicka-Urban and Katarzyna Osior-Szot (Warsaw), who extracted and analyzed the Polish data.



The composition of the datasets followed the principle that they should include a comparable number of observations for both IPFV1 and IPFV2 stems for each language and each period. This aim was not equally achieved (see the Polish and Czech datasets period by period in Table 2), but at least for Polish the number of observations per period is comparable. Regardless of the differences in size for period×language, this design makes sure that tokens of competing ipfv stems can be dispersed over a large range of functions and forms, so that potentially divergent weights of IPFV1 and IPFV2 in a comparison of the three periods could be revealed. Moreover, the chosen statistical methods (in particular, the tree-based techniques) are in general considered to be sufficiently robust to cope with differences in sizes of subsamples.

#### 4.1.2 Why this choice of triplet?

Let us now turn to the lexical items chosen for analysis. We chose Cz. *dělit* – *rozdělovat*, Pol. *dzielić* – *rozdzielać* ‘divide, separate’, because the triplets which they belong to are among the most frequent since 1750. These stems belong to the basic vocabulary of both languages, they have a maximal range of identical meanings, and there was no reason to assume that these meanings changed considerably over the last 250+ years. However, for each language the two ipfv members of the triplets do not completely overlap in their pattern of meaning alternations (polysemy). A broader semantic scope characterizes IPFV1 (Pol. *dzielić*, Cz. *dělit*), which has the additional meaning ‘isolate, separate by some distance’:

##### Czech

- (21) Prahu *dělí* od Brna nejen 200 kilometrů, ale také řada rozdílů, tradic i předsudků na obou stranách.  
 ‘Prague is *separated* from Brno not only by 200 kilometers, but also by a set of differences, traditions and prejudices on both sides.’ (lit. ‘Not only 200 km *separate* Prague from Brno, but also...’) (Mladá fronta DNES. 2010)

##### Polish

- (22) Tu akurat niezbyt mu się spodobało—ołtarz był ogromny, a Papieża *dzielili* od wiernych zbyt duża odległość.  
 ‘He didn’t like it here—the altar was huge, but the Pope *was too far from* the faithful people.’ (lit. ‘... the distance which *separated* the Pope from the faithful was too big.’)  
 (P. Zuchniewicz: Jan Paweł II: „Będę szedł naprzód”. Powieść biograficzna. 2009)

For this reason, we analyzed only those cases in which IPFV1 and IPFV2 had an identical lexical meaning, namely ‘divide, separate’; the underlined meaning can be considered the more general and frequent one. Additionally, we had to remove corpus hits with the reflexive clitic Pol. *się*, Cz. *se*. All stems of these triplets derive anticausatives (e.g., Pol. *Ich drogi się (roz)dzielily* ‘Their paths separated’) and closely related meanings associated to derivations which, by changing the argument structure, yield new lexical items. The simplex stem occurs in even more, and less trivial, meanings of lexical derivation (e.g., reciprocals with three arguments, as in Pol. *Rozbitkowie dzielili się resztkami prowiantu* ‘The shipwrecked people shared the remains of their provision’, or Cz. *Nějak se s nimi musíme o tuto planetu dělit* ‘Somehow we have to share this planet with them’), for which pfv stems with different prefixes occur as aspectual partners. Such cases were eliminated from the samples.

The ‘price’ we paid for this additional step in data selection was an exclusion of tokens in which the reflexive clitic served only as a means of agent demotion from the syntax, i.e. in passive-like uses as in the particularly frequent Polish ‘reflexive impersonal’, in which only the nominative subject is blocked, but the underlying argument (= human agent) remains implied as an argument. Compare, for instance, (23a), which is denotationally equivalent to the unmarked active construction with a nominative subject, cf. (23b):

- (23) a. Zapałk-ę dzieli-ł-o się na czworo  
 match-ACC divide[IPFV]-PST/NVIR-N REFL on four  
 ‘The match was divided into four.’ / ‘One divided the match into four.’
- b. Koledz-y dzieli-l-i zapałk-ę na czworo  
 colleague[M]-NOM.PL divide[IPFV]-PST.VIR-PL.VIR match-ACC on four  
 ‘The colleagues divided the match into four.’

This construction, as well as tokens with the reflexive clitic employed in the agreeing passive (in Czech and in older Polish), would have to be considered to be paradigmatic forms of the transitive verb, which do not change lexical meaning. Unfortunately, there is no way to formally distinguish these paradigmatic forms from occurrences in which the reflexive clitic marks lexical derivation, so that they could not be automatically retrieved. As a manual check of each token with Pol. *się* / Cz. *se* (lexical derivation or paradigmatic form?) proved unfeasible as well, we did not include these tokens in our set of samples.<sup>19</sup>

Bearing this in mind, let us now justify why we chose to contrast the IPFV1 *dzielić* / *dělit* with the IPFV2 *rozdzielać* / *rozdělovat* and not with IPFV2 with another prefix. A broader range of meaning alternations is typical of unprefixed verbs, thus also of the simplex (= IPFV1) in a potential triplet. Many simplex stems have more than one Natural Perfective (see Sect. 3), there is thus a potential for more than one triplet per simplex. Pol. *dzielić* can be considered to have two, maybe even more Natural Perfectives, each with its IPFV2; compare Pol. *dzielić* – *rozdzielić* – *rozdzielać* with *dzielić* – *podzielić* – *podzielać* and *dzielić* – *oddzielić* – *oddzielać* and their Czech cognates. However, the triplet with the prefix *roz-* can be considered to be the best choice for our purposes. While substantiating this choice, it should be emphasized that our considerations are not meant as a full-fledged analysis of the semantics and valency patterns of the involved stems; we just want to justify our choice in order to give an exemplary usage-based analysis. We have restricted ourselves to the relevant stems in Polish, since, apart from a few minor differences, the argument is the same for the Czech case.<sup>20</sup>

To begin with, contemporary *podzielić* is a very common NatPerf of *dzielić* with the meaning ‘divide, cut (into pieces)’, see (24)–(25):

- (24) Każda minuta ma swój sens w podtrzymywaniu porządku. Czas *dzieli* na części jak jabłko.  
 ‘Every minute has its sense in maintaining order. It *divides* the time *into parts* like an apple.’ (K. Kofta: Lewa, wspomnienie prawej. 2003)
- (25) Całość *podzieliłem* na cztery równe części.  
 ‘I *divided* the whole *into four equal parts*.’ (M. Sokołowski: *Gady*. 2007)

However, the contemporary IPFV2 *podzielać* is not used in this meaning; instead, it tends to be restricted to the meaning ‘share (sb’s opinion, fate)’, which implies human subjects and

<sup>19</sup>In Polish, there is another subject impersonal, which can be unambiguously identified from its morphology (namely, the suffix *-no / -to*). Such tokens were eliminated from the samples.

<sup>20</sup>We would like to thank Marek Łaziński for his support for this semantic analysis.

abstract objects (26). In this meaning it forms an aspect pair with PFV *podzielić* (27), while IPFV1 *dzielić* occurs in this meaning only rarely; see however (28):

- (26) Zastanawiała się, czy Hildegard *podziela* jej *sympatię* [...].  
 ‘She wondered if Hildegard *shared* her *sympathy* ...’  
 (E. Kujawska: *Dom Małgorzaty*. 2007)
- (27) Lambros Kastoriadis *podzielił los* wielu greckich rodzin, które przybyły do Polski, gdy skończyła się druga wojna światowa [...].  
 ‘Lambros Kastoriadis *shared the fate* of many Greek families who came to Poland after Second World War.’  
 (M. Czubaj: *21:37*. 2010)
- (28) Te *przekonania dzielił z* nim brat, malarz amator...  
 ‘His brother, an amateur painter, *shared these beliefs* with him.’  
 (K. Kolińska: *Orzeszkowa* ... 1996)

As a consequence, the meaning of the IPFV2 *podzielać* now differs from the IPFV1 *dzielić*. If there ever was a true triplet *dzielić – podzielić – podzielać*, it has now almost split into two aspect pairs with the PFV *podzielić* being polysemic:

- (29) *dzielić – podzielić*1                    ‘divide, cut’  
                   *podzielić*2 – *podzielać*    ‘share (opinion, conviction, etc., or fate).’

The first meaning of *podzielić* is realized in (30), in which it occurs like a synonym to *rozdzielić* used in the same context:

- (30) Niemcy—twierdzi Prodi—mur *rozdzielił* na dwa państwa, podczas gdy u nas *podzielił* naród na skazanych na pozostawanie na zawsze w opozycji komunistów i neofaszystów...  
 ‘Prodi says, that the wall *has divided* Germany into two parts, while here, it *has divided* the nation into communists and neo-fascists which are doomed to stay in in opposition forever ...’  
 (Gazeta Wyborcza. 1993)

Admittedly, there is a certain surplus of meaning contributed by *roz-*, which is difficult to explicate, but contrary to *podzielić*, *rozdzielić* implies that the unit divided should better be kept together. The simplex *dzielić* is rather indifferent to this extra feature, see (31), and this is what makes it convenient for different NatPerfs and a better ipfv partner in an aspect pair with *podzielić* than with *rozdzielić*. Since, however, the IPFV2 *podzielać* does not share this meaning (see above), there is no resulting triplet with the prefix *po-*.

- (31) [...] granica, która *dzieliła* oba światy, zacierała się lub wyostrzała.  
 ‘the border that *divided* both worlds became blurred or sharpened.’  
 (Gazeta Wyborcza. 1992)

Moreover, the stems prefixed with *roz-* are compatible with an indication of two opposite entities that are separated by a borderline, see the pattern in (33b), and this feature can be observed for *dzielić* as well; compare (31) and (32):

- (32) Jakże ciekawą kreską jest granica, która *rozdziela* wiedzę od mądrości.  
 ‘How subtle is the borderline that *divides* knowledge from wisdom.’  
 (Gazeta Radomszczańska. 2009)

Consequently, *rozdzielić – rozdzielać* has two basic valency patterns, which correlate with meaning differences:

- (33) a. *X rozdziela Y między Z / Z-om*. See (34)–(35).  
 Y = some substance or collection of (homogeneous) things.  
 Z = collective (usually human) referent which is the goal (or beneficiary) of the action;
- b. *X rozdziela Y od Z*. See (32).  
 Y and Z = two units that were conceived of as distinct already before the dividing action (Y from Z).
- (34) marynarz *rozdziela* żywność krajowcom.  
 ‘the seaman *distributes* food to the local inhabitants.’  
 (G. Methea: IV Rzeczpospolita. 2005)
- (35) [...] obie panie niosły jeszcze inne sztuki młodzieżowej garderoby, zgodnie *dzieląc* ciężar między siebie.  
 ‘both ladies carried other piece of youth clothing, *dividing* the burden between each other in harmony.’  
 (M. Musierowicz: Dziecko piątku. 1993)

There is a third pattern, much less frequent in contemporary speech (and the NKJP), which, from a semantic point of view, is somehow intermediate between the two aforementioned main patterns:

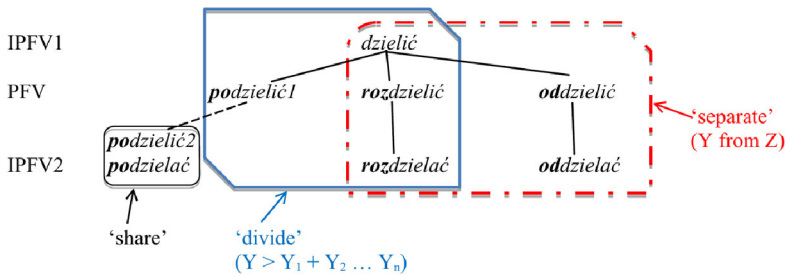
- (35c) *X rozdziela Y (na Z)*. (See ex. (36).)  
 Y = complex entity which by division becomes multiplied ( $Y > Y_1 + Y_2 \dots Y_n$ ).  
 This multiple entity can be indicated by Z.
- (36) [...] tu babcia wyciąga serwetkę, *rozdziela* ją na trzy warstwy i na jednej rysuje wypukłe czoło [...].  
 ‘here, my grandmother pulls out the a napkin, *divides* it into three layers and draws a convex forehead [...].’  
 (J. Dehnel: Lala. 2008)

For our analysis this intermediate usage was practically irrelevant. The second pattern (33b) makes *rozdzielać* similar to *oddzielać*, as in (37), and we can also find it with the simplex *dzielić* (see (38)):

- (37) Wspomniałem o rzece Jordan, która *oddziela od* siebie dwie grupy kościołów.  
 ‘I mentioned the Jordan River, which *separates* two groups of churches.’  
 (I. Karpowicz: Nowy kwiat cesarza (i Pszczoły). 2007)
- (38) Turystów *dzieli* od tubylców pięć metrów.  $\surd$ oddziela  
 ‘Tourists *are* five meters *away* from the locals.’ (lit. ‘Five meters *separate* the tourists from the locals.’)  
 (M. Przybyłek: Gamedec. Zabaweczki. Błyski. 2008)

In particular, *dzielić* and *oddzielać* easily replace each other if a distance is indicated that separates the two objects, as in (38). This observation applies to the Czech cognates as well:

- (39) Pod lodí se nachází hlubina Challenger, jejíž dno od hladiny *dělí* téměř 11 kilometrů ledové vody.  
 ‘Below the ship is the depth Challenger, whose bottom *is* 11 kilometers *away* from the surface of icy water.’  
 (Víkend HN. 2012)
- (40) Od japonského Hokkaida je *odděluje* poměrně úzký (15–25 km) Kunaširský průliv, Tanfil’jevův ostrov leží přímo před japonským mysem Nošappu.  
 ‘*Separated* by the narrow Kunashir Strait (15–25 km) from the Japanese Hokkaido, Tanfiljev island is located directly in front of Japan’s Cape Noshappu.’  
 (Respekt. 1991)



**Fig. 2** Meaning relations between Pol. *dzielić* and its closest synonyms

In contrast, *dzielić* can hardly replace *oddzielać* if the process (or act) of separation is focused on, as in the following example:

- (41) *Nacina i podważa skórę, słonię pomału oddziela od mięsa* [...]. ?dzieli.  
 ‘He cuts and gently pulls up the skin, slowly *separates* the fat from the meat.’  
 (B. Sławiński: *Królowa Tiramisu*. 2008)

These facts might be conditioned by the clear preference of Pol. *dzielić* / Cz. *dělit* for stative contexts (see Sect. 4.2).

Our survey of the lexical relations and syntactic patterns between *dzielić* and its closest prefixed synonyms (= NatPerfs) and imperfective derivatives of the latter ones (= IPFV2) can be summarized as in Fig. 2.

In sum: the triplet *dzielić* – *rozdzielić* – *rozdzielać* exhibits the largest meaning range; it overlaps with *podzielić1* for the meaning ‘divide’ and with *oddzielić* for the meaning ‘separate’. However, the former has an IPFV2, which is only marginally used in the same meaning, while the latter does have a frequent synonymous IPFV2, but for *oddzielić* – *oddzielać* the meaning distance from *dzielić* is slightly greater than for *rozdzielić* – *rozdzielać*. These conclusions drawn for Pol. *dzielić* / Cz. *dělit* and their possible triplets let us decide in favor of the triplet with the prefix *roz-*; but we also decided to analyze only those examples in which the meaning ranges of IPFV1 and IPFV2 overlap.

## 4.2 Data analysis

The tokens of Cz. *dělit* – *rozdělovat* and Pol. *dzielić* – *rozdzielać* were manually coded for the variables presented in (IVa) and for canonical functions of ipfv aspect listed in (IVb):

### (IVa) Grammatical and semantic features

- grammatical forms: tenses, moods, non-finite forms
- polarity:  $\pm$  negation
- subject / direct object: SG / PL
- form of object(s): NP, pronoun, PP
- referential status of arguments (subject, object): specific – generic
- ontological status of arguments: human, (other) animate, inanimate, abstract
- argument syntactically realized (excluding ellipsis)?: yes – no

### (IVb) Canonical functions of ipfv aspect

- progressive (processual), telic: yes / no / indefinite
- general-factual (existential: ‘at least once X’)

- iterative (only limited number of repetitions)
- habitual (unlimited iteration, regular or irregular)
- dispositional ('is able to S')
- two-way action
- stative

The variables in (IVa) contain information on grammatical forms of the verb stems, on polarity and properties of the arguments. Altogether 17 categorial variables were annotated describing syntactic and semantic features of the IPFV1 and IPFV2, plus three variables targeting grammatical forms (conditional, infinitive and participles). In addition, we accounted for the referential and ontological status of the arguments and their grammatical realization. It is these features which turned out to be more relevant than the paradigmatic forms of the verb or its polarity (see Sects. 4.2.2–4.2.3). Table 3 gives some examples illustrating how canonical aspect functions and referential and ontological status were coded.

**Table 3** Potentially relevant factors for the choice of IPFV 1 or IPFV2

Example	Canonical function of ipfv aspect	Ontological status / referential status
Czech		
(42) [...] čerpat finanční pomoc z operačního programu Životní prostředí, který <i>rozděluje</i> evropské peníze. 'draw financial help from the Operational Program Životní prostředí, which <i>distributes</i> European money.' (Dobry den s kurýrem. 2010)	habitual	subject: abstract / specific object: abstract / generic
(43) Státní peníze doposud <i>rozdělují</i> úředníci [...]. 'Officials still <i>distribute</i> money of the state.' (Mladá fronta DNES. 1993)	habitual	subject: human / generic object: abstract / generic
(44) Vztah k rostlinám <i>dělí</i> lidi na tři skupiny. 'The relationship to plants <i>divides</i> people into three groups.' (Doma DNES. 2012)	stative	subject: abstract / specific object: human / generic
(45) Svůj život <i>dělí</i> na život před dítětem a po něm. 'He <i>divides</i> his life into a period before and after the baby.' (Blesk pro ženy. 2012)	stative	subject: human / specific object: human / specific
(46) Svůj život <i>rozděluje</i> na dobu „před Keňou a po ní“. 'He <i>divides</i> his life into a period before visiting Kenia and after it.' (Blesk. 2011)	stative	subject: human / specific object: abstract / specific
(47) Já jsem nikdy zastupitele <i>nedělila</i> na muže a ženy. 'I have never <i>divided</i> representatives into men and women.' (i.e. 'the gender of the representatives was not important for me') (Deníky Moravia. 2004)	doubtful	subject: human / specific object: human / generic
(48) Analogie je pouto, které spojuje a současně <i>rozděluje</i> to, co je k sobě navzájem v analogickém vztahu. 'Analogy is a bond that at the same time unites and <i>divides</i> things which are in an analogous relation to each other.' (Lidové noviny. 1996)	stative	subject: abstract / specific object: abstract / generic

**Table 3** (Continued)

Example	Canonical function of ipfv aspect	Ontological status / referential status
Polish		
(49) Papieską pielgrzymkę od pierwszych letnich strajków roku 80 <i>dzielił</i> dokładnie rok. 'The papal pilgrimage was exactly one year away from the first summer strikes in 1980.' (A. Michnik, J. Tischner, J. Żakowski: Między panem a Plebanem. 1995)	stative	subject: abstract / specific object: abstract / specific
(50) Może <i>dzieliła</i> nas przepaść intelektualna? 'Maybe an intellectual abyss <i>separated</i> us?' (A. Bojarska, M. Bojarska: Siostry B. 1996)	stative	subject: abstract / doubtful object: human / specific
(51) W codziennej praktyce muzycznej <i>dzielimy</i> oktawę maksymalnie na 12 półtonów. 'In everyday musical practice, we <i>divide</i> the octave into a maximum of 12 semitones.' (K. Bilica: Wokół Chopina i Polski: siedem szkiców. 2005)	habitual	subject: human / generic object: abstract / generic
(52) Różnica stopy życia, ciężarów życia <i>rozdziela</i> nawet szczerze przyjaźnie. 'The difference in the standard of living and the burden of life <i>separates</i> (tears apart) even sincere friendships.' (A. Kamińska: Dzienniki 1927–1969. 2007)	habitual	subject: abstract / specific object: abstract / generic
(53) Miasta dawały większą szansę przeżycia, ponieważ tu <i>rozdzielano</i> pomoc zagraniczną. 'In the the cities the survival chances were bigger, because foreign aid was <i>distributed</i> here.' (R. Kapuściński: Heban. 2004)	habitual	subject: human / generic object: abstract / generic
(54) Ich sąsiadujące ze sobą domki <i>rozdziela</i> tylko linia kopalnianej wąskotorówki do przewozu towarów. 'Their neighboring houses were <i>separated</i> only by a mine narrow-gauge railway for the transport of goods.' (M. Szejnert: Czarny ogród. 2007)	stative	subject: abstract / specific object: inanimate / specific

Importantly, for any of these predefined categories tokens were labelled as 'doubtful' ('na' = 'not applicable' in the legends of Figs. 3 and 5 and Tables 4d, 5d) if no sufficiently clear decision could be made. For instance, in (47), determining the aspect function is troublesome: general-factual, habitual and stative function 'compete', since we cannot clearly establish whether the predication applies to a series of different, sufficiently distinct time intervals; the reason mainly lies in sentential negation (see Sect. 3.1). Another case in point is (50): since this is a yes / no-question (epistemically modified by *może* 'maybe'), the abstract subject-NP (*przepaść intelektualna* 'intellectual divide') is not even existentially quantified, we thus hesitated to assign it a specific status, even if the object-NP is specific. The case differs for the abstract subject-NPs in (52) and (54), which occur in assertions: in (52) the difference in life standards (*różnica stopy życia*) is referentially bound by the generic object-NP, while in (54) the abstract line is metonymically related to the mine narrow-gauge (*wąskotorówka*) and can be pointed at in real space.

One more remark is in order. In the chosen usage-based (bottom-up) approach we were not particularly interested in a clear-cut distinction between arguments and adjuncts. We therefore also took PPs whose status as argument realizations of the given verbs might be

debatable into account. In particular, this concerns PPs of the type Pol. *według*+genitive ‘according to’, which code the basis on which something is divided; for instance:

- (55) —Od dwóch lat dotacje dla uczelni podległych Ministerstwu Edukacji *rozdzielane są według specjalnego algorytmu.*  
 ‘—Since two years, subsidies for the universities reporting to the Ministry of Education *have been distributed* according to a special algorithm.’  
 (Gazeta Wyborcza. 1994)

Thus, overall, we started from a fine-grained grid. However, many variables proved irrelevant, so they subsequently were skipped or conflated. Among the canonical functions of ipfv aspect only those showed relevance that are not crossed out in (IVc):

(IVc): Canonical functions of ipfv aspect: sufficiently relevant functions

- progressive (processual), ~~telic / atelic / indefinite (as for telicity)~~
- general-factual (existential: ‘at least once X’)
- ~~iterative (only limited number of repetitions)~~
- habitual (incl. unlimited iteration)
- ~~dispositional (‘is able to X’)~~
- ~~two-way action~~
- stative

Overall, the habitual and the stative functions were the most frequent ones, the progressive and the general-factual functions were less frequent, other functions proved unimportant.

#### 4.2.1 Method

In order to check which factors affect the choice between IPFV1 and IPFV2, we applied binomial logistic regression, conditional inference trees and random forests. The two last methods are particularly well suited in cases of data sparsity and an interdependence of variables, which makes them a valuable addition to commonly used regression techniques. After running several trees, we selected the variables for our regression analysis. Then we created the random forest, allowing us to obtain conditional importance scores for the predictors (Levshina 2015, pp. 291–298) and the final tree plot that illustrates significant splits among variables.

#### 4.2.2 Analysis of Cz. *dělit* – *rozdělovat*

In the case of Czech, apart from relevant canonical aspect functions (see IVb), three very specific variables from the inventory of grammatical and semantic features (see IVa) proved relevant for the model selection: referential status of the subject, referential status of abstract objects expressed by a noun phrase (NP object), and referential status of the object expressed by a prepositional phrase (PP object); cf. Table 4a.

Estimate values below 0 indicate that the variable favors IPFV1 (*dělit*), and accordingly positive values enhance the chances for IPFV2 (*rozdělovat*). Table 4b shows goodness-of-fit statistics. The concordance index C measures the predictive accuracy of the model: as the C value approaches 1, the model becomes better at predicting the correct outcome. The value obtained for the present model (0.76) indicates that our model has an acceptable predictive power, with some field for potential improvements.

In order to check for multicollinearity in our data, we calculated the VIF-scores for each variable (Table 4c): low VIF-scores indicate that the assumption of no multicollinearity is



**Table 4a** Table of predictors for *Cz. dělit – rozdělovat*

Predictor	Estimate (log odds)
Intercept	-0.5047 ( $p = 0.3420$ )
Referential status of subject = specific referential status	0.9724 ( $p = 0.0183$ )
Referential status of subject = rg (generic referential status)	1.4655 ( $p = 0.0002$ )
Canonical function = habitual	0.5423 ( $p = 0.1091$ )
Canonical function = processual	1.1002 ( $p = 0.0256$ )
Canonical function = stative	-0.5436 ( $p = 0.1048$ )
Referential status of abstract NP object = specific referential status	-0.8504 ( $p = 0.0004$ )
Referential status of abstract NP object = generic referential status	-0.5601 ( $p = 0.0619$ )
Referential status of NP object = specific referential status	-0.7847 ( $p = 0.0148$ )
Referential status of NP object = generic referential status	0.2806 ( $p = 0.3208$ )
Referential status of PP object = specific referential status	0.6119 ( $p = 0.2171$ )
Referential status of PP object = generic referential status	-1.1642 ( $p < 0.0001$ )

**Table 4b** Summary statistics for *Cz. dělit – rozdělovat*

C	0.760
Dxy	0.519

**Table 4c** VIF-scores

Referential status of the subject = generic	5.153
Referential status of the subject = specific	5.16
Canonical function = habitual	4.92
Canonical function = processive	1.65
Canonical function = stative	5.00
Referential status of abstract NP object = generic	1.28
Referential status of abstract NP object = specific	2.93
Referential status of NP object = generic	2.36
Referential status of NP object = specific	3.55
Referential status of PP object = generic	1.10
Referential status of PP object = specific	1.10

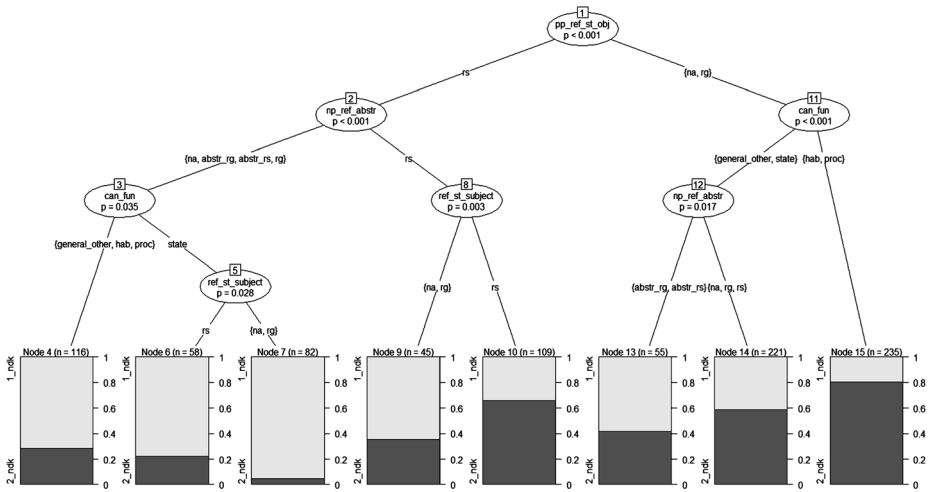
met. The VIF-scores are not higher than 10, some variables, however, exceed 5, which is considered a more conservative threshold.

The other assumptions of logistic regression were met as well. The database does not contain quantitative variables, so there is no need to test for linear relationships between the logits and the quantitative predictors.

Figure 3 presents the final tree plot with splits at significant levels. The names of the variables are explained in Table 4d.

The tree is analyzed according to the order of the nodes, so we will first comment on the left and the middle part (nodes 1–10).

At the bottom of the plot eight bins can be seen (nodes 4, 6, 7, 9, 10, 13, 14, and 15), each bin shows the proportions of IPFV1 (1\_ndk, *dělit*) and IPFV2 (2\_ndk, *rozdělovat*). Above each bin the number of observations which match the conditions of the splits is listed. For



**Fig. 3** Conditional inference tree: Czech

**Table 4d** Legend—variables and tags

Variable acronym	Variable	Tag	Explanation
pp_ref_st_obj	Referential status of the PP object	rs	Specific referential status
		rg	Generic referential status
		na	Not applicable, because there is no PP object
np_ref_abstr	Referential status and abstractness of NP object	rs	Specific referential status of non-abstract NP object
		rg	Generic referential status of non-abstract NP object
		abstr_rs	Specific referential status of abstract NP object
		abstr_rg	Generic referential status of abstract NP object
		na	Not applicable, because there is no NP object
can_fun	Canonical function	state	Stative
		hab	Habitual
		proc	Progressive
		general_other	Iterative and general-factive (with doubtful cases)
ref_st_subject	Referential status of the subject	rs	Specific referential status
		rg	Generic referential status
		na	Not applicable, because there is no (non-elliptical) subject

example, within node 4, 70% of the observations matching the conditions of the splits are contexts with IPFV1.

Surprisingly, the most important property is whether the verb occurs with a PP (56) or not (57):

- (56) Svůj život dělí na život před dítětem a po něm.  
‘He divides his life into a period before and after the baby.’

(Blesk pro ženy. 2012)

- (57) Obchod zastupitele *rozděloval*.  
 ‘The business *divided* the representatives.’ (Deníky Moravia. 2013)

The first split (node 1) concerns the referential status of the PP object. The left branch shows that contexts with specific referential status of PP objects (node 2) are split into two branches. The left one contains contexts with abstract NP objects with specific or generic referential status and non-abstract NP objects with generic status (node 3) as in (56) above.

Node 3 separates the contexts with habitual, progressive and other functions (node 4 with 116 observations) vs contexts with a stative function (node 5). Node 4 represents contexts such as in (58):

- (58) Dvanáct dní po povodni začne vedení Ústeckého kraje *dělit* finanční pomoc mezi ty, které vyplavila povodeň [...].  
 ‘Twelve days after the flood, the management of the Ústí nad Labem region will start *to divide* the financial aid among those who were washed out by the flood.’ (Mladá fronta DNES. 2009)

Node 5 is split into the contexts in which the referential status of the subject was specific (node 6 with 58 observations), as in (59):

- (59) [...] Zimbabwe povážlivě názorově *rozděluje* Afriku na straně jedné a vyspělé západní státy na straně druhé [...].  
 ‘[...] Zimbabwe judiciously *divides* Africa on the one hand and developed Western states on the other [...].’ (Hospodářské noviny. 2002)

Node 7 unites contexts in which the subject was either absent or has a generic referential status, as in (60):

- (60) Po 11. březnu terorismus už nelze *dělit* na mezinárodní a národní.  
 ‘After March 11, terrorism can no longer *be divided* into an international and a national one.’ (Respekt. 2004)

As the bins show, the majority of the observations on the left side of the tree concern the IPFV1 *dělit*.

Let us now take a closer look at the middle part of the plot, which branches to the right from node 2. Contexts with a specific referential status of a NP object can be found under node 8:

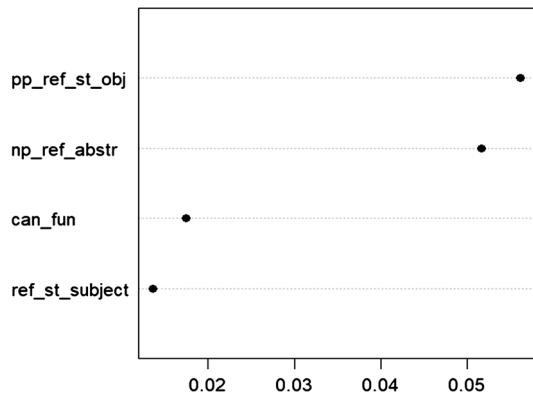
- (61) Vladimir Putin *dělí* své oponenty na poctivé opozičníky, kteří sice mají konstruktivní výhrady k režimu, ale milují svou vlast, a na zrádce národa, kteří Rusko nenávidí.  
 ‘Vladimir Putin *divides* his opponents into honest ones, who have constructive objections toward the regime, but love their homeland, and traitors to a nation who hate Russia.’ (Reflex. 2015)

These contexts split into two branches regarding the referential status of the subject. Node 9 contains 45 observations and represents cases in which a subject is absent (‘na’) or generic. In node 10, which contains 109 observations, one can find contexts with a specific referential status of the subject, as in (61) above.

The right branch of node 2 represents cases in which the object is either generic or absent (‘na’), as in (62):

- (62) Všechno, co nás *rozděluje* a sjednocuje.  
 ‘Everything that *divides* us and unites us.’  
 (from a heading—A. Lustig: Dům vrácené ozvěny. 1968)

**Fig. 4** Conditional importance of variables for Czech data



Let us now return to the top under node 1 and go from there to the right side of the tree. Node 11 separates tokens with a habitual or progressive function (node 15) from those with a stative function and others (node 12). Node 12 is split into contexts in which the referential status of the subject is generic or there is no subject at all (node 13 with 55 observations) and contexts with a specific referential status of the subject (node 14 with 221 observations); see, for instance, example (2) above.

Node 15 contains 235 observations, the majority of them are contexts with IPFV2 *rozdělovat*, as in (63):

- (63) Státní peníze doposud *rozdělují* úředníci a je to tedy závislé na jejich libovůli a na tom, kdo s kým je jaký kamarád.  
 ‘State money is still *distributed* by officials and this, thus, depends on their arbitrariness and who is friend with whom.’  
 (Mladá fronta DNES. 1993)

The C-score of the model equals 0.77 (see Table 4b). Thus, the presence or absence of PP objects underpins our model. IPFV1 prefers contexts with a PP. The interesting fact is that the separation of stative and habitual function appears twice, both splits make *dělit* in stative contexts more preferable. This suggests that the opposition may have a consistent effect across different contexts. The other relevant variable is referential status, but due to the conflation of the form of objects and its referential status we do not really know to what extent this is by chance or of more serious relevance.

Figure 4 presents the conditional importance scores of each of the variables included in the logistic regression model. The closer to 0 the less relevant the variable is.

#### 4.2.3 Analysis of Pol. *dzielić* – *rozdzielać*

After running several trees for the Polish data set we selected four variables: ontological status of the subject, canonical function, number of the NP object and ontological status of the PP object. This set of reference levels represents the most typical scenario for IPFV2 *rozdzielać*. Negative values increase the chance of IPFV1 occurring, which means that *dzielić* is more likely to appear in contexts with stative function and with inanimate PP objects (see Table 5a).

A high C-score indicates that our model discriminates well (Table 5b). Low VIF-scores indicate that the assumption of lacking multicollinearity is met (Table 5c).

Figure 5 presents the final tree plot with significant splits for the Polish database. The names of the variables are explained in the legend (Table 5d). Similar to the Czech tree, the Polish one is analyzed following the order of the nodes.

**Table 5a** Table of predictors for Pol. *dzielić – rozdzielać*

Predictor	Estimate ( <i>p</i> -value)
Intercept	2.4175 ( <i>p</i> < 0.0001)
Ontological status of the subject = human and animate	0.4782 ( <i>p</i> = 0.2092)
Ontological status of the subject = inanimate	-0.5621 ( <i>p</i> = 0.1531)
Canonical function = habitual	-0.1601 ( <i>p</i> = 0.5739)
Canonical function = stative	-0.7613 ( <i>p</i> = 0.0116)
Number of NP object = plural	-0.9380 ( <i>p</i> = 0.0021)
Number of NP object = singular	-1.5672 ( <i>p</i> < 0.0001)
Ontological status of PP object = human and animate	-1.1505 ( <i>p</i> = 0.0001)
Ontological status of PP object = inanimate	-1.5512 ( <i>p</i> < 0.0001)

**Table 5b** Summary statistics

C	0.795
Dxy	0.591

**Table 5c** VIF-scores

Ontological status of the subject = human and animate	3.45
Ontological status of the subject = inanimate	5.16
Canonical function = habitual	1.98
Canonical function = stative	2.63
Number of NP object = plural	2.71
Number of NP object = singular	2.46
Ontological status of PP object = human and animate	1.13
Ontological status of PP object = inanimate	1.30

The first split concerns contexts in which the ontological status of the PP object is animate or inanimate (node 2, ex. (64)) and contexts in which the ontological status of the PP object was marked as not applicable ('na', ex. (65)); see node 7:

- (64) [...] *sprawiedliwie dzielącym* ich łóżko na dwie połowy.  
 '[...] fairly *dividing* their bed into two halves.'

(H. Samson: *Pułapka na motyla*. 2000)

- (65) *Dzielily* nas schody.  
 'The stairs *separated* us.'

(K. Ostrowska: *Sny-klucz*. 2000)

This means that the proper contrast between the observations is the presence (64) vs absence (65) of a PP object.

Node 2 divides the contexts into those with stative function (node 3) and those with habitual, progressive and other functions (node 4). Node 3 contains 189 observations, the majority of them are contexts with IPFV1 *dzielić*, such as in (64)–(65) above. Node 4 separates tokens in the last-mentioned type of context with respect to the number of the object. Contexts with a singular object (see ex. (66)) are listed under node 5 with 41 observations:

- (66) Powinno być oczywiste, że partia nie *rozdziela* stanowisk pomiędzy swoich.

'It should be obvious that the party does not *distribute* the positions between its members.'

(Polityka. 2004)

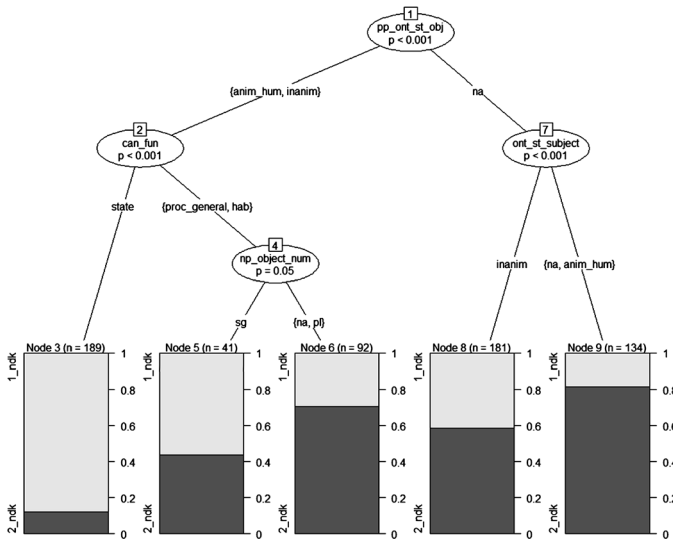


Fig. 5 Conditional inference tree: Polish

Table 5d Legend—variables and tags

Variable acronym	Variable	Tags	Explanation
pp_ont_st_obj	Ontological status of PP object	anim_hum	Animate or humane ontological status of the PP object
		inanim	Inanimate ontological status of the PP object
		na	Not applicable, because there is no PP object
can_fun	Canonical function	state	stative canonical function
		hab	habitual canonical function
		proc_general	conflated processive, generalfactive, iterative functions and doubtful cases
np_object_num	Number of the NP object	sg	singular
		pl	plural
		na	not applicable, because there is no NP object
ont_st_subject	Ontological status of the subject	anim_hum	Animate or humane ontological status of the subject
		inanim	Inanimate ontological status of the subject
		na	Not applicable, because there is no subject

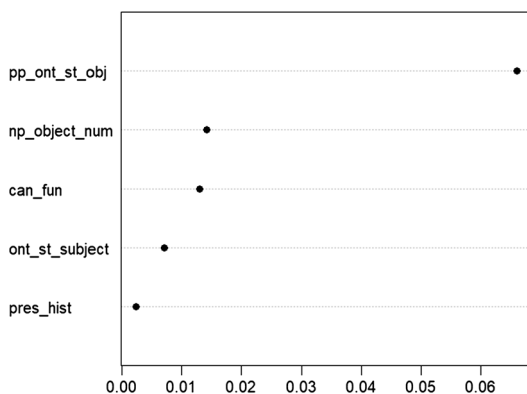
Node 6 contains 92 observations without a subject or with a plural subject as in sentence (67):

(67) [...] które będą *dzielić* różne dobra pomiędzy różnych ludzi, na różny sposób i z różnych powodów [...].

‘[...] which will *distribute* the different goods between different people in different ways and for different reasons.’

(E. Wnuk-Lipiński: *Demokratyczna rekonstrukcja: z socjologii radykalnej zmiany społecznej*. 1996)

**Fig. 6** Conditional importance of variables for Polish data



Let us turn now to the right branch under node 1. Node 7 separates observations according to the ontological status of the subject. The left branch represents the contexts with inanimate subjects (node 8). It contains 181 observations and represents contexts such as in (65) above. The right branch contains contexts with animate (incl. human) subjects or without a subject (node 9). Here, one finds contexts as in (68):

- (68) *Dzieliłem oglądane obrazy metodologicznie (...).*  
 'I divided the pictures up methodologically.'  
 (M. Miller: *Pierwszy milion czyli Chłopcy z Mielczarskiego*. 1999)

It is worth noting that the majority of contexts without a PP object concern the IPFV2 *rozdzielać*, which is similar to the results from the analysis of the Czech cognates.

The C-score of the model equals 0.795. The absence or presence of a PP object is the basis of the model. In general, when a PP object is involved, the IPFV1 *dzielić* is preferred. The stative function clusters against the habitual and progressive functions. As for the Czech cognates, the stative function makes IPFV1 more likely.

The conditional importance scores for the variables are presented in Fig. 6; the C-score is slightly higher for the random forest model ( $C = 0.83$ ).

### 4.3 Discussion of results

Let us sum up this case study. First of all, diachrony (from 1750 till today) has turned out to be no significant factor; in other words: no discernable changes in functional distribution have occurred over time. Moreover, there is hardly any difference in the distributional properties of the cognate ipfv stems in Czech and Polish, and in neither language do different grammatical forms (inflections of finite and non-finite forms) have any seizeable impact on the choice of IPFV1 or IPFV2. That is, variation among those kinds of complex word forms which are traditionally considered to be members of a verb paradigm (including infinitives and participles), do not play any palpable role in the choice of the morphologically related and synonymous ipfv stems.

Turning to the other variables, we have to admit that the semantic variables which were conflated with the syntactic variables, such as NP object and PP object, may blur the results a bit, but presence vs absence of a PP object made the differences between the two ipfv stems more prominent. The presence of a PP object increases the chances of IPFV1 in both languages. This is indicative of a salient relationship between the respective ipfv stem and verb valency and/or clause syntax.

Moreover, in both languages habitual and progressive functions are grouped against the stative function. The latter makes IPFV1 (Pol. *dzielić* / Cz. *dělit*) more likely. Therefore, at least at this very coarse level, the distribution of canonical functions of the ipfv aspect clearly tends toward complementary sharework between IPFV1 and IPFV2, which to some extent supports the ‘division of labour’ hypothesis between two ipfv stems.

At the same time, more research is necessary to better understand the factors behind the usage of IPFV1 and IPFV2 forms, as well as to obtain broader generalisations. Further research should most certainly involve a more fine-grained treatment of information on valency and clausal syntax, with a focus on the association of IPFV1 and IPFV2 with specific types of PPs. Of course, this kind of information is at best indirectly related to distinctions in terms of canonical aspect functions.

The significance of these interim conclusions based on the functional distribution of IPFV1 and IPFV2 is not immediately evident. On the one hand, the fact that syntactic valency (presence/absence of a PP, i.e. a peripheral argument) and referential (Czech) or ontological (Polish) properties of the object NP have greater predictive power than canonical functions may be taken as indicative that IPFV1 and IPFV2 prove more easily substitutable for each other in terms of typical functions of ipfv aspect irrespective of alternations of lexical meaning. The two ipfv stems can thus be considered as really equivalent partners of the pfv stem. On the other hand, among the canonical functions of ipfv aspect, we at least observed a clear split between the stative function and the rest (in favor of IPFV1), and this lends support to the assumption that the two ipfv stems tend toward complementary distribution among canonical aspect functions. In addition, a different preference for the stative vs habitual function can be interpreted as showing that IPFV2, i.e. the prefixed stem, is better suited for situations whose internal temporal structure can be seen as fragmented into distinct subintervals, while IPFV1 prefers situations that are void of such fragmentation. Whether this distributional split is ultimately conditioned by the prefix, remains an open question. The role of the prefix might likewise be connected to the different preferences in the coding of arguments and adjuncts.

## 5 Conclusions and outlook

The conclusions drawn above are based on an in-depth study for just one case. So, what can these findings tell us about the Polish and Czech aspect systems? It is certainly unjustified to speculate about a particular part of that system, namely the role of aspect triplets, *in toto*. Systematic studies concerning this issue are still lacking; those studies which have raised this issue are practically exclusively concerned with contemporary Russian and have rather downplayed the significance of triplets for the entire system.<sup>21</sup> Regardless of this, it is evident that triplets do not form a homogeneous class, either in terms of the range of alternations of lexical meaning between their members or in terms of aspect functions. Therefore, generalizations based on just one case study, even if a central representative of this class was chosen, would be premature.

Curiously, our findings allow us to say more about the paradigmatic nature of the involved stems in relation to the entire aspect system, exactly because differences in forms typically associated with inflection (tense, person-number) or with operations that change the syntactic class (participles etc.) have turned out irrelevant for the choice of IPFV1 vs IPFV2. These distinctions of complex word forms are highly regular and they are virtually not accompanied

<sup>21</sup> Apart from Janda et al. (2013, §6) cf. Apresjan (1988, 1995), Zaliznjak et al. (2015, pp. 228–247), also the discussion in Wiemer (2017, pp. 235–244, 2019b, pp. 49–56).



by even slight alternations in lexical meaning. We know this to be the case not only for the units from our case study, but presume this to be a general property of Polish and Czech (or Russian) verbs for system-internal reasons which were explained in Sects. 2–3. This knowledge about the regularity and the usual insensitivity to lexical meaning ascribed to ‘typical inflection’ justifies a more general claim concerning the paradigmatic organization of stems related by derivational affixes that just change aspect membership, but do not alter lexical meaning. Our usage-based analysis of a core representative of aspect triplets showed that the involved ipfv stems are insensitive to typically inflectional properties, but they partially complement each other when it comes to functional distinctions relevant for actionality and to preferences in the coding of arguments and adjuncts.

That being said, let us return to the notion of paradigm. A paradigm is usually assumed as a set of word forms subdivided according to criteria which are considered to distinguish finite and non-finite forms. However, paradigms also imply that the word forms match certain combinatorial restrictions in predictable ways. Normally such constraints are couched in syntactic terms and concern different levels of constituency (compare, for instance, limitations in choosing pfv vs ipfv verbs in the scope of phasal or modal verbs), but they should be extended to constraints concerning categorial distinctions as, for instance, the distribution of pfv and ipfv stems in negated and unnegated imperatives (or equivalent constructions), their use in converbs (adverbial participles), and constructions marking the future, passives or other agent-demoting operations. These constraints are more or less well-described in the aspectological literature, but, here, we have already entered into a third dimension which—apart from distinctions of word forms and their combinability with other word forms and constituents—concerns the inventory of aspect functions (considered canonical for a particular Slavic language). These functions show up particularly when the choice between pfv and ipfv stem offers some “leeway”, i.e. none of them is strictly “forbidden” (ungrammatical), so that relative freedom of choice triggers differences in interpretation (according to the levels mentioned in §2). This creates functional distinctions for which choice of aspect often provides minimal pair conditions.

Therefore, if the general idea behind paradigms is based not only on the morphological regularity of patterns, but also on the predictability with which variation of form corresponds to variation in meaning, and which simultaneously guarantees the identity of lexical units (i.e. of lexical meaning and the range of meaning alternation), the Slavic PFV : IPFV opposition requires us to account for three building blocks:

- (i) inventory of (finite and non-finite) word forms,
- (ii) set of their combinatorial restrictions,
- (iii) function inventory (partially following from the relation between (i) and (ii)).

In order to do justice to the distribution of ipfv and pfv stems considered to represent an identical lexical concept we have to capture the replacement conditions between these stems and their overlaps; this also concerns possible ‘variants’ in form of two (or more) ipfv stems that kind of compete with regard to the building blocks (ii) and (iii). Our account of the paradigm of aspect choice has to include this competition, in analogy to what “usual” inflectional paradigms reveal as (free, arbitrary or meaningful) variation in the phonological realization of, say, case distinctions. Compare, for instance, the choice between {a} vs {u} in the genitive singular of masculine nouns; in Russian this choice more often than not is meaningful: {u} indicates indeterminate quantity; compare, e.g., (*vkus*) *čaj-a* ‘(taste of) tea’ vs (*nalil*) *čaj-u* ‘(he poured in some) tea’. Its status thus differs from the superficially identical choice in Polish, where this distribution is not governed by some sufficiently reliable and, thus, predictable semantic distinctions.

Such analogies could be considerably extended, but for the point to be made here it suffices to emphasize that nothing, in principle, inhibits us to revise the notion of paradigm in such a way that it fits the requirements of the morphological ‘outfit’ of the Slavic aspect opposition which, among other things, contains lots of triplets. Moreover, nothing indicates that their amount has been decreasing, nor does a theoretical account of the system require that a decrease must, in principle, take place.<sup>22</sup> The issue is rather whether we can observe changes in distribution between competing stems and what speakers make out of them: are they discriminated meaningfully, and does this discrimination add up to some systematicity? Only a systematic token-based investigation would enable us to pinpoint a more general tendency to complementary distribution, it cannot be discerned with the naked eye. As always with distributions, we can give more or less informed guesses, but distributions cannot be stated just by pointing at a few examples (and counterexamples). Establishing them in an objectifiable way is a very time-consuming enterprise (even with sufficiently well-annotated corpora). For research practice this means that the properties of the system in terms of a comprehensive notion of paradigm based on choice of aspect and stem type can hardly be checked in larger parts, let alone in its entirety. These are, as it were, the physical restrictions of research, but they are not an argument against revising the notion of paradigm in the sense as it has been advocated for here.

We need not be pessimistic though, even if a usage-based approach to testing hypotheses about the behavior of aspect triplets—and thus, indirectly, about the consistency of the aspect system—is very labor-intensive. Analyses like the one we carried out here could possibly be performed for a manageable amount of triplets after possible candidates have been identified. Core representatives (as in the case analyzed above) should be contrasted with cases for which semantic identity and/or diachronic stability are less clear, and it is probably justifiable to skip ‘usual’ inflectional distinctions such as tense or person-number, at least for the core representatives, since these distinctions appear to be non-discriminative for the choice of IPFV1 vs IPFV2.

## Sources

### Czech

ČNK—Czech National Corpus (*korpus syn7* and *Diakorp 6*): <https://korpus.cz/>.

Křen, M., Cvrček, V., Čapka, T., Čermáková, A., Hnátková, M., Chlumská, L., Jelínek, T., Kovářiková, D., Petkevič, V., Procházka, P., Skoumalová, H., Škrabal, M., Truneček, P., Vondříčka, P., Zasina, A.: *Korpus SYN, verze 7 z 29. 11. 2018*. Ústav Českého národního korpusu FF UK, Prague 2017. Available at <http://www.korpus.cz>.

Kučera, K., Řehořková, A., Stluka, M.: *DIAKORP: Diachronní korpus, verze 6 z 18. 12. 2015*. Ústav Českého národního korpusu FF UK, Prague 2015. Available at <http://www.korpus.cz>.

### Polish

Electronic corpus of 17th and 18th century Polish texts (*KorBa*): <http://clip.ipipan.waw.pl/KORBA>.

<sup>22</sup>Dwelling upon the question of how to count triplets and how their inventory might have changed through time would exceed the range of this contribution. This issue does not differ much from the notorious question of how to count aspect pairs. The crucial points for triplets are (a) how to establish Natural Perfectives and (b) how productive is secondary imperfectivization (see Sect. 3.2). Slavic languages seem to differ particularly with respect to the latter.

NKJP—Polish National Corpus: <http://nkjp.pl/>.

Corpus of 12 mln tokens (texts from 1750–1917) compiled for *DiAsPol*.

Wikisource: [https://pl.wikisource.org/wiki/Wikiźródła:Strona\\_główna](https://pl.wikisource.org/wiki/Wikiźródła:Strona_główna).

Wolne Lektury: <https://wolnelektury.pl/>.

**Funding Note** Open Access funding enabled and organized by Projekt DEAL.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Apresjan, Ju. D. (1988). Morfoložičeskaja informacija dlja tolkovogo slovarja. In Ju. N. Karaulov (Ed.), *Slovarnyje kategorii. Sbornik statej* (pp. 31–53). Moskva.
- Apresjan, Ju. D. (1995). Traktovka izbytočnyx aspektual'nyx paradigm v tolkovom slovarje. In Ju. D. Apresjan, *Izbrannye trudy. Tom II: Integral'noe opisanie jazyka i sistemnaja leksikografija* (pp. 102–113). Moskva.
- Benacchio (2004); Benackio, R. Glagol'nyj vid v imperative v južnoslavjanskix jazykax. In Ju. D. Apresjan (Ed.), *Sokrovennye smysly. Slovo. Tekst. Kul'tura. Sbornik statej v čest' N. D. Arutjunovoj* (pp. 267–275). Moskva.
- Breu, W. (1984). Grammatische Aspektkategorie und verbale Einheit. In W. Girke & H. Jachnow (Eds.), *Aspekte der Slavistik. Festschrift für Josef Schrenk* (Slavistische Beiträge, 180, pp. 7–25). München.
- Breu, W. (2000). Zur Position des Slavischen in einer Typologie des Verbalaspekts (Form, Funktion, Ebenenhierarchie und lexikalische Interaktion). In W. Breu (Ed.), *Probleme der Interaktion von Lexik und Aspekt (ILA)* (Linguistische Arbeiten, 412, pp. 21–54). Tübingen.
- Dickey, S. M. (2000). *Parameters of Slavic aspect. A cognitive approach*. Stanford.
- Divjak, D., & Gries, S. T. (2006). Ways of trying in Russian: clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory*, 2–1, 23–60. <https://doi.org/10.1515/CLLT.2006.002>.
- Geeraerts, D., Grondelaers, S., & Bakema, P. (1994). *The structure of lexical variation. Meaning, naming, and context* (Cognitive Linguistics Research, 5). Berlin, New York. <https://doi.org/10.1515/9783110873061>.
- Glynn, D. (2010). Testing the hypothesis. Objectivity and verification in usage-based Cognitive Semantics. In D. Glynn & K. Fischer (Eds.), *Quantitative methods in Cognitive Semantics. Corpus driven approaches* (Cognitive Linguistics Research, 46, pp. 239–269). Berlin, New York.
- Isačenko, A. V. (2003). *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim. Morfoložija, čast' vtoraja*. Bratislava.
- Janda, L. A. (2007). Aspectual clusters of Russian verbs. *Studies in Language*, 31(3), 607–648.
- Janda, L. A., & Lyashevskaya, O. (2011). Grammatical profiles and the interaction of the lexicon with aspect, tense, and mood in Russian. *Cognitive Linguistics*, 22(4), 719–763. <https://doi.org/10.1515/COGL.2011.027>.
- Janda, L. A., et al. (2013). *Why Russian aspectual prefixes Aren't Empty. Prefixes as verbal classifiers*. Bloomington.
- Klavan, J. (2014). A multifactorial corpus analysis of grammatical synonymy. The Estonian adessive and the adposition *peal* 'on'. In D. Glynn & J. Robinson (Eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy* (Human Cognitive Processing, 43, pp. 253–278). Amsterdam, Philadelphia.
- Łaziński, M. (2020). *Wykłady o aspekcie polskiego czasownika*. Warszawa.
- Lehmann, V. (1988). Der russische Aspekt und die lexikalische Bedeutung des Verbs. *Zeitschrift für slavische Philologie*, 48(1), 170–181.
- Lehmann, V. (1999). Aspekt. In H. Jachnow (Ed.), *Handbuch der sprachwissenschaftlichen Russistik und ihrer Grenzdisziplinen* (pp. 214–242). Wiesbaden.

- Levshina, N. (2015). *How to do Linguistics with R. Data exploration and statistical analysis*. Amsterdam, Philadelphia.
- Maslov, Ju. S. (2004[1948]). Vid i leksičeskoe značenje glagola v sovremennom ruskom literaturnom jazyke. In Ju. S. Maslov, *Izbrannye trudy. Aspektologija. Obščee jazykoznanie* (pp. 71–90, reprinted from: *Izvestija AN SSSR*, t. 7, vyp. 4, 303–316). Moskva.
- Mattiola, S. (2019). *Typology of pluractional constructions in the languages of the world* (Typological Studies in Language, 125). Amsterdam, Philadelphia.
- Mende, et al. (2011): Mende, Ju., Born-Rauxeneker, E., Brjugeman, N., Dipong, X., Kukla, Ju., & Leman, F., *Vid i akcional'nost' russkogo glagola. Opyt slovarja* (Slavolinguistica, 14). München.
- Plungjan, V. A. (2000). *Obščaja morfologija. Vvedenie v problematiku*. Moskva.
- Plungjan, V. A. (2011). *Vvedenie v grammatičeskiju semantiku: grammatičeskie značenja i grammatičeskie sistemy jazykov mira*. Moskva.
- PWN Oxford: J. Linde-Usiekniewicz (Ed.) (2004). *Wielki słownik polsko-angielski PWN-Oxford*. Warszawa.
- Šluinskij, A. B. (2006). K tipologii predikatnoj množestvennosti: organizacija semantičeskoj zony. *Voprosy jazykoznanija*, 1, 46–75.
- Tatevosov, S. G. (2015). *Akcional'nost' v leksike i grammatike. Glagol i struktura sobytija*. Moskva.
- Tatevosov, S. G. (2016). *Glagol'nye klassy i tipologija akcional'nosti*. Moskva.
- van Schooneveld, C. H. (1958). The so called 'préverbes vides' and neutralization. In *Dutch contributions to the Fourth International Congress of Slavists. Moscow, September 1958* (Slavistische Drukken en Herdrukken, pp. 159–161). 'S-Gravenhage.
- Vey, M. (1952). Les préverbes 'vides' en tchèque moderne. *Revue des Études Slaves*, 29(1–4), 82–107.
- Wiemer, B. (2006). Vimer, B., O razgraničenii grammatičeskix i leksičeskix protivopostavljenij v glagol'nom slovoobrazovanii ili: čemu mogut naučit'sja aspektologi na primere *sja-glagolov*? In V. Lehmann (Ed.), *Glagol'nyj vid i leksikografija. Semantika i struktura slavjanskogo vida IV* (Slavolinguistica, 7, pp. 97–123). München.
- Wiemer, B. (2008). Zur innerslavischen Variation bei der Aspektwahl und der Gewichtung ihrer Faktoren. In S. Kempgen, K. Gutschmidt, U. Jekutsch, & L. Udolph (Eds.), *Deutsche Beiträge zum 14. Internationalen Slavistenkongress. Ohrid 2008* (Die Welt der Slaven. Sammelbände, 32, pp. 383–409). München.
- Wiemer, B. (2015). O roli vida v oblasti kratnosti i pragmatičeskix funkcij (ėskiz s točki zrenija xronotopii). In R. Benacchio (Ed.), *Glagol'nyj vid: grammatičeskoe značenje i kontekst. Verbal aspect: Grammatical meaning and context* (Die Welt der Slaven. Sammelbände, 56, pp. 585–609). München.
- Wiemer, B. (2017). O roli pristinok i suffiksov na rannix i pozdnix etapax istorii slavjanskogo vida. In R. Benacchio, A. Muro, & S. Slavkova (Eds.), *The role of prefixes in the formation of aspectuality. Issues of grammaticalization* (pp. 219–253). Florence.
- Wiemer, B. (2019a). Tipologija akcional'nosti: svojstva finitnoj klauzy, klassifikacija glagolov i edinyj podxod k slovoizmenitel'nomu i derivacionnomu vidu. *Voprosy jazykoznanija*, 1, 93–129. (Review article on: S. G. Tatevosov, *Akcional'nost' v leksike i grammatike (Glagol i struktura sobytija)*. Moskva 2015, ISBN 978-5-9905762-5-4 and S. G. Tatevosov, *Glagol'nye klassy i tipologija akcional'nosti*. Moskva 2016, ISBN 978-5-94457-238-7). <https://doi.org/10.31857/S0373658X0003595-3>.
- Wiemer, B. (2019b). O semantičeski invariantnom i grammatičeski trivial'nom v ruskom vide In D. V. Gerasimov, S. Ju. Dmitrenko, & N. M. Zaika (Eds.), *Sbornik statej k 85-letiju V. S. Xrakovskogo* (pp. 43–66). Moskva.
- Wiemer, B. (2020). O sinonimax i suppletivizme neskol'ko inače, to est': s točki zrenija osnov. In L. L. Iomdin et al. (Eds.), *Ot semantičeskix kvarkov do vselejnoj v alfavitnom porjadke. Sbornik statej k 90-letiju Jurija Derenikoviča Apresjana. / From Semantic Quarks to the Universe in Alphabetical Order. A Festschrift to commemorate the 90th Anniversary of Academician Yuri Derenikovich Apresjan* (Trudy Instituta russkogo jazyka im. V. V. Vinogradova, XXIV, pp. 146–169). Moskva.
- Wiemer, B., & Seržant, I. A. (2017). Diachrony and typology of Slavic aspect: What does morphology tell us? In W. Bisang & A. Malchukov (Eds.), *Unity and diversity in grammaticalization scenarios* (Studies in Diversity Linguistics, 16, pp. 239–307). Berlin. <https://doi.org/10.5281/zenodo.823246>.
- Wood, E. J. (2007). *The semantic typology of pluractionality* (unpubl. PhD thesis, University of California). Berkeley. Retrieved from <https://scholarship.org/uc/item/2hc3b5pb> (24 September 2020).
- Zaliznjak, A. A., & Šmelev, A. D. (1997). *Lekcii po ruskoj aspektologii* (Slavistische Beiträge, 353. Studienhilfen, 7). München.
- Zaliznjak, A. A., Mikaëljan, I. L., & Šmelev, A. D. (2015). *Russkaja aspektologija: v zaščitu vidovoj pary*. Moskva.