UNIVERSITY OF BIRMINGHAM

On the nature and organisation of morphological categories

Divjak, Dagmar; Testini, Irene; Milin, Petar

On the nature and organisation of morphological categories: verbal aspect through the lens of associative learning

Dagmar Divjak[1,2] · Irene Testini[1] · Petar Milin[1]

Abstract

The process by which awareness and/or knowledge of linguistic categories arises from exposure to patterns in data alone, known as emergence, is the corner stone of usage-based approaches to language. The present paper zooms in on the types of patterns that language users may detect in the input to determine the content, and hence the nature, of the hypothesised morphological category of aspect.

The large-scale corpus and computational studies we present focus on the morphological encoding of temporal information as exemplified by aspect (imperfective/perfective) in Polish. Aspect is so heavily grammaticalized that it is marked on every verb form, yielding the practice of positing infinitival verb pairs ('do' = 'robićimpf/zrobićpf') to represent a complete aspectual paradigm. As has been shown for nominal declension, however, aspectual usage appears uneven, with 90% of verbs strongly preferring one aspect over the other. This makes the theoretical aspectual paradigm in practice very gappy, triggering an acute sense of partialness in usage. Operationalising emergence as learnability, we simulate learning to use aspect from exposure with a computational implementation of the Rescorla-Wager rule of associative learning. We find that paradigmatic gappiness in usage does not diminish learnability; to the contrary, a very high prediction accuracy is achieved using as cues only the verb and its tense; contextual information does not further improve performance. Aspect emerges as a strongly lexical phenomenon. Hence, the question of cognitive reality of aspectual categories, as an example of morphological categories in general, should be reformulated to ask which continuous cues must be learned to enable categorisation of aspectual outcomes. We discuss how the gappiness of the paradigm plays a crucial role in this process, and how an iteratively learned, continuously developing association presents a possible mechanism by which language users process their experience of cue-outcome co-occurrences and learn to use morphological forms, without the need for abstractions.

Keywords Verbal aspect · Paradigm structure · Emergence · Associative learning · Polish

---

Extended author information available on the last page of the article

Springer

1 Introduction

The idea of emergence of structure from exposure to usage, which is core to usage-based approaches to language, is attractive, yet remains open to criticism. Despite its popularity, no procedure has been proposed that could be used to test the process by which full-fledged abstract categories emerge from exposure to usage on authentic language data. In this paper, we begin to address this challenge by tackling the first step in the process of emergence: we focus on the input that language users receive to identify the building blocks, and hence the nature, of morphological categories.

Our case study zooms in on the morphological encoding of temporal information, and more in particular on the category of aspect, as exemplified by Polish. Slavic aspect constitutes a notorious challenge for linguists, in particular in terms of its use. Even though aspect has been studied extensively, and existing approaches have made substantial contributions to our understanding of the category of aspect, a comprehensive model of aspect that would capture aspectual use has remained out of reach. Following Divjak et al. (2021) we consider learnability from exposure to raw data as precondition for the cognitive plausibility of any posited abstract structures. Using a computational algorithm that is informed by research on animal and human learning (Baayen et al., 2011) we investigate what patterns linguistic categories *might* contain if they emerge from usage and what information they *must* contain for them to be useful in guiding usage. This knowledge allows us to contrast the "in-principle" system of morphological theory, which includes all that is possible, with the "in-practice" system attested in actual usage, which highlights what is probable.

In what follows, we will first briefly introduce the theoretical discussions with which this study engages (Sect. 1.1), and survey existing theoretical accounts of Slavic aspect (Sect. 1.2) as well as research that has examined the mental reality of the proposed categories (Section 1.3) before moving onto a description of the current study (Section 1.4). We will then demonstrate how a computational approach, relying on algorithms that are grounded in knowledge about learning, can address theoretical linguistic questions and improve the description of language as it is used.

1.1 Emergence from use: can the complete arise from the partial?

Emergence from usage forms the corner stone of usage-based approaches to language. It is the process by which awareness and/or knowledge of linguistic structures arises from exposure to patterns in data alone (Langacker, 1987a): language users gradually build their linguistic system by accumulating and organising experience with the input they receive. Once patterns are established, categories of linguistic knowledge can be formed through the processes of generalisation and abstraction over patterns. Inflectional paradigms, such as nominal declensions and verbal conjugations, represent an interesting case as they illustrate the emergence of a whole cluster of related forms, each of which specifies a particular function/meaning (for a recent discussion see Blevins, 2016). For example, the Latin nominal declension distinguishes five cases across two grammatical numbers, typically presented as a 5×2 grid of word forms, some cells of which contain forms that are identical, i.e., syncretic.

Morphological word-and-paradigm models (e.g., Matthews, 1965, 1991; Blevins, 2006, 2016) agree with usage-based theories (e.g., Bybee, 1988, 1995; Langacker, 1988, 2018) that paradigmatic patterns do not reduce to static morphological representations that are combined through productive rules. For usage-based theories, morphological patterns and paradigms have the same theoretical status as words and the mental lexicon consists of words that "are related to other words via sets of lexical connections between identical and similar phonological and semantic features" (cf. Bybee, 1995, p. 428). For word-and-paradigm models, words are not only related but are also the basic elements of analysis (cf., Blevins, 2016), and word paradigms receive the status of natural linguistic units, just like words (cf., Matthews, 1991).

However, input is not uniform for all lexemes. It is well-known that the frequency distribution of forms in language is rather skewed, with a minority of forms making up the bulk of language use (Zipf, 1949). Because of this, learners do not encounter all possible forms of every lexeme, which makes it difficult to state that they have complete paradigms available for every lexeme. This is particularly acute in morphologically rich languages where some forms may not be attested at all, e.g., due to the semantic incompatibility between the noun's lexical semantics and the functions or meanings associated with the paradigm cell. The concept of *paradigm size*, i.e., the number of attested forms of a given lexeme, has long occupied psycholinguists (for example, Traficante & Burani, 2003; see also Bertram et al., 2000 and Moscoso del Prado Martín et al., 2004a). Measures have been developed to capture the skewness of an inflected form's distribution in actual usage: inflectional entropy quantifies the degree of unevenness in the frequency of a lexeme's inflected forms (Moscoso del Prado Martın et al., 2004b), while inflectional relative entropy expresses the typicality of said unevenness with respect to the whole inflected class to which a given lexeme belongs (Milin et al., 2009). Although missing inflected forms are quite common, language users can usually supply forms that are not attested for one lexeme, based on similarity to attested forms of other lexemes. The ability to apply proportional analogy is often argued to be the mechanism that users rely upon when producing unseen forms (for a first demonstration and use of the Wug Test see Berko, 1958). This hypothesised cognitive mechanism appears as rather reliable since naïve language users can apply it even to flex nonce words that are morphotactically compliant (see Jovanović et al., 2008; also see Milin et al., 2011).

It has been suggested that inflectional paradigms such as verbal conjugations and nominal declensions have set-theoretic properties (Stump & Finkel, 2013), or schematicity (Booij, 2010, 2017) and, because of that, it has been argued that they should be considered as natural linguistic units (cf., Matthews, 1991; Blevins, 2006, 2016). Such a position might be taken to suggest the neatness and completeness of these units (see Ackerman et al., 2009). At the same time, paradigm defectiveness (Sims, 2015) is well-documented in the literature (in particular, see Bonami & Beniamine, 2016). Does usage of inflected forms, then, lead to contrasting "forces", i.e., a sense of completeness on the one hand, and an experience of partialness of attestation on the other hand (cf., Blevins et al., 2017)?

While the cell-filling problem has been discussed at length for nominal inflection, it was not standardly considered for verbal inflection. Nevertheless, recent work shows that verbal inflectional paradigms in morphologically rich languages are also

gappy. For example, Ackerman et al. (2009) mention verbal systems in languages such as Tundra Nenets or Mordvin, which due to their inherent complexity have a considerable number of unattested cells (see also further work in Ackerman & Malouf, 2013). Stoll et al. (2012) describe how the language of very young speakers of Chintang is characterised by a significantly higher noun-to-verb ratio than the language of adults because of the immense complexity of Chintang verbal morphology. This phenomenon is not restricted to languages with small pools of speakers: similar findings have been reported for Finnish (Karlsson, 1985, 1986) and even for widely spoken languages such as Russian (Lyashevskaya, 2013), for which all open word classes have been claimed to exhibit this behaviour (Janda & Tyers, 2021). These findings suggest that paradigms, and complete paradigms in particular, are a theoretical construct more than a descriptive reality, not only for nouns, but also for verbs and the so-called paradigm cell filling problem has now been explored in detail for verbs in Romance (Maiden & O'Neill, 2010), French (Bonami & Beniamine, 2016), European Portuguese (Beniamine et al., 2021), Murrinhpatha (Mansfield, 2015), Navaho (Beniamine et al., 2017), Chatino (Beniamine, 2021), Latin (Bonami & Pellegrini, 2022), Pitjantjatjara (Wilmoth & Mansfield, 2021), Asama (Lévêque & Pellard, 2023). Hence, there may be more unity in how nominal and verbal knowledge develops from usage and is processed and represented than has hitherto been assumed. We will explore this hypothesis in more depth for the case of Slavic aspect which has been claimed to have grammaticalized to an unusual extent.

1.2 Slavic aspect: challenges of form and function

Aspectual marking is a widespread phenomenon. In a sample of 222 languages studied by Dahl and Velupillai (2013), almost half (101, to be precise) marked the aspectual imperfective/perfective distinction overtly. It is well known that, even among the many languages that do mark aspect, the Slavic languages stand out (Dahl, 1985, p. 69). Different from most languages, Slavic does not use consistent inflectional aspectual markers. Slavic aspectual morphology is highly irregular, and clear inflectional categories cannot easily be distinguished (cf., Grzegorczykowa et al., 1998 for Polish; Townsend, 1975 for Russian).

Imperfective and perfective variants relate to each other in one of three ways: through prefixation, infixation or stem changes and suppletion. Take the example of Polish, which captures the essence of Slavic aspectual morphology. Imperfective and perfective verb forms differ most often from each other in terms of prefixation, e.g., *pisać_impf* versus *napisać_pf* (write). Since Structuralism (but note before, see Grech, 1827), these two variants have been thought to create pairs (although some have argued for larger clusters, see Janda, 2007a; Maslov, 1948). In fact, it is possible for numerous different perfective forms to relate to the same imperfective verb, with prefixes generally adding or changing some shade of meaning; compare here *napisać_pf* (write) with *podpisać_pf* (sign) or *przepisać_pf* (advise, order). These prefixed verbs can, in turn, form their own, so-called secondary imperfective *podpisywać_impf* (sign). Prefixation does not affect the choice of verbal conjugation class, and the same conjugation classes are found across imperfective and perfective. A second way for imperfective and perfective verbs to relate to each other is by means

of stem changes as in *wracać_impf* versus *wrócić_pf* (return), infixation like *spotykać_impf versus spotkać_pf* (meet) or both as in *zaczynać_impf* versus *zacząć_pf* (begin). Suppletion plays only a minor role in the Slavic aspectual marking system, as in *brać_impf* versus *wziąć_pf* (take). Note that imperfective and perfective verbs of the latter two types may belong to different conjugation classes. Finally, there is a very small group of verbs, called "biaspectual', which is virtually exclusively made up of (recent) loanwords (Janda, 2007b and references therein) and for which one form functions as both imperfective and perfective. While the morphological relationship between imperfective and perfective has been found not to affect aspectual behaviour (Janda & Lyashevskaya, 2011), formal variability has contributed to the opinion that Slavic aspectual categories are not proper inflectional categories; instead, it is felt that they have something distinctly derivational (Dahl, 1985, p. 85, 89), a discussion which remains unsettled to this day (Gorbova, 2017; Percov, 1998).

The Slavic aspectual system has grammaticalised to an unusual extent and any verbal form bears aspectual marking; this is reminiscent of the nominal system, where each form bears case marking. Due to the extent of the grammaticalization of aspect, aspect has become the cornerstone of the verbal paradigm (see Table 1). If the prototypical Slavic nominal paradigm distinguishes six or seven cases (nominative, accusative, genitive, dative, instrumental, locative/prepositional, and possibly vocative) across two grammatical numbers (singular and plural), then the prototypical Slavic verbal paradigm distinguishes, across two aspects, at least three tenses (past, present, future), three moods (indicative, imperative, conditional), two voices (active/passive) and two numbers (singular and plural) with each three persons (first, second and third); aspect is additionally marked on the non-finite forms such as the infinitive, participle and gerund. Inflectionally, there are notable aspectual gaps that apply without exception: perfective verbs lack a present tense and have a simple future tense while imperfective verbs have a present tense but require an auxiliary to express the future. Similarly, only imperfectives have active present participles and present gerunds while passive past participles and past gerunds are exclusive to perfectives.[1]

The function of Slavic aspect appears equally complex. Although there is general agreement that aspect "has something to do with time" (Dahl, 1985, p. 25), the exact nature of this relation has remained elusive. Aspect appears as a non-deictic category that resists being pinned down with respect to the exact dimension(s) of experience it encodes. Two rather different approaches have dominated research on Slavic aspect (we refer to Binnick, 1991 and Sasse, 2002 for comprehensive overviews). One approach highlights the "different ways of viewing [. . .] a situation" and is known as grammatical or viewpoint aspect, which has dominated work on Slavic aspect. The other approach focuses on capturing the "internal temporal constituency of a situation" and leads to aspect being analysed as lexical or situational aspect.[2]

Grammatical or viewpoint approaches assume that aspect encodes, in the meaning of imperfective and perfective, a basic opposition that is primary and governs its

---

[1]There are further restrictions on the co-occurrence of certain moods, tenses and persons that are unrelated to aspect and will not be discussed here.

[2]Discourse theories of aspect remain out of the scope of this study; for overviews, see Hopper (1982).

Table 1 The Polish verbal paradigm (cells marked * require gender; those marked ** additionally require case; passives impose restrictions on person)

Mood	Voice	Tense	Aspect	
			Imperfective	Perfective
Finite forms				
Indicative	Active/Passive	Past*	✓	✓
		Present	✓	–
		Future	✓ (complex)*	✓ (simplex)
Imperative	–	–	✓	✓
Conditional	–	–	✓	✓
Non-finite forms				
Infinitive	–	–	✓	✓
Gerund	–	Past	–	✓
	–	Present	✓	–
Participle*/**	Active	Past	–	✓
		Present	✓	–
	Passive	Past	–	✓
		Present	✓	–

usage: the grammatical distinction is fundamental and leaves a recognizable mark on the lexical items to which it is applied. The overarching concept of boundedness, the temporal demarcation of the situation expressed by the verb,[3] has often been named as the concept to motivate the perfective and govern all its uses, leaving the unbounded imperfective to absorb any conflicting meanings (Sasse, 2002). Of course, like many grammatical categories, aspect is highly polysemous, making it quite challenging and rather reductionist to explain every instance of usage across the many languages that mark aspect formally in terms of whether the situation is presented as including at least one of its temporal boundaries, or not, and a range of concepts has been suggested to supplement the notion of boundedness.

Lexical approaches to aspect assume that aspectual usage is governed largely by lexical factors. Approaching aspect as a lexical category has its roots in the work by Maslov (1948) and later Vendler (1957) who noted that the meaning of a verb implicitly limits its usage; it remains an attractive proposal (see Croft, 2012). On a lexical approach to aspect, perfective and imperfective do not possess an invariant meaning that is primordial. Instead, the type of action expressed by the verb determines the meaning of the aspectual opposition and explains and predicts aspectual usage. Crucially, on a lexical account, aspect is considered to be regulated by categories of lexical items *and* their complements, which explains the variation in aspectual behaviour that individual lexical items may display (for a first solid elaboration, see Dowty, 1979).

---

[3]The concept of boundedness has been interpreted in several ways. For example, it has been described as involving (1) the highlighting of a boundary or the lack thereof, (2) a focus on the complexity of the situation as opposed to a summary view of its totality or (3) a view on the situation from the outside or from within.

Table 2 Some aspectual pairs and triplets from http://www.diaspol.uw.edu.pl/baza/

Imperfective [infinitive]	Perfective [infinitive]	Secondary imperfective [infinitive]	Translation
adaptować	**za**adaptować	**za**adapto**wy**wać	adapt
akceptować	**za**akceptować	**za**akcepto**wy**wać	accept
aklimatyzować	**za**aklimatyzować	**za**aklimatyzo**wy**wać	acclimatize
alienować	**wy**alienować	**wy**alieno**wy**wać	alienate
angażować	**za**angażować	**za**angażo**wy**wać	employ
artykułować	**wy**artykułować	**wy**artykuło**wy**wać	articulate
asfaltować	**wy**asfaltować	**wy**asfalto**wy**wać	asphalt

The fact that young children across languages as varied as Chinese, English, French, Greek, Italian, Japanese, Polish and Turkish use past or perfective forms first and mainly with telic predicates and present or imperfective forms with atelic predicates has been taken to show that young L1 learners[4] are influenced by lexical aspectual classes of verbs (and predicates) when they use tense-aspect markings in their target language; this is known as the Aspect First Hypothesis (Andersen & Shirai, 1994). This observation has given rise to claims about the primacy, if not innateness (Bickerton, 1981), of certain linguistic distinctions, notably aspect. A more straightforward interpretation of the data is, however, that learners accurately replicate the distributions they hear in the (maternal) input (Shirai & Andersen, 1995): adult native speakers' use of tense/aspect morphology is congruent with the Aspect First Hypothesis, and this raises the impression that children are aware of and influenced by lexical aspectual classes of verbs in their use of tense/aspect morphology, while in reality they are merely mirroring the input. We will return to the idea of prototypical exemplars as anchors for category meaning in the discussion of the findings of our computational learning model.

1.3 Mental correlates of aspectual paradigms

The dominant view on aspect as a grammatical phenomenon, illustrated in Table 1, has led to aspectual forms typically being presented in dictionaries and grammars as forming pairs, with one infinitive for each aspect, as shown in Table 2. In linguistic research, these pairs are often supplemented with a third form, the so-called secondary imperfective, which uses an infix to form an imperfective from a typically prefixed, perfective form, as in column three in Table 1b.

This traditional presentation assumes that language users know which forms are imperfective and which are perfective, and begs the question of whether the overarching aspectual categories of imperfective and perfective itself can be learned from exposure to usage. For long, little was known about the distribution of tense/aspect marking in larger datasets. Recently, studies have charted the aspectual behaviour of verbs across millions of sentences to address longstanding issues in aspectology.

---

[4]The same pattern has been shown to hold for L2 learners (see Bardovi-Harlig, 2000 for a review) including heritage learners (see Polinsky, 2006 for heritage Russian).

Eckhoff et al. (2017) use data from the small but manually disambiguated morphological standard of the Russian National Corpus on about 200 verbs (each with a lower frequency threshold of 50) across 3 genres. Information contained in the frequency profiles for each verb, i.e., how often each verb occurs in one of nine TAM classes, is across the genres on average 92.7% accurate in sorting verbs according to aspect. This result is statistically indistinguishable from the accuracy of the prediction of aspect based on morphological marking (Eckhoff et al., 2017: 870), i.e., considering *simplex verbs* (verbs lacking overt aspectual markers) as imperfective (which is wrong for perfective simplex verbs), prefixed verbs as perfective (which is wrong for prefixed indeterminate verbs) and verbs suffixed in -*y/iva, -va*, and −(*j*)*a* as imperfective (which is always correct). In other words, there are several distributional paths that lead to the detection of imperfective and perfective as grammatical categories, and these categories could thus be considered learnable from input.

However, the abstract grammatical categories these patterns are supposed to mark may not be obvious to linguistically untrained language users: the fact that patterns are detectable in input does not mean they are detected and exploited by language users (Divjak, 2015). Indeed, there is little, if any, support for the cognitive reality of the horizontal type of organisation, that is, of verbs into pairs or triplets. Despite the vast amount of descriptive literature on Slavic aspect, the study of how the highly grammaticalized version of aspect that typifies the Slavic languages is processed and represented mentally is in its infancy (Feldman et al., 2002; Roussakova et al., 2002; Anstatt & Clasmeier, 2012; Gattnar, 2013; Clasmeier, 2015) and inconsistency of findings abounds: while some insights can be gleaned, it remains unclear whether verbs are stored as pairs and whether the posited aspectual meaning differences are detected and applied.

Roussakova et al. (2002), for example, dealt directly with the question of whether the members of an aspectual pair are stored and processed as separate lexemes or as forms of the same lexeme in the mental lexicon of native Russian speakers. They conducted a speech production experiment and asked their subjects (children of various ages and adults) to produce the past tense form of a verb given in the present tense (e.g., stimulus *rešaet* 'he / she / it decides' > expected reaction *rešal* 'decided'). In this way, the authors tested 10 different aspectual pairs and then analysed 'mistakes', i.e., when subjects used the opposite aspectual form of the stimulus verb in their answers. As the probability of mistakes differed heavily between the tested items, the researchers concluded that the "degree of intimacy in the mental representations of the members of aspectual pairs differs considerably" (Roussakova et al., 2002, p. 308); thus, some pairs may be stored as one lexeme, but certainly not all of them. Interestingly, the smallest children showed the highest probability of correct answers, which implied that the mental representations of the members of an aspectual pair are acquired separately, only to then be combined during the development of language competence.

Feldman et al. (2002) showed for Serbian that there are differences in processing latencies for target word verbs depending on whether they are primed by inflectionally vs. derivationally related forms; e.g., PIŠU (imperfective 3 plural / [they] *write*) – PIŠEM (imperfective 1 singular / [I] *write*) versus ZAPIŠEM (perfective 1 singular / [I] *noted down*) – PIŠEM (imperfective 1 singular / [I] *write*). The facilitation was

stronger with inflectionally related primes than with derivationally related ones, and the latter types of primes show further differentiation depending on their semantic transparency (more priming if transparent vs. less priming if opaque). Indirectly, the study confirms the cognitive plausibility of the assumption that inflected verb forms cherish paradigmatic relationships. Furthermore, the semantic relationship plays an important role and appears to be graded, which also means lexeme specific.

Recently more sensitive techniques were introduced, including self-paced reading on Polish data (Klimek-Jankowska et al., 2018), eye-tracking on Russian (Bott & Gattnar, 2015; Klimek-Jankowska et al., 2018) and EEG on Polish (Błaszczak et al., 2014; Błaszczak & Klimek-Jankowska, 2016; Klimek-Jankowska & Błaszczak, 2020). Yet the research foci of these studies differ from the traditional theoretical discussions and, here too, findings have remained equivocal. Klimek-Jankowska and Błaszczak (2020)'s findings support the view that the domain of aspectual interpretation in Polish is the verb phrase, not just the verb; that the verb phrase would play a role is expected on a lexical aspectual approach. Earlier, Bott and Gattnar (2015), using eye-tracking, had reached the opposite conclusion for Russian where the domain of aspectual interpretation appears to be just the verb. We will return to these contradictory findings in the discussion of the findings of our computational learning model.

1.4 This study: learning to use aspect in context

While behavioural experiments provide support for the existence of certain linguistic structures (Rissman & Majid, 2019; Rissman & Lupyan, 2022), no formal procedure has been identified or proposed that would take us from exposure to usage through to full-fledged abstract category and that could be used to test emergence directly on actual data. As such, the idea of emergence which is core to usage-based approaches remains open to criticism. In this paper, we begin to address this challenge, and focus on two overarching questions of linguistic knowledge, illustrated on the basis of the category of aspect: 1) what might linguistic categories be made of if they emerge from usage, and 2) what information must linguistic categories minimally contain for them to be useful in guiding usage.

To answer these questions in an empirically refutable way we follow the proposal of Divjak et al. (2021) who operationalise emergence through learnability. In brief, only information or knowledge that can be learned from exposure to usage can be said to have emerged from usage. This operationalisation of emergence is weak rather than strong: we aim to determine whether the basic building blocks of abstractions that are theoretically motivated are learnable from information available to naïve language users (for more discussion see Divjak & Milin, 2022). In other words, our approach is constrained top-down by theory, and bottom-up by data. Strong emergence would be constrained bottom-up only, and exclusively driven by the data that might give rise to "natural" clusters or abstractions, regardless of whether they exist in the theory. In more technical terms, weak versus strong emergence roughly equates to the difference between supervised and unsupervised machine learning – i.e., the difference between classification into existing groups versus group discovery (cf., Love, 2002; Jebara, 2012).

We report on two studies, 1) an exploratory, large-scale corpus-based study in which we map out the natural distribution of aspectual forms as encountered in a large dataset, and 2) a computational learning simulation in which we model how aspect could be learned from exposure to different dimensions of usage. The present study does not aim to determine whether aspectual categories are explicitly learned, but whether their determining features can be learned by a simple algorithm using raw data (cf., Milin et al., 2016, 2023). It enables us to evaluate the types of cues and their informativity in learning to use the preferred aspectual form and to establish whether aspectual *usage* can be learned from exposure to the ambient language only, without the need for the abstract semantic labels.

2 Study 1 - Mapping out usage

2.1 Methods

To map out usage, we used the Araneum corpus (Benko, 2014), which comprises 841,894,001 word tokens scraped from the Internet; the corpus is formatted such that each line contains a word, a punctuation mark or an end of sentence/paragraph mark. It is automatically lemmatised and annotated for parts of speech following the National Corpus of Polish (NKJP) labelling conventions [https://www.sketchengine.eu/polish-nkjp-part-of-speech-tagset/].

From the corpus we extracted sentences by considering '.', ';', '?' and '!' as end of sentence markers, and any other punctuation as clauses markers. For each sentence, we noted their verbs, infinitive form, and part of speech tag as provided in the corpus file, and we also stored the position of the verb within the sentence. Since sentences may contain multiple verbs, and we consider each sentence+verb a data point, some sentences are repeated in our dataset, but each of these repetitions contains information about a different verb.

It is also important to note that, since our anlaysis focus on Tense and Aspect (TA) combinations in the indicative mood, we removed data points that did not contain a verb in the indicative mood. Where a verb's NKJP tag was labelled as 'bedzie', indicating a form of the verb 'być' ('be') we relied on a set of heuristics <https://github.com/ooominds/Polish-Aspect> to store it as a standalone verb or as part of another verb if it was an instance of 'być' ('be') used as auxiliary, e.g., as part of a periphrastic imperfective future. We excluded verbs tagged as 'impt' (imperative), 'imps' (impersonal), 'pcon' (contemporary adverbial participle), 'pant' (anterior adverbial participle), 'pact' (active adjectival participle), 'ppas' (passive adjectival participle), 'ger' (gerund), 'winien' (winien-like verbs that inflect in an atypical way, lacking some verbal flexemes).

We also handcrafted a set of rules <https://github.com/ooominds/Polish-Aspect> to annotate sentences for TA Tags; recall here that the perfective is incompatible with the present tense, and only occurs in future and past while the imperfective aspect is used in all three tenses. These rules exploited the tags extracted from Araneum, as well as lexical cues in the sentence, such as the presence of 'być' ('be') or whether the verb had certain morphological markers. Instances of sentences whose TA Tags were

either 'infinitive', 'aglt' (agglutinate być) or 'other' were also discarded at this stage. Therefore our dataset consists of 55,912,059 data points, of which 33,937,962 unique sentences. For the purpose of both studies reported below, simulations, distributions and frequencies are calculated on the corpus with repetitions, i.e., every verb in every sentence is considered.

Because the infinitive of Polish verbs carries aspectual marking, it cannot be used as a cue for learning. To overcome this difficulty, we use what we will refer to as "superlemmas", i.e., aspectual pairs consisting of both the perfective infinitive and the imperfective infinitive (ignoring the very rare secondary imperfectives), to serve as a placeholder for the meaning of the verb, while allowing us to not reveal the aspectual marking of the target verb to the algorithm. Note that this decision does not influence the results: we could have replaced all verbs with numeric placeholders instead, but this would have made it harder for human analysts to inspect the model output. For the current purposes we relied on a list of aspectual pairs, compiled from two sources: 1809 pairs culled from Project Diaspol (http://www.diaspol.uw.edu.pl/baza/) and 822 natural (in the sense of Janda, 2007a), non-ambiguous pairs from Princeton WordNet 3.1 for Polish (Dziób et al., 2017; for details see Borowski, 2022 Borowski, 2022). From these combined lists, which contained only 94 overlapping verbs, a total of 1,765 unique pairs of aspectual counterparts occurs in our corpus. Our focus on paired verbs also removed singletons such as 'być' ('be') and 'mieć' ('have'), as they do not have an aspectual counterpart; note that this reduction did not affect the original distribution of TA tags (Fig. 2). The lists also included reflexive verbs, but Araneum's lemmas are not marked for reflexive forms; therefore, we looked for sentence markers of reflexive forms such as *się* two words before or after the verb of interest and corrected the infinitives to account for reflexives. Overall, these operations reduced the size of our corpus to 18,788,976 sentences.

2.2 Results and Discussion

Focusing on aspectual usage in the indicative mood, we charted usage of 1,765 pairs across 18,788,976 sentences from the Araneum Polonicum corpus, excluding infinitives, participles, gerunds and agglutinatives. Even though it is generally thought that "[a]spect is one of the most pervasive and characteristic grammatical categories in Russian, requiring its speakers constantly to choose between Perfective and Imperfective" (Janda & Reynolds, 2019: 491), large-scale tracking of aspectual usage shows that there is, in fact, rather little choice. Figure 1a plots types, with the values on the X-axis representing bias bins that capture, in 10% increments, how often a particular aspectual pair occurred in the imperfective or the perfective. Pairs that fall in the 0 – 0.1 bin are pairs where the imperfective counterpart is use in at least 90% of cases. The values on the Y-axis express the number of pairs attested in each bin. Figure 1 reveals that the majority of our 1,765 pairs are biased towards occurring in one aspect only: 34% appears virtually exclusively in the imperfective and 23% in the perfective. More specifically, 71% of all imperfective-biased verbs occur in the imperfective in more than 90% of all cases. Only 194 unique pairs have a bias in the range of 0.4 – 0.6, i.e., in these pairs, the proportion of one of the aspects in a sum of frequencies ranges from 40% to 60%. In other words, only 11% of pairs can be described as being
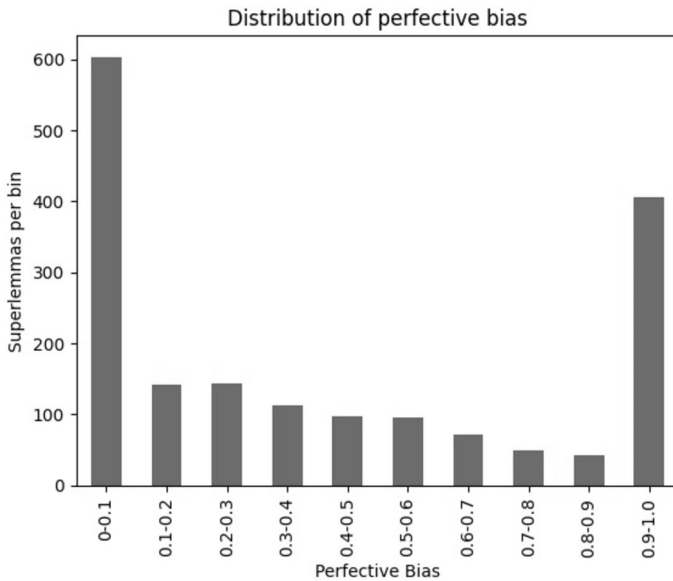
Fig. 1 Distribution of perfective bias in aspectual pairs attested in the Araneum corpus – type frequency or number of unique pairs

distributed evenly between the two aspects, even though that is the default position assumed by the dominant view on aspect. We note here that 41 superlemmas occur only once, hence, strictly speaking their preference for one or the other aspect cannot be determined. Of these, 28 occurred only once and in the imperfective, meaning they are included in the 0 – 0.1 bin; 12 occurred only once and in the perfective, meaning they are included in the 0.9 – 1 bin.

Figure 2 plots tokens per bias bin, with the values on the Y-axis expressed in millions of sentences. This reveals that the frequency of the pairs in the extreme bins in uneven, with those pairs preferring the perfective occurring much less frequently than those pairs preferring the imperfective. Furthermore, pairs that actually consists of two verbs that are used comparably frequently in their respective aspects are not only relatively limited in number, they are also not of high frequency. The equiprobably pairs in the 40%-60% range account for 1,657,421 out of 10 million sentences only, while overall, 3,848,416 sentences alone are construed with verbs that virtually exclusively occur in either imperfective (3,226,803) or perfective (621,613). Given the lower frequency of the perfective overall, the frequency of pairs that prefer to be used in the perfective in 60%-90% of all cases remains low, totalling 1,709,697. However, the higher frequency of the imperfective sees the total of sentences reach 3,406,079 in which those verbs are used that prefer the imperfective in 60%-90% of all cases.

Because the perfective only occurs in past and future tense, it may be worth having a closer look at the distribution of aspect over the tenses. Figure 3 visualizes the distribution of tense/aspect tags in the full Araneum of 55,912,059 sentences (left panel) and in the subsample of 18,788,976 sentences, obtained after filtering by our list of
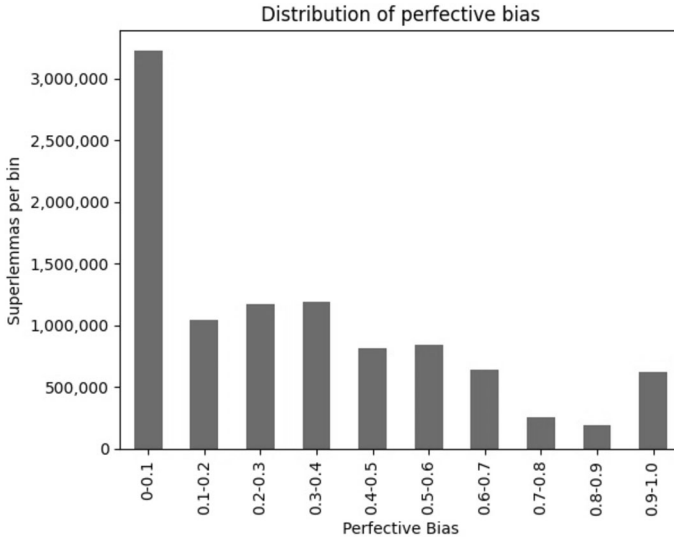
Fig. 2 Distribution of perfective bias in aspectual pairs attested in the Araneum corpus – token frequency.
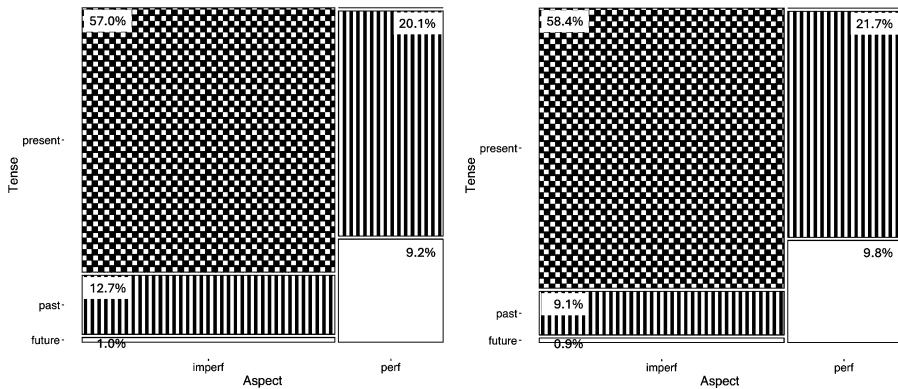
Fig. 3 (Left) Distribution of TA tags in Araneum; (Right) Distribution of TA tags after superlemma filtering. White = future, lined = past, checkered = present

superlemmas (right panel). Usage is clearly biased towards the imperfective (70.7%), and the bulk of these cases is consumed by the present tense. In the future, the perfective accounts for the lion share of usage (9.2% versus 1.0% for the imperfective). Only in the past do we see any sign of a potential choice between imperfective and perfective, yet with perfective (20.1%) still occurring nearly twice as often as imperfective (12.7%) in our data.

What causes aspectual bias? Is it something in the structure of the event that the verb expresses, or is it a by-product of reporting preferences, that is, might a verb's attraction to the imperfective be explained by the fact that the event it expresses is typically commented on while the event is taking place? In Fig. 4, the leftmost bar

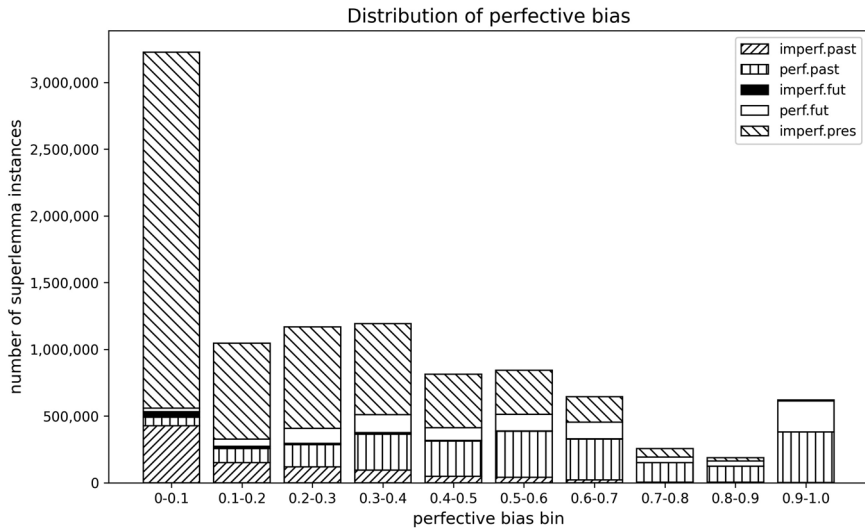**Fig. 4** Token frequencies for all tense and aspect combinations, per perfective bias bin.

represents all the tokens of those verbs that occur predominantly in the imperfective. It is clear that the imperfective accounts for a substantial proportion of tokens across tenses, yet the number of imperfective past tokens is also large. The opposite holds true of verbs in the bin at the other extreme, which contains verbs that strongly prefer the perfective, regardless of tense. From left to right, we see the proportion of imperfective occurrences diminish, not only in the present but also in past and future, and the proportion of perfectives increase, in the past but also in the future.

Looking beyond token frequencies at how verb types are spread out over the five possible TA combinations, we see in Fig. 5 that very few types in the leftmost bin are attested in the perfective past or future; these are clearly events that have a conceptual affinity with the imperfective. As we move towards the right-hand side of the plot, we gradually see the proportions of types for each TA combination equalize, with the exception of the rightmost bin. Reading Fig. 5 in conjunction with Fig. 4 reveals an interesting relation between type and token frequency. In the future, there are more types that prefer imperfective, but overall there are more tokens of the types that prefer perfective. In the past, there are not only more types that prefer perfective, but overall there also are more tokens of the types that prefer perfective. A lower type but higher token frequency has been found to promote verb-specific construction learning (Goldberg et al., 2004), while few occurrences of many types would encourage generalization (Wonnacott et al., 2008: 188–189). The implications of the usage distribution for learnability (cf. Divjak, 2017) is a topic we will return to in the discussion.

When we zoom in on non-present tense, as is done in Fig. 6, we see a very similar picture across past and future, with the TA proportions on the left-hand side appearing as the inverse of those on the right-hand side. In between, there is a rather equal number of types occurring in each TA combination, with more types attested for the imperfective future on the left, which gradually loses its dominance as we move to
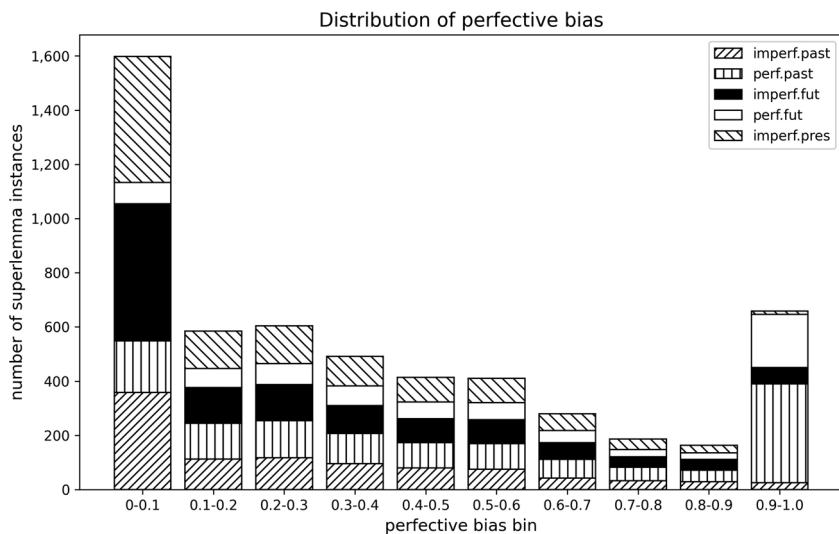
Fig. 5 Aspect/tense co-occurrence with types, per perfective bias bin.

the right. The perfective past is the TA combination for which most types are attested, starting from the second perfective bias bin on the left. Reducing our data set even further to only the past tense sees the same pattern repeating itself: while about a third of types that are strongly biased towards the imperfective are also attested in the perfective past (190 out of 603), of those types that are biased towards the perfective, only about 21% is attested in an imperfective past or future form (out of 406, 27 types are imperfective past and 60 are imperfective future). In all other bins slightly more types are attested with the perfective past than with the imperfective past; of the 615 unique types in bins ranging from 0.2 to 0.9, 596 appear in the perfective past (97%), while 473 appear in the imperfective past (77%).

In sum, charting aspectual usage makes aspect looks like a lexical phenomenon, rather than a grammatical one. For a phenomenon to be considered grammatical, one would have expected the bulk of the lemmata to show considerable usage of forms of either aspect. Yet, the distribution of lemmata over aspectual bias bins reveals a very strong lexical restriction on aspectual usage. This suggests that aspect is more lexical in nature than even a lexical approach to aspect (with its focus on semantic classes) would have assumed.

3 Study 2 - Simulating learning

If the hypothesis that aspect is lexical holds true, it should be possible to learn aspect from lexical information. We test this hypothesis by letting our computational algorithm learn from contextual data and select the aspectual paradigm that would harbour the verb form that would fit the context best. After all, if we assume that the goal of language use is successful communication, then language proficiency is the
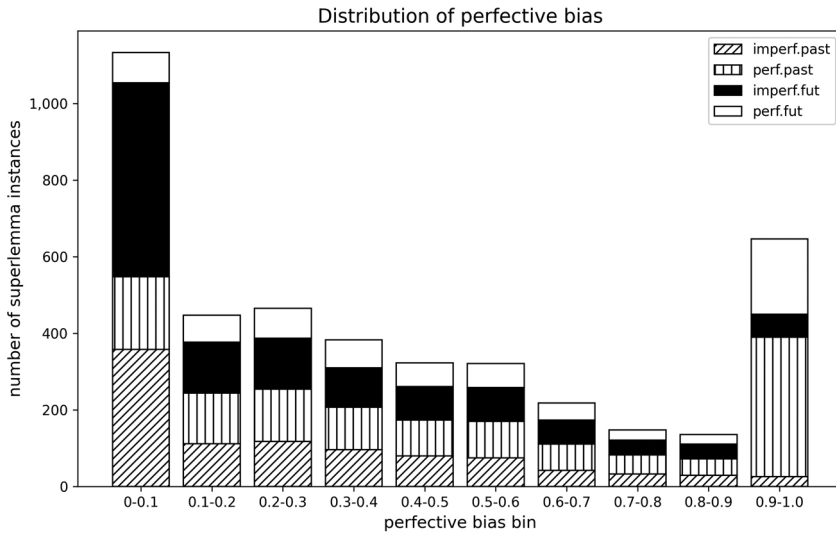
Fig. 6 Aspect/tense co-occurrence with types, per perfective bias bin, for past and future.

ability to select, from among a large(r) set, a form that would be acceptable given the context.

3.1 Methods

For the learning simulation, we employed the Naïve Discriminative Learning model (NDL; Baayen et al., 2011), which implements the Rescorla and Wagner (1972) rule of associative, error-correction learning computationally. This rule is also known as the Widrow and Hoff (1960) or Delta rule from work on Artificial Neural Networks (ANN) and Signal Processing. In simple terms, the rule governs the process of gradual, iterative learning about relationships between cues and outcomes in the environment (see Rescorla, 1988), in discrete time-steps. What this algorithm trades in terms of performance, at least when compared to recent large language models, it gains in cognitive plausibility of its architecture and hence in the resemblance of its performance to that of human participants (Romain et al., 2022). The original argument of Widrow himself is also relevant here: the "inability to realize all functions is in a sense a strength [...] because it [...] improves its ability to generalize" to new and unseen patterns (Widrow & Lehr, 1990, p. 1422; also see Milin et al., 2023).

Each step is represented as a learning event to which the rule is applied. A learning event consists of learning cues and outcomes. In NDL terminology, cues are the information we make available to the algorithm, while the outcomes are the information we want to learn about. Cues and outcomes can be present or absent: if a cue and an outcome are both present, then the rule requires strengthening of their association; however, if a cue is present but a given outcome is not, then their association must weaken. The strengthening or weakening is a small change to the previous state of the cue-outcome association. In other words, the process of learning builds "knowledge"

which changes dynamically as a learner receives new information (i.e., an event, with cues and outcomes).

Knowledge is distributed as a network of associations, expressed or quantified as association weights. At any moment, the state of the network which represents the knowledge acquired thus far can be probed by checking which outcome would be preferred, given a set of cues. When the current knowledge – i.e., the weights given the present cues and a given outcome, leads to the correct outcome, the learning rule makes the used weights "heavier" (strengthening), otherwise, error-correction must be applied to "lighten" the weights (weakening). Learning proceeds as new events take place and given informative cues and after enough time, which is the essence of experience, the error becomes the smallest possible (cf., Widrow & Hoff, 1960).

For modelling purposes, every learning event is made up of a single Polish sentence as retrieved from our corpus.[5] NDL requires a frame of cues and outcomes for learning. We employed 3 different families of cues for each learning event (sentence) namely superlemmas, n-grams and tense, each separately, in sets of two and all three together.

Recall that superlemmas merely represent the meaning of the verb whose aspect we were predicting. Typically, lemmata are used for this purpose, but since the infinitives of Polish verbs are also marked for aspect, we substituted them with so-called superlemmas, i.e., the aspectual pair the verb belongs to, to represent the meaning of the verb regardless of aspect.

Context has been hypothesized to play an important role in the choice of aspect (Janda & Reynolds, 2019) and the findings by Bermel and Kořánová (2008) and Cvrček and Fidler (2017) indicate that the context might contain a great deal of information about aspect, in the form of lexical items that co-occur with aspectually marked verbs.[6] To represent the context, we extracted n-grams, groups of 1 to 4 contiguous words in a sentence. Consider for example, the following sentence (1), where the target verb "otrzymały" ('received') is in the perfective:

(1)   Od samego początku w naszej gminie szkoły **otrzymały** swój budżet.
      From self start   in our district   schools received   their budget.

      From the very start, the schools in our district received their own budgets

The n-grams are all possible combinations of 1, 2, 3 and 4 consecutive words formed by all the words occurring in the sentence, excluding the target verb. Words within an n-gram are separated with hashtags and the n-grams themselves are separated from each other with underscores. For sentence (1), this results in the following:

1-grams: a single word forms a 1-gram; 1-grams are separated by underscores
       Od_samego_początku_w_naszej_gminie_szkoły_swój_budżet_

---

[5]Some parts of our corpus are not actually sentences, but titles or subtitles. Rather than removing such parts we adjoined them to the first following sentence. From a computational point of view this is, in fact, beneficial: firstly, utterances such as titles or subtitles are generally rare and, thus, they do not cause disturbance; secondly, these utterances add a fraction of noise that facilitates generalisation.

[6]Cvrček and Fidler (2017) tabulated lemmas in terms of how often they occur in close proximity to verbs, using a window of three words before and after. Bermel and Kořánová (2008) and Cvrček and Fidler (2017) investigate the relationship of each element to the target verb separately and independent of all other elements (and in the case of Cvrček and Fidler also independent of their position).

2-grams: two words together (joined by a #) form a 2-gram; 2-grams are separated by underscores Od#samego_samego#początku_początku#w_w#naszej_ naszej#gminie_gminie#szkoły_szkoły#swój_swój#budżet_

3-grams: three words together (joined by a #) form a 3-gram; 3-grams are separated by underscores Od#samego#początku_samego#początku#w_początku#w# naszej_w#naszej#gminie_naszej#gminie#szkoły_gminie#szkoły#swój_ szkoły#swój#budżet_

4-grams: four words together (joined by a #) form a 4-gram; 4-grams are separated by underscores Od#samego#początku#w_samego#początku#w#naszej_ początku#w#naszej#gminie_w#naszej#gminie#szkoły_naszej#gminie# szkoły#swój_gminie#szkoły#swój#budżet

We also included information about tense to capture an essential dimension of communication which will be clear to the interlocutors taking part in an authentic communicative situation, namely information regarding the relation between the time of the event and the time of speaking (before, during, after). Specifications of time are not necessary as long as there is common ground (term used here in the non-technical sense) between the interlocutors. We implemented tense as a 3-way system consisting of past, present and future, and as a 2-way system consisting of past and non-past. The 3-way system captures a naïve understanding of experiential time and its relation to the time of speaking (before, during, after), which is present in children before the age of five (Friedman, 1978). Recall also that, on a 3-way tense system, aspect and tense do not occur in all possible combinations: only the imperfective can be used to form a present tense.[7] The 2-way system represents the linguistic view on tense in Polish and other Slavic languages that do not formally distinguish between present and future by means of a different set of suffixes, yielding a contrast between what has happened and what has not yet happened at the time of speaking.

As outcome we included the aspectual labels "imperfective" or "perfective". Note that these are only shallow labels – placeholders, that do not contain or point to any further linguistic information. For example, in the superlemma OTRZYMAĆ# OTRZYMYWAĆ ('receive'), the lemma OTRZYMAĆ would carry the label perfective while the lemma OTRZYMYWAĆ would carry the label imperfective. Each lemma is associated with all the forms it takes in our corpus, for example, the imperfective lemma OTRZYMYWAĆ unites all imperfective forms that occur in our corpus. In other words, while we use the labels imperfective and perfective, they do not represent abstract categories and do not come with any linguistic baggage. Instead, they simply unite all the forms that are associated with one lemma (and occur in our corpus). Because of this, our simulations do not provide evidence for the existence of an abstract category of aspect in the (human or machine) learner, nor for the availability of the equivalent of the linguistic knowledge associated with the aspectual labels imperfective and perfective; they merely demonstrate the ability to select

---

[7]As a reviewer pointed out, an alternative 3-way tense distinction is possible where morphology guides the categorisation: this would leave us with a past which is morphologically marked in the same way for imperfective and perfective forms, a non-past which is morphologically marked in the same way for imperfective and perfective forms but the future is linked with the imperfective only due to the requirement of the auxiliary *be*. This system has a less transparent form-function mapping and was therefore dispreferred. We do, however, report the modelling results for this solution.

Table 3 Overview of all cue combinations and samples used for training

CUES	Tense classification	Samples
ngrams		All and Equiprobable
superlemmas		All and Equiprobable
ngrams + lemmas		All and Equiprobable
tenses	3-way and 2-way	All and Equiprobable
ngrams + tenses	3-way and 2-way	All and Equiprobable
superlemmas + tenses	3-way and 2-way	All and Equiprobable
superlemmas + ngrams + tenses	3-way and 2-way	All and Equiprobable

a form associated with the appropriate lemma, given the specific context provided as input information (learning cues). For example, for our sentence (1) above, the available set of cues (i.e., the specific context) and the outcome would look as follows: Cues:

- Superlemma: OTRZYMAĆ#OTRZYMYWAĆ; OTRZYMAĆ (perfective) is the corresponding infinitive for the target verb (same aspect), but the model is not told this. It is provided only with a meaning placeholder consisting of the imperfective and perfective infinitive of the target verb, to represent the meaning "receive"
- Ngrams: the other words in the sentence, split up into n-grams of length one, two, three and four words (underscores separate ngrams, while hashtags separate parts of an n-gram) Od_samego_początku_w_naszej_gminie_szkoły_swój_budżet_Od# samego_samego#początku_początku#w_w#naszej_naszej#gminie_gminie# szkoły_szkoły#swój_swój#budżet_Od#samego#początku_samego#początku#w_ początku#w#naszej_w#naszej#gminie_naszej#   gminie#szkoły_gminie#szkoły# swój_szkoły#swój#budżet_Od#samego#początku#w_samego#początku#w# naszej_początku#w#naszej#gminie_w#naszej#gminie#szkoły_naszej#gminie# szkoły#swój_gminie#szkoły#swój#budżet
- Tense: PAST; this is the tense in which the target word was used in the training sentence. Recall that we used two different classifications of tense.

Outcome:

- perfective, the target lemma's aspectual tag

This process was repeated for all sentences in our training set, meaning that all sentences had their target verb removed and replaced by a superlemma and a marker for tense, while all other words were combined into n-grams of lengths one through four.

We created various lists of cues, combining the aforementioned list of ngrams, the superlemmas and the tenses and trained on all cue combinations listed in Table 3.

Recall that filtering the corpus based on verbs for which we had superlemmas had reduced it to 18,788,976 sentences. Balancing the conflicting requirements of considering the environment (simulations that take longer to run have larger carbon footprints) and ensuring reliability (larger datasets yield more reliable results), we sampled 10,000,000 sentences; the model was then trained on 8,000,000 sentences and tested on 2,000,000. To reduce size and memory overhead further, and to speed

up the training, we created a file containing a list of ngrams to use for learning. We excluded anything with a frequency lower than 10, and for every n ranging from 1 to 4, we kept the 10,000 most frequent ngrams, for a total of 40,000 unique cues. Data preparation, extraction of word ngrams and selection of 40,000 most frequent ones, took about 30 hours. For this task we relied on the University of Birmingham, UK High-Performance Computing infrastructure, using 4 parallel threads and 240 GB of memory.

We then trained a Naive Discriminative Learning Model, implemented in Python, using the pyndl library for python 3.8, to predict the two aspects, to enforce learning of our target outcomes – a form of a particular aspect (perfective or imperfective). As shown in Table 2, we ran separate models, some including all 1,765 pairs, others focusing on those 194 pairs that showed equiprobable aspectual behaviour, and yet others still looking at past tense only or using a balanced TA sample. We also ran separate models with 3-way tense cues and with 2-way tense cues. Performance was measured using accuracy and F1 score. We relied on the same High Performance Computing infrastructure, this time with 16 parallel threads and 32 GB of memory.

3.2 Results & Discussion

Table 4 displays F1 scores for each simulation we ran; for each type of simulation, the highest scoring model, in its simplest form, is boldfaced. The F1 score is a machine learning metric that can be used in classification models, in particular models run on imbalanced data. It is a combination of two simpler performance metrics, precision (how many of the positives are true positives, i.e., how many items that the model identifies as instances of X are indeed instances of X) and recall (how many true positives were identified, i.e., how many of the total number of instances of X were identified by the model as instances of X). The F1 score is the harmonic mean of the precision and recall, and is recommended for imbalanced datasets precisely because it considers both precision and recall (rather than recording how often the model is correct in predicting one class over the entire dataset, as Accuracy scores do). An F1 score reaches its best value at 1 and its worst score at 0. The relative contribution of precision and recall to the F1 score are equal. The formula for the F1 score is:

$$F1 = 2 \times \frac{(\text{precision} \times \text{recall})}{(\text{precision} + \text{recall})}$$

where precision is the measure of the correctly identified positive cases from all the predicted positive cases and recall is the measure of the correctly identified positive cases from all the actual positive cases. The F1 value is high only if both precision and recall are high. Medium F1 scores are obtained if one metric is high, and the other one low.

While for all 1,765 pairs taken together the results for imperfective return an F1 that is above 0.8 on any type of cue, there is a marked improvement for the perfective with ngrams alone giving an F1 score of 0.41, superlemmas alone resulting in 0.65 and 3-way tense information alone 0.86. This suggests that ngrams and superlemmas encode (slightly) different types of information, but neither of these cues alone does as well as tense on its own. Our algorithm achieves an F1 of 0.95 predicting

Table 4 Summary of results for all simulation runs across types of simulations (all verbs, equiprobable verbs and balanced TA representation)

Bins	CUES	F1 IMPERFECTIVE	F1 PERFECTIVE
All			
	ngrams	0.81	0.41
	superlemmas	0.86	0.65
	ngrams + tenses (2-way)	0.86	0.69
	tenses (2-way)	0.86	0.70
	ngrams + superlemmas	0.87	0.69
	superlemmas + tenses (2-way)	0.91	0.79
	superlemmas + ngrams + tenses (2-way)	0.91	0.79
	tenses (3-way)	0.92	0.86
	ngrams + tenses (3-way)	0.93	0.86
	superlemmas + tenses (3-way)	**0.95**	**0.9**
	superlemmas + ngrams + tenses (3-way)	0.95	0.9
Equi			
	superlemmas	0.55	0.56
	ngrams	0.7	0.63
	ngrams + superlemmas	0.7	0.64
	tenses (2-way)	0.83	0.80
	superlemmas + 2-way tenses	0.83	0.80
	ngrams + tenses (2-way)	0.83	0.82
	superlemmas + ngrams + tenses (2-way)	0.84	0.82
	tenses (3-way)	**0.94**	**0.94**
	superlemmas + 3-way tenses	0.94	0.94
	ngrams + tenses (3-way)	0.94	0.94
	superlemmas + ngrams + tenses (3-way)	0.94	0.94
Balanced			
	tenses 2-way	0	0.57
	superlemmas + tenses (2-way)	0	0.57
	ngrams + tenses (2-way)	0	0.57
	ngrams + superlemmas + tenses (2-way)	0	0.57
	ngrams	0.01	0.57
	ngrams + superlemmas	0.18	0.6
	superlemmas	0.48	0.65
	tenses (3-way)	**0.5**	**0.67**
	ngrams + tenses (3-way)	0.5	0.67
	superlemmas + tenses (3-way)	0.5	0.67
	ngrams + superlemmas + tenses (3-way)	0.5	0.67

imperfective and 0.90 predicting perfective with a combination of superlemmas and 3-way[8] tense information. The added complexity, brought in by combining two types of cues, is justified in terms of improvement of model fit.[9] However, combining all three types of cues does not improve prediction accuracy further. We therefore consider the model containing 3-way tense cues and superlemmas as our best model.

Things look rather different for the subset of 194 equiprobable pairs: as expected, superlemmas do not contribute much here and consequently ngrams might shoulder more of the learning weight. On their own, ngrams predict imperfective less well (0.7 versus 0.81) but perfective better (0.63 versus 0.41) than for all pairs. Yet, they are strongly outperformed by tense which obtains 0.94 for imperfective and for perfective when used as only cue. Adding superlemmas and ngrams into the training mix does not yield further gains.

We also ran simulations on a balanced subset of our corpus, which contains 460,320 instances, with 92,064 instances of each TA combination; this number was set by the total number of instances that was available fo the least frequent tense/aspect combination, i.e. the imperfective future. In a balanced sample, F1 and Accuracy scores are considerably lower, and the best simulation runs on 3-way tense cues; this yields an F1 of 0.50 for imperfective and of 0.67 for perfective. It appears that learning how to use aspect benefits significantly from the differences in frequency with which cues and outcomes are attested (see also Romain et al., 2022 for similar findings for the English tense/aspect system).

In addition to comparing the F1 scores for the different models, our simple 2-layer model allows us to inspect the cues that are most strongly associated with the imperfective and perfective on each of our models (Table 5). Focusing on the 20 top cues for imperfective and perfective on the simplest best-performing model for all verbs (superlemmas + 3-way tense) we see that the model has correctly idenitfied present tense as strongest cue for the imperfective and past and future tense as strongest cues for the perfective. These tense cues are followed by a list of verbs that either occur predominantly in the imperfective (left column) or perfective (right column); note that the superlemma is provided, which contains both imperfective and perfective partner, in random order.

For equiprobable verbs, the 3-way tense is the simplest and best-performing model. However, when we look at the tenses + ngrams model which performs equally

---

[8]If we use an alternative 3-way tense distinction where both past and non-past have imperfective and perfective forms but future only has imperfective forms, as suggested by one of the reviewers, we obtain a slightly worse F1 of 0.91 for imperfective and of 0.79 for perfective, which is identical to the results obtained for the same including a 2-way tense distinction (F1 of 0.91 for imperfective and of 0.79 for perfective).

[9]To estimate whether the added complexity is justified in terms of improvement in model fit, we estimated the Akaike Information Criterion. AIC estimation is not perfectly straightforward, as one can assume either a single free parameter of the Widrow-Hoff/Rescorla-Wagner learning rule (i.e., the learning rate) for all trained models, or a "regression-style" number of parameters, in which case each weight is a parameter in the model. The latter option is more conservative but leads to large differences between models with very few parameters (e.g., tense cues only) or with a large number of parameters (e.g., ngram and superlemma cues). Nevertheless, some comparisons are meaningful. In particular, the model with tense cues only vs. the model with tense and superlemmas are strictly nested which makes a meaningful comparison possible. The Evidence Ratio, based on the difference between respective AICs (Evidence Ratio $= exp\,(abs\,(AIC_1 - AIC_2)/2) = 3.42\mathrm{E} + 27$), is in favour of the more complex model.

Table 5 20 top cues for each aspect in the superlemmas and 3-way tenses learning simulation with all verbs

Cue - imperfective	Weight	Cue - perfective	Weight
PRESENT	0.924920997	FUTURE	0.852556106
SKONSULTOWAĆ#KON-SULTOWAĆ (consult)	0.608229853	PAST	0.665321874
GOŚCIĆ#UGOŚCIĆ (host)	0.582353941	MIEŚCIĆ#ZAMIEŚCIĆ (fit)	0.331207945
NOCOWAĆ#PRZENO-COWAĆ (spend the night)	0.555785909	PRZERODZIĆ#PRZER-ADZAĆ (transform)	0.330757439
TULIĆ#UTULIĆ (cuddle)	0.532796214	TOCZYĆ#STOCZYĆ (conduct, carry out)	0.329093725
BIELIĆ#WYBIELIĆ (bleach)	0.519474539	TWIERDZIĆ#STWIERDZIĆ (claim)	0.328690407
RECYTOWAĆ#WYRE-CYTOWAĆ (recite)	0.516613172	KRYĆ#UKRYĆ (hide)	0.328420349
SZKOLIĆ#PRZESZKOLIĆ (train)	0.464932258	CHYLIĆ#SIĘ#POCHYLIĆ#SIĘ (droop)	0.328157652
DZIELIĆ#SIĘ#ROZDZIELIĆ#SIĘ (share)	0.443790459	PRZECIWIĆ#SIĘ#SPRZE-CIWIĆ#SIĘ (object)	0.327512192
SZKOLIĆ#SIĘ#PRZESZKOLIĆ#SIĘ (train oneself)	0.435524959	JAWIĆ#SIĘ#OB-JAWIĆ#SIĘ (appear)	0.326800887
KREŚLIĆ#SKREŚLIĆ (delete)	0.426691027	BUDZIĆ#WZBUDZIĆ (awaken)	0.32666851
DZIELIĆ#ROZDZIELIĆ (divide)	0.425067827	GARNĄĆ#ZGARNĄĆ (gather)	0.32578302
WYĆ#ZAWYĆ (howl)	0.413533694	SYPAĆ#SIĘ#POSYPAĆ#SIĘ (crumble)	0.325568507
ZWALCZYĆ#ZWAL-CZAĆ (fight)	0.409253514	TULIĆ#PRZYTULIĆ (hug)	0.325520927
POZOWAĆ#UPOZOWAĆ (pose)	0.4051071	NAŁOŻYĆ#NAKŁADAĆ (apply)	0.32491584
ZAGŁOSOWAĆ#GŁOS-OWAĆ (vote)	0.403183768	ZATRACIĆ#ZATRACAĆ (lose)	0.324487054
WSŁUCHAĆ#WSŁUCHIWAĆ (listen)	0.400330496	PRZEDŁOŻYĆ#PRZEDKŁADAĆ (submit)	0.324378561
ROZESŁAĆ#ROZSYŁAĆ (distribute)	0.399352585	PRZYDARZYĆ#PRZY-DARZAĆ (happen)	0.32430827
PRZETWORZYĆ#PRZETWARZAĆ (process)	0.396162598	MRZEĆ#ZAMRZEĆ (die)	0.323711654
WALIĆ#SIĘ#ZWALIĆ#SIĘ (tumble)	0.39448843	CHOWAĆ#SIĘ#WYCHOWAĆ#SIĘ (raise)	0.322384791

well but is slightly more complex, we notice that expressions of time signalling duration or repetition figure frequently among the 100 top cues for the imperfective (the cues with weights > 0.1 are displayed in Table 6, and relevant phrases are highlighted). The situation for the perfective looks very different: focus seems to be on cues that express modality, such as necessity, obligation, possibility, or aim. There are a few expressions that signal temporal relations of singularity or finality (highlighted). Compared to the corresponding model for all verbs, more ngrams expressing temporal information occur within the top cues for equiprobable verbs; given that the training sample for all pairs was bigger, more cues ended up having stronger weights.

While it would be unnatural to assume that learners are ever exposed to a sample of aspectual usage in the past tense only, we did train our model on such a subsample, for linguistic purposes. After all, our usage charts (Figs. 1 and 2) showed that aspectual choice or competition is limited to the past tense, where the usage proportion of perfectives versus imperfective is 2 to 1. Training on this subsample would reveal what guides usage, if tense is eliminated. Table 7 displays F1 scores for 6 different simulations, using either ngrams only as cues, superlemmas only, or a combination of both, for all verbs as well as for the equiprobable subsample. While we obtain relatively good results for all verbs for the imperfective with ngrams only (0.81), ngrams are not good indicators of perfective use (0.45). The situation is reversed for superlemmas only, which offers better cues for perfective usage (0.90) than for imperfective usage (0.73). Combining both types of cues improves the results marginally for the imperfective (0.75), but performance remains below the level achieved by ngrams only. In other words, it appears that aspectual usage in the past is guided by lexical information for the perfective, but by contextual information for the imperfective. The differential dependence of lexical items on contextual support should be taken into account in future experimental work on aspect as it may well be the cause of the inconsistensies that have been documented as to whether the domain of aspectual interpretation is the verb or the verb phrase.

Looking at the past tense data only in more detail, we recalculated the perfective bias to consider instances that are equiprobable in the past, and we identified 109 pairs, or 6.2% of all pair types, that display this behavior. In these cases, where neither meaning nor tense mediates aspectual choice, we see that ngrams play an important role in guiding aspectual usage. However, with the highest F1 scores at around 0.65, F1 scores for these simulations are by far the poorest. This comes as no surprise as these verbs, in the past tense, truly represent the most equiprobable scenario: the verbs themselves occur equally frequently in each aspect and tense, which is otherwise a strong cue, is kept constant.

4 General discussion

In this study we set out to explore what information linguistic categories, in this case aspectual categories, would contain if they emerged from exposure to usage only, and what information they must contain to be able to guide usage. We took a data-driven approach to aspectual usage, analysing data from Polish. Modelling the data using a computational algorithm, founded on psychological principles of learning,

Table 6 Cues with weights > 0.1 in the tenses and ngrams learning simulation with equiprobable verbs

Cue - imperfective	Weight	Cue - perfective	Weight
PRES	0.985882182	FUT	0.949930303
PAST	0.215667306	PAST	0.784331757
kilkakrotnie (several times)	0.206594476	będą#mogli	0.185942098
kto#by	0.184874307	będziemy#musieli	0.180584636
wielokrotnie (repeatedly, frequently)	0.179937884	będzie#musiał	0.146832588
Pamiętam	0.162332478	będziemy#mogli	0.143844827
musieli	0.153663434	będę#miał	0.140831616
opowiadał	0.147308628	będą#musieli	0.140730475
będzie#o	0.146838345	będziesz#mógł	0.139516947
co#chwilę (every now and then)	0.142186359	będą#mogły	0.131500799
Robili	0.125868838	będziesz#miał	0.116555538
polegał#na#tym#że	0.125659142	na#zawsze (forever)	0.108559949
systematycznie (regularly)	0.123928538	staje#się#coraz (is becoming increasingly)	0.103041433
Będę	0.123620187	będzie#miał	0.099806443
w#cieniu	0.121721047	nie#jeszcze (not yet)	0.090196997
o#nic#nie	0.115814723	będą#mieli	0.08999105
długo#się (long)	0.112291478	spojrzał	0.087990363
Robił	0.10939012	będzie#w#stanie	0.08687219
pamiętam	0.108166661	żeby#było	0.081531924
przez#lata	0.106373197	na#mnie#duże#wrażenie	0.080806392
Stopniowo (gradually)	0.106159291	nie#będzie#w#stanie	0.080235187
będzie#za	0.105211185	nigdy#nie#będzie (never will be)	0.078998929
Konsekwentnie (consistently)	0.105177323	się#do#powstania	0.078785068
mogli	0.105159092	będzie#mógł	0.078172201
mecz#w	0.10463201	się#do#niewoli	0.078171742
Przez#cały	0.104326408	gola	0.078129907
nikt#nie#będzie	0.104296184	uchwałę	0.077994039
pod#okiem	0.103428439	od#czasu (since)	0.077590816
wiele#razy (many times)	0.10340625	będzie#mogła	0.077483926
do#niedawna	0.103226692	do#roku#życia	0.077441044
Razy (times)	0.102941229	choć#raz (at least once)	0.077301521
co#miesiąc (every month)	0.102084683	inicjatywę	0.07504836
który#będzie	0.100908009	Ostatecznie (finally)	0.07498379
próby	0.100335393	będą#miały	0.074289739

shed new light on what information language users need to extract from input to be able to use an aspectually appropriate form in context. The findings have implications for the way in which morphological categories and the paradigms they head up are conceived within approaches to linguistics that emphasize the cognitive reality of the categories they acknowledge.

Table 7 Results of the simulation runs for past tense data

Bins	CUES	F1 IMPERFECTIVE	F1 PERFECTIVE	Accuracy
All				
	ngrams	0.45	0.81	0.72
	superlemmas	0.73	**0.90**	0.85
	ngrams + superlemmas	**0.75**	**0.90**	0.86
Equi				
	superlemmas	0.55	0.52	0.53
	ngrams	**0.66**	0.64	0.65
	ngrams + superlemmas	**0.66**	**0.65**	0.65

4.1 Aspectual use is lexical, and mediated by tense

Mapping out usage for our set of 1,765 aspectual pairs across 18,788,976 sentences revealed that usage is biased towards the imperfective (70.7%) in general, the bulk being consumed by the present tense. Only in the past do we see any sign of a potential "choice" between imperfective and perfective, albeit with the perfective still occurring twice as often as the imperfective in our data. Furthermore, tracking the aspectual usage preference for each of our aspectual verb pairs reveals that more than half of unique pairs (1008 in total, or 967 if we exclude those that only occur once) are strongly biased towards occurring in one aspect only with at least 90% of attested usage in the preferred aspect. Usage is distributed more evenly between the two aspects for only 194 or 10.99% of unique pairs. In other words, aspect is, in essence, a radically lexical phenomenon: while it seems possible to create the imperfective or perfective counterpart of virtually any verb, and present them as aspectual pairs, this does not imply that both aspectual partners are used equally frequently. Instead, most verbal concepts have a very strong bias towards being reported on in one aspect or the other. This preference is mediated by tense, or the time at which an event is typically reported on, i.e., after, during or before taking place. It appears that the meaning of a verb consists of information about when to use a particular label, and this implies knowing what type of event it applies to, but also when that event is typically reported on referenced.

The idea of aspectual partners that has pervaded the literature on Slavic aspect since Jakobson (1932) and dominates the way in which aspect has traditionally been approached thus applies to a small subset of the verbal lexicon only. It is on this small subset that the grammatical view on aspect comes into its own: if aspect is a grammatical category and applies to every verb, resulting in an imperfectice and a perfective partner for every verbal concept, then the appropriate form is to be selected and this process is thought to be guided by the aspectual meaning that is being rendered. The fact that the discipline has focused precisely on this "alternating" subset seems to be a consequence of Structuralist theory and methodology which dominated the linguistic scene during the first half of the 20th century. Following Jakobson (1932), the minimal pair technique was applied to the study of aspect. Minimal pairs are words or sentences that differ only in one dimension, such as the imperfective (2) versus perfective (3) form of the same lexeme, while keeping the context constant.

(2) ***Wczoraj czytałem książkę.***
yesterday read; IMPF.1SG;PST;IPFV book
'I was reading a book yesterday.'

(3) ***Wczoraj przeczytałem książkę.***
yesterday read; PF.1SG;PST;IPFV book
'I read a book yesterday.'

This manipulation foregrounds the contribution aspectual morphology makes to meaning, while neutralizing the importance of the lexical verb and the context. Because of this, attention is drawn to that part of the verb inventory that allows both aspects equally. With large amounts of data at our fingertips, this dominant view is being challenged. It could, of course be argued that neither structuralists nor generativists make claims as to the statistical distribution of forms or the statistical dependencies between features. Yet, structuralist work did yield detailed clasifications of the contexts of usage for each aspect, often based on examples extracted from literary works (Maslov, 1948; Isačenko, 1960; Bondarko, 1971). However, their theory considered variation in usage between elements of a pair irrelevant. The same argument can be made for generative approaches: modern generative approaches rely on frequency information for parameter setting, and this information can only be obtained from usage. Yet, here too, generative theory declares the details of the usage dimension irrelevant, hence they are not considered further. Usage-based linguistics differs radically in this respect and puts frequency of usage, as proxy for experience, central on the agenda (but see Divjak, 2019 for a nuanced overview of the role fo frequency in theory-building).

The distribution of imperfective and perfective variants in actual usage points in the direction of a strong link between lexical item and aspectual variant. Rather than there being a basic opposition in the meaning of imperfective and perfective that pervades all its applications, regardless of lexical item, it is the lexical item that gives the aspectual opposition between imperfective and perfective its meaning. That is, the type of action expressed by the verb determines the meaning of the aspectual opposition and explains and predicts aspectual usage for the majority of verbs. However, our data points towards a lexical basis for aspect in a radical way. Rather than there being a limited number of semantic groups (Maslov, 1948; Vendler, 1957), our data suggest that aspectual preferences are stored lexeme by lexeme. In other words, aspect and lexical meaning are inseparable for the majority of verbs.

It could be argued that there are downsides to adopting a radically lexical account of aspect, in particular in terms of the challenges it poses for memory. Although storing information at the lexical level goes against the principle of economy of description, often coveted by linguists, it does not violate what we know about the storage capacity of the brain (see Herculano-Houzel, 2009 for a critical assessment of existing proposals). It has long been known that words vary in the frequency with which their various forms occur in actual usage (see Kaeding, 1898 for German; Thorndike & Lorge, 1944 for English; Steinfeldt, 1963 for Russian verbs) and that speakers of a language are aware of the frequency with which forms occur, even explicitly (for an overview, see Gernsbacher, 1984). On a usage-based account, this is expected as children learners learn words from exposure: the mappings between a form and the

experience it refers to are presented and learned through exposure to instances in a particular context of use. This means that the forms that are retained are not the base forms listed in a dictionary, but actual usage forms that contain the morphosyntactic marking required by the context. In a language like English that displays only very modest form variation, there may not be any difference between a word in isolation versus a word in context. However, using data from the Russian National Corpus, Lyashevskaya (2013) lists a large number of verb biases, paying special attention to verbs that are never used in the imperative.

It is well known that, in order to use a word correctly, a great deal of specific knowledge is required. In the verbal domain, differences in and knowledge of verb co-occurrence with types of subject or object, tense, aspect or mood have given rise to the Behavioral Profiling methodology (see Divjak, 2010 for an overview) by which usage is charted along a large number of dimensions to map out the characteristics of the individual senses of a polysemous word to pinpoint the differences between near-synonyms. Psycholinguistic research on morphologically richer languages has supplied us with ample evidence that L1 users (a) have knowledge of word forms and their relationships (see Milin et al., 2009 for evidence from Serbian), (b) store high frequency forms separately (Alegre & Gordon, 1999 put the boundary at 6 occurrences per million words for English, but this may be higher for morphologically rich languages such as Finnish, see Soveri et al., 2007), and (c) are able to determine whether near-synonyms are exchangeable in context based on usage information related to tense, aspect and mood preferences, among other things (Divjak, 2010).[10] The strong aspectual preferences of the individual lemmas are very apparent in the incomplete language acquisition observed in heritage speakers: if only one lemma is retained, it tends to be that member of the aspectual pair that denotes the more common, and hence more frequent, conceptualization of the event (Polinsky, 2008, p. 276).

The answer to our theoretical question thus seems to be that usage of inflected forms indeed introduces conflicting tendencies when it comes to organising the knowledge this yields at a more abstract level: there is a sense of completeness of what a given verbal paradigm should consist of, and a sense of partialness caused by the actual usage of all possible forms. The cell-filling problem (Ackerman et al., 2009), thus, becomes a feature rather than a problem – a language default. In fact, a lexical account of aspect seems to reflect a more general "organising" principle, which we can also observe in the lexically specific bias that nouns show for grammatical number (Baayen et al., 1997) or the lexically specific distribution of nominal cases (Milin et al., 2009). In that sense it might not be unreasonable to say that the available language input shows consistency, and as such makes structurally similar input available to a learning system (a human or a machine). This consistency could in principle even drive the "evolution of learning machines" (Poggio, 2012), something that Gallistel (2000) refers to as being "adaptively specialised", by challenging them to become gradually more attuned to relevant input.

---

[10]Intriguingly, due to the large number of different forms that each require a specific context, some word forms in morphologically rich languages may occur so rarely that they become hard to produce even for native speakers, although their formation can be considered as fully rule-governed. Lack of exposure to certain forms inhibits their production (Sims, 2015).

4.2 Usage of aspect can be learned by associating cues and outcomes in skewed input

To test the hypothesis that aspectual usage in Polish is lexical, we trained a computational learning algorithm to learn, from exposure to usage, to select the aspectual variant that fits the context. Taking all 1,765 verb pairs into consideration, our algorithm achieves an F1 of 0.95 predicting imperfective and 0.90 for perfective with a combination of superlemmas and tense information only. For the 194 equiprobable pairs, i.e., pairs of which both the imperfective and the perfective are used roughly equally often, the 3-way tense distinction offers the simplest and best-performing model, with an F1 of 0.93 for imperfective and of 0.94 for perfective. And for past utterances only, perfectives are best cued by superlemmas (F1 = 0.90), but imperfectives by contextual ngrams (F1 = 0.81). The present results concur with such an inference: verbs seem to show a "personal taste" (a bias) for occurring in one particular aspect. That taste seems to be determined by a combination of narrative time (tense) or the preferred time at which an event is reported on, and lexical meaning or the essence of the even that is being reported on.

While it may not be surprising that aspectual usage is learnable from exposure to words, in isolation or in chunks (after all, this is how children are thought to master the phenomenon, at least on a usage-based approach, see Stoll, 2001), it is impressive that a simple 2-layer computational network manages to do the same. The addition of tense information does improve model performance, yet this information arguably remains "naïve" and accessible to learners from a very young age: it merely encodes whether an event has already taken place, is currently taking place or still has to take place, compared to the now. The *emergence-as-learnability* of aspect is, thus, possible, constrained by the theory (of what ought to be learned) and the data (available to naïve language users). Given that the theory supervises the process by setting the learning targets, it is important to emphasise that our results show that the algorithm learned aspectual cues and not necessarily and explicitly aspectual categories as such. The fact that a complex morphological phenomenon can be captured by a "simple" associative process does not detract from the achievement of our model, and does not imply that the phenomenon itself does not have a cognitive dimension: instead, associative learning offers an explanatory mechanism for that complex cognitive phenomenon. It should not come as a surprise that simple principles can lie at the core of colossal complexities. The same applies to the principle of associative learning that has been argued to underly complex behaviour and higher cognition (Hanus, 2016; Mondragon, 2022). At the same time, the simple Widrow-Hoff or Delta element is the crucial building block of more complex and powerful learning systems, such as Connectionists Parallel-Distributed Processing models (Rumelhart et al., 1986) and various Deep Learning models (cf., Golovko, 2017; Irie et al., 2022). Hence, with internalised distributed representations (i.e., embeddings) and complex processing architectures one can expect further advancements and refinement of our understanding of language complexity and dynamics. Recent studies, for example, show that state-of-the-art learning systems can develop pseudo-localised representations of particular linguistic features (e.g., Varda & Marelli, 2023).

In our study, we have focused on the linguistic dimension and what it offers in terms of inputs to be associated with outputs, leaving lived experience out of the

equation. Yet we found that lexical and temporal dimensions alone are so rich in information that they achieve an F1 of 0.9 (perfective) and 0.95 (imperfecive) in using aspectual forms correctly. If users are to build up knowledge about language from exposure to usage, to attain the coveted state where they are able to "use words and grammatical forms because of what they mean", they must have formed associations between elements of the input, and the desired outputs.

Meaning stems from complex and dynamic interactions in and with the environment (similar ideas can be found in theoretical work as diverse as, e.g., mediation theory: Osgood, 1966, 1980; embodied cognition: Barsalou, 1999, Barsalou & Wiemer-Hastings, 2005). How we use words is the results of the synergy of many forces such as our own experiences of events, but also our observation of others experiencing these events; our experiences of how these actions are expressed verbally and in what kind of wider contexts they are talked about. The mechanism behind said forces might well be that of a simple association of cues and outcomes over many of our lived and observed experiences. The sheer number of such experiences, the number of potential cues in them and the rich dynamics of their competition, all of that together, is singularly complex, ever-changing and, hence, immensely adaptive.

4.3 Aspect as an abstract category is surplus to requirement

If aspectual use is learned from exposure to input, the distribution of which has very heavy tails, what does the morphological organisation of aspectual knowledge look like? Do language users (need to) distil a concept of aspect that has two possible values, imperfective and perfective?

Our findings are fully compliant with the general views of usage-based (cf., Bybee, 1988, 1995; Langacker, 1988, 2018) and word-and-paradigm approaches (cf., Matthews, 1965, 1991; Blevins, 2006, 2016). No general combinatorial rule(s) over static morphological units is needed to explain the attested patterns of usage. Instead, as Bybee (1995) points out, the dynamics of word usage reflects varying degrees of lexical strength, where those with greater strength retain independence and, in the context of the present study, can serve as anchors within their respective paradigms. From a word-and-paradigm point of view, this is why paradigms, however gappy, should be considered natural units of linguistic analysis (cf., Matthews, 1991; Blevins, 2016), with their unique set-theoretic (Stump & Finkel, 2013) or schematic (Booij, 2017) characteristics. The patchy or skewed usage of imperfective versus perfective aspectual forms calls into existence what we have called the emergence of contrasting forces: the potential or promise of paradigmatic completeness (where non-attested forms of a lexeme can be provided by analogy to attested forms of a lexeme) in spite of the defectivity or gappiness that characterizes usage (cf., Sims, 2015; Ackerman et al., 2009). These contrasting forces could provide guidance to answer the following pressing question: if aspectual use is learned together with the lexeme, (how) does any knowledge emerge that can be applied productively, to unseen items?

It is important to recall here that a lexical approach does not preclude language users from detecting commonalities among lexemes and usage-based theories make explicit allowances for abstract categories to exist alongside concrete, lexical entries (Langacker, 1987b). From our model, we can identify a verb or a set of verbs that

frequently occur(s) in the imperfective (present) or perfective (past) and that could function as prototype for the category. Chcieć ('want'), for example, is strongly attracted to the imperfective across all tenses, while robić ('do') is frequent in the imperfective present but gravitates towards the perfective for past and future. It is plausible to assume that the (virtually) exclusively imperfective or perfective verbs function as anchors that would provide language users with a wealth of information from which they could distil the function of imperfective and perfective. Shirai and Andersen (1995) proposed that prototypes could be used for anchoring. The exclusive co-occurrence of imperfective aspect and present tense would provide learners with a strong sense of what meaning, if any at all, the imperfective aspect conveys; note that a user's notion of "imperfective meaning" here might not be more than a vague sense of all the information that is typically accessible when an event or action is reported on that is taking place at the moment of speaking. According to Laskowski (1984), in Polish, the aspectual opposition can be clearly defined only for telic verbs. In other cases, even though pairs can be formed, what is expressed changes and becomes irregular, as it depends on the semantics of the verb. This assumption of vagueness of any assumed aspectual meaning at the category level is strengthened by the observation that contextual elements differ for verbs in the equiprobable zone: it does appear to be the case that, often, the meaning conveyed by the form on its own is insufficient and the intended interpretation needs to be amplified. This can be achieved by adding contextual elements that express ongoingness, for example. In other words, while abstractions are possible, they appear as surplus to requirement: they are neither necessary, nor sufficient.

Interestingly, our findings also resolve a long-standing question relating to the role that these contextual items play. Janda and Reynolds (2019) proposed that "anchoring in redundant contexts likely facilitates the independence of construal in contexts with less redundancy": users would learn the function of aspect by encountering aspect in contexts, enriched with pointers. For example, adverbs expressing repetition or duration tend to occur alongside imperfective verbs, and could be used as cues to the function of the imperfective. Unfortunately, "known cues for aspect are absent in the majority of such contexts", leaving "the question of what exactly makes contexts redundant" unanswered. This study shows that it is the lexemes themselves that make the context redundant; because aspectual preference is unique for each lexeme, cues for aspect can be absent in the majority of cases. We found that a model enriched with n-grams did not improve performance, but it allowed us to compare the types of contextual elements that are important for the two types of verbs (biased versus equiprobable): compared to the corresponding model for all verbs, more ngrams expressing temporal information occur within the top cues for the use of the imperfective aspect in case of equiprobable pairs. In other words, contextual markers may not occur often, but they do occur where they are needed, that is in cases where the lexeme itself is not already strongly biased towards one aspect or the other. We see this again when we zoom in on the past tense only, where contextual elements cue the imperfective. Recall that the imperfective is the less frequently used aspectual form in the past tense, with the perfective occurring twice as often. In this case, too, contextual elements come out strong: as was found for the English tense/aspect system (Romain et al., 2022), contextual elements are used with forms that are less usual

and therefore harder to process. They provide the contextualisation that is required to interpret the form correctly.

The skewed distribution we observed might convince even sceptics that, to use verbs correctly in the majority of cases, language users do not need more than lexical knowledge, paired with tense, i.e., knowledge of when events are typically reported on relative to time of happening. However, they would likely argue that language users need *some* understanding of the contrast to deal with the equiprobables. In usage-based circles, the choice of one aspectual form over another, where allowed, is considered as an instance of construal (see Croft, 2012 for the theoretical elaboration of aspectual construal and Janda & Reynolds, 2019 for an application). The concept of construal (Langacker, 1987a; Talmy, 1988) refers to the contribution made by language to the perception and conception of an event: the linguistic choices we make reflect or affect our conceptualization of the event we communicate. Recall that in a limited number of cases, i.e., 11% of pairs across all tenses (see Figs. 1a and 1b), both imperfective and perfective conceptualisation appear equally felicitous. Yet a closer look reveals that tense plays an important role here, and contributes hugely to selecting the preferred aspect. When we zoom in on past-tense data only, aspectual meaning, if any, comes into its own, but the number of contexts in which this is the case is rather limited: together they account for a mere 155,496 usage instances, or 0.02% of our 10 million sentence sample. Contra Janda and Reynolds (2019), in our data, construal appears as an add-on to the system, not as its core. It is available in a relatively small number of cases only and unless these instances are very salient, they would not provide a firm basis for learning the meaning contrast that has been said to govern aspectual usage, in particular because the context does not provide much additional support either, with contextual ngrams doing little to improve model performance. What are the implications of these findings for the cognitive reality of the concept of aspect or its subcategories of imperfective and perfective, and by extension for morphological categories and their paradigms in general? Categorisation is a rather active domain of research within cognitive psychology, which has brought forth several families of models. Strikingly, the most successful models of categorisation such as Nosofsky's exemplar model (1986) and the SUSTAIN model (Love et al., 2004) use continuous rather than discrete representations to produce discrete behavioural outcomes; that is, computational models that are not provided with labels are nevertheless able to categorise instances using discrete labels. This has been coined "categorisation without categories" (Ramscar & Port, 2015). Hence, the question regarding the cognitive reality of aspectual categories, as an example of morphological categories, could or should be reformulated to ask which continuous properties must be learned given that language users (need to) produce discrete language outcomes, in this case, a particular verb form in a particular context. The simple model presented here shows that iteratively learned, continuously developing association weights are capable of doing just that: the principles of associative learning offer a mechanism that enables language users to digest their experience and retain the information they need to use morphological forms appropriately. The paradigms that emerge from usage do not look like their idealised theoretical counterparts. They are, by their very nature, defective. Hence, concern should not so much be with how unattested cells are filled, but instead with how usage is mastered given

the gappy nature of the input and the kind of system this skewed distribution gives rise to.

Data Availability The data used in this study is freely available from http://ucts.uniba.sk/aranea_about/_ polonicum.html.

Declarations

Competing Interests The authors declare that they have no conflict of interest.

References

Ackerman, F., & Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language*, *89*, 429–464.

Ackerman, F., Blevins, J. P., & Malouf, R. (2009). Parts and wholes: Patterns of relatedness in complex morphological systems and why they matter. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar: Form and acquisition*, Oxford: Oxford University Press.

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, *40*, 41–61.

Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition*, *16*, 133–156.

Anstatt, T., & Clasmeier, C. (2012). Wie häufig ist poplakat'? Subjektive Frequenz und russischer Verbalaspekt. *Wiener Slawistischer Almanach*, *70*, 129–163.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*, 94–117.

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*, 438–481.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. *Grounding cognition: The role of perception and action in memory, language, and thought*, 129–163.

Beniamine, S. (2021). One lexeme, many classes: Inflection class systems as lattices. In M. Sailer & B. Crysmann (Eds.), *One-to-many relations in morphology, syntax and semantics*. Language Science Press.

Beniamine, S., Bonami, O., & McDonough, J. (2017). When segmentation helps: Implicative structure and morph boundaries in the Navajo verb. In *Proceedings of the first international symposium on morphology* (pp. 11–15).

Beniamine, S., Bonami, O., & Luís, A. R. (2021). The fine implicative structure of European Portuguese conjugation. *Isogloss. Open Journal of Romance Linguistics*, *7*, 1–35.

Benko, V. (2014). Yet another family of (comparable) web corpora. In *17th international conference text, speech and dialogue* (pp. 257–264).

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*, 150–177.

Bermel, N., & Kořánová, I. (2008). From adverb to verb: Aspectual choice in the teaching of Czech as a foreign language. *Between Texts, Languages, and Cultures*, 53–70.

Bertram, R., Baayen, R. H., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language*, *42*, 390–405.

Binnick, R. (1991). *Time and the verb: A guide to tense and aspect*. Oxford: Oxford University Press.

Błaszczak, J., & Klimek-Jankowska, D. (2016). Aspectual coercion versus blocking: Experimental evidence from an ERP study on Polish converbs. In J. Błaszczak, A. Giannakidou, D. Klimek-Jankowska, & K. Migdalski (Eds.), *Mood, aspect, modality revisited. New answers to old questions*. Chicago.

Błaszczak, J., Jabłońska, P., & Klimek-Jankowska, D. (2014). An ERP study on aspectual mismatches in con- verbial contexts in Polish. In V. Torrens & L. Escobar (Eds.), *The processing of lexicon and morphosyntax*. Newcastle upon Tyne.

Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics*, *42*, 531–573.

Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford: Oxford University Press.

Blevins, J. P., Milin, P., & Ramscar, M. (2017). The Zipfian paradigm cell filling problem. *Perspectives on morphological organization*. Brill.

Bonami, O., & Beniamine, S. (2016). Joint predictiveness in inflectional paradigms. *Word Structure*, *9*, 156–182.

Bonami, O., & Pellegrini, M. (2022). Derivation predicting inflection: A quantitative study of the relation between derivational history and inflectional behavior in Latin. *Studies in Languages*, *46*, 753–792.

Bondarko, A. V. (1971). *Vid I vremja russkogo glagola (The aspect and tense of the Russian verb)*. Moscow: Prosveščenie.

Booij, G. (2010). Construction morphology. *Language and Linguistics Compass*, *4*, 543–555.

Booij, G. (2017). The construction of words. In B. Dancygier (Ed.), *The Cambridge handbook of cognitive linguistics*, Cambridge: Cambridge University Press.

Borowski, M. (2022). *Learning from use: An error-driven approach to Polish aspect*. Birmingham: University of Birmingham.

Bott, O., & Gattnar, A. (2015). The cross-linguistic processing of aspect–an eyetracking study on the time course of aspectual interpretation in Russian and German. *Language, Cognition and Neuroscience*, *30*, 877–898.

Bybee, J. (1988). Morphology as lexical organization. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology: Approaches in modern linguistics*, Leiden: Brill.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*, 425–455.

Clasmeier, C. (2015). *Die mentale Repräsentation von Aspektpartnerschaften russischer Verben*. Kubon & Sagner.

Croft, W. (2012). *Verbs. Aspect and causal structure*. Oxford: Oxford University Press.

Cvrček, V., & Fidler, M. (2017). Probing aspectual context with keyword analysis.

Dahl, Ö. (1985). *Tense and aspect systems*. Oxford: Blackwell.

Dahl, Ö., & Velupillai, V. (2013). Tense and aspect. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*, Leipzig: Max Planck Institute for Evolutionary Anthropology.

Divjak, D. (2010). *Structuring the lexicon*. De Gruyter Mouton.

Divjak, D. (2015). Four challenges for usage-based linguistics. Change of paradigms: New paradoxes. *Recontextualizing Language and Linguistics*, *31*, 297–309.

Divjak, D. (2019). *Frequency in language: Memory, attention and learning*. Cambridge: Cambridge University Press.

Divjak, D., & Milin, P. (2022). In *Ten lectures on language as cognition: A multi-method approach*, Leiden, Brill.

Divjak, D., Milin, P., Ez-Zizi, A., Józefowski, J., & Adam, C. (2021). What is learned from exposure: An error-driven approach to productivity in language. *Language, Cognition and Neuroscience*, *36*, 60–83.

Dowty, D. R. (1979). In *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*, Dordrecht, Reidel.

Dziób, A., Piasecki, M., Maziarz, M., Wieczorek, J., & Dobrowolska-Pigoń, M. (2017). Towards revised system of verb wordnet relations for Polish. In *Proceedings of the LDK workshops: OntoLex, TIAD and challenges for wordnets*.

Eckhoff, H. M., Janda, L. A., & Lyashevskaya, O. (2017). Predicting Russian aspect by frequency across genres. *The Slavic and East European Journal*, 844–875.

Feldman, L. B., Barac-Cikoja, D., & Kostić, A. (2002). Semantic aspects of morphological processing: Transparency effects in Serbian. *Memory & Cognition*, *30*, 629–636.

Gallistel, R. C. (2000). The replacement of general-purpose learning models with adaptively specialized learning modules. *The Cognitive Neurosciences*, *2*, 1179–1191.

Gattnar, A. (2013). Competitive types of verb in iterative contexts depending on the type and position of the quantifier. *Voprosy Yazykoznaniya*, 52–68.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology. General*, *113*, 256.

Goldberg, A. E., Casenhiser, D., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, *15*, 289–316.

Golovko, V. (2017). Deep learning: An overview and main paradigms. *Optical Memory and Neural Networks*, *26*, 1–17.

Gorbova, E. V. (2017). Русское видообразование: словоизменение, словоклассификация или набор квазиграммем?(еще раз о болевых точках русской аспектологии). Voprosy Jazykoznania, 24–52.

Grech, N. I. (1827). In *Prostrannaja russkaja grammatika [extended Russian grammar]*, Peterburg.

Grzegorczykowa, R., Laskowski, R., & Wróbel, H. (Eds.) (1998). *Gramatyka współczesnego języka polskiego. Morfologia*. Warszawa: Wydawnictwo Naukowe PWN.

Hanus, D. (2016). Causal reasoning versus associative learning: A useful dichotomy or a strawman battle in comparative psychology? *Journal of Comparative Psychology*, *130*, 241.

Herculano-Houzel, S. (2009). The human brain in numbers: A linearly scaled-up primate brain. *Frontiers in Human Neuroscience*, 31.

Irie, K., Faccio, F., & Schmidhuber, J. (2022). Neural differential equations for learning to program neural nets through continuous learning rules. *Advances in Neural Information Processing Systems*, *35*, 38614–38628.

Isačenko, A. V. (1960). *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim – Čast' vtoraja: Morfologija (the grammatical structure of Russian in comparison with Slovak – part two: Morphology)*. Bratislava: Izdatel'stvo akademii nauk.

Jakobson, R. (1932). *The structure of the Russian verb. Russian and Slavic grammar studies* (pp. 1931–1981). Berlin: Mouton.

Janda, L. A. (2007). Aspectual clusters of Russian verbs. *Studies in Language. International Journal Sponsored by the Foundation "Foundations of Language"*, *31*, 607–648.

Janda, L. A. (2007). *What makes Russian bi-aspectual verbs special. Cognitive paths into the Slavic domain*. Berlin: Mouton de Gruyter.

Janda, L. A., & Lyashevskaya, O. (2011). Grammatical profiles and the interaction of the lexicon with aspect. Tense, and mood in Russian.

Janda, L. A., & Reynolds, R. J. (2019). Construal vs. redundancy: Russian aspect in context. *Cognitive Linguistics*, *30*, 467–497.

Janda, L. A., & Tyers, F. M. (2021). Less is more: Why all paradigms are defective, and why that is a good thing. *Corpus Linguistics and Linguistic Theory*, *17*, 109–141.

Jebara, T. (2012). *Machine learning: Discriminative and generative*. Berlin: Springer.

Jovanović, T., Filipović-Đurđević, D., & Milin, P. (2008). Kognitivna obrada alomorfije u srpskom jeziku [the cognitive processing of the allomorphy in Serbian]. *Psihologija*, *41*, 87–101.

Kaeding, F. W. (1898). Häufigkeitswörterbuch der deutschen Sprache: Festgestellt durch einen Arbeitsausschuss der deutschen Stenographiesysteme. Selbstverlag des Herausgebers; E.S. Mittler & Sohn.

Karlsson, F. (1985). Paradigms and word forms. *Studia Gramatyczne*, *7*, 135–154.

Karlsson, F. (1986). Frequency considerations in morphology. *STUF – Language Typology and Universals*, *39*, 19–28.

Klimek-Jankowska, D., & Błaszczak, J. (2020). How incremental is the processing of perfective and imperfective aspect in Polish? An exploratory event-related potential study. *Journal of Slavic Linguistics*, *28*, 23–69.

Klimek-Jankowska, D., Czypionka, A., Witkowski, W., & Błaszczak, J. (2018). The time course of processing perfective and imperfective aspect in Polish: Evidence from self-paced reading and eye-tracking experiments. *Acta Linguistica Academica*, *65*, 293–351.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford: Stanford university press.

Langacker, R. W. (1987). Nouns and verbs. *Language*, *63*, 53–94.

Langacker, R. W. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics*, Amsterdam: John Benjamins.

Langacker, R. W. (2018). Morphology in cognitive grammar. In J. Audring & F. Masini (Eds.), *The Oxford handbook of morphological theory*, London: Oxford University Press.

Laskowski, R. (1984). Kategorie morfologiczne języka polskiego – charakterystyka funkcjonalna. In R. Grzegorczykowa, R. Laskowski, & H. Wróbel (Eds.), *Gramatyka współczesnego języka polskiego. Morfologia*, Warszawa PWN.

Lévêque, D., & Pellard, T. (2023). The implicative structure of Asama verb paradigms: A quantitative study of segmental and suprasegmental alternations. *Morphology*, 1–26.

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review*, 9, 829–835.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, 111, 309.

Lyashevskaya, O. (2013). *Frequency dictionary of inflectional paradigms: Core Russian vocabulary. Basic research program. Working papers. Humanities*. Moscow: National Research University Higher School of Economics (HSE).

Maiden, M., & O'Neill, P. (2010). Morphomic defectiveness. In M. Baerman, G. Corbett, & D. Brown (Eds.), *Defective paradigms: Missing forms and what they tell us*, London: OUP/British Academy.

Mansfield, J. (2015). Morphotactic variation, prosodic domains and the changing structure of the Murrinh-patha verb. *Asia-Pacific Language Variation*, 1, 163–189.

Maslov, J. S. (1948). Vid I leksiçeskoe znaçenie glagola v russkom yazyke. *Izvestiya Akademii Nauk SSSR. Otdelenie literatury i yazyka*, 7, 303–316.

Matthews, P. H. (1965). The inflectional component of a word-and-paradigm grammar. *Journal of Linguistics*, 1, 139–171.

Matthews, P. H. (1991). *Morphology*. Cambridge: Cambridge University Press.

Milin, P., Đurđević, D. F., & del Prado Martín, F. M. (2009). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language*, 60, 50–64.

Milin, P., Keuleers, E., & Đurđević, D. (2011). Allomorphic responses in Serbian pseudo-nouns as a result of analogical learning. *Acta Linguistica Hungarica*, 58, 65–84.

Milin, P., Divjak, D., Dimitrijević, S., & Baayen, R. H. (2016). Towards cognitively plausible data science in language research. *Cognitive Linguistics*, 27, 507–526.

Milin, P., Tucker, B. V., & Divjak, D. (2023). A learning perspective on the emergence of abstractions: The curious case of phone (me) s. *Language and Cognition*, 1–23.

Mondragon, E. (2022) 23 November 2022.

Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., & Baayen, R. H. (2004a). Morphological family size in a morphologically rich language: The case of Finnish compared with Dutch and Hebrew. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 30, 1271.

Moscoso del Prado Martın, F., Kostić, A., & Baayen, R. H. (2004b). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition*, 94, 1–18.

Osgood, C. E. (1966). Meaning cannot be rm? *Journal of Memory and Language*, 5, 402–407.

Osgood, C. E. (1980). *Lectures on language performance*. New York: Springer.

Percov, N. V. (1998). Russkij vid: Slovoizmenenie ili slovoobrazovanie. *Tipologija vida: Problemy, poiski, rešenija*, 343–355.

Poggio, T. (2012). The levels of understanding framework, revised. *Perception*, 41, 1017–1023.

Polinsky, M. (2008). Without aspect. In G. G. Corbett & M. Noona (Eds.), *Case and grammatical relations: Studies in honor of Bernard Comrie*, Amsterdam: John Benjamins.

Ramscar, M., & Port, R. (2015). Categorization (without categories). In E. Dabrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics*, Hague: The De Gruyter Mouton.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *The American Psychologist*, 43, 151–160.

Rescorla, R. A., & Wagner, R. A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In H. Black & W. F. Proksay (Eds.), *Classical conditioning II*, New York: Appleton-Century-Crofts.

Rissman, L., & Lupyan, G. (2022). A dissociation between conceptual prominence and explicit category learning: Evidence from agent and patient event roles. *Journal of Experimental Psychology. General*, 151, 1707.

Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review*, 26, 1850–1869.

Romain, L., Ez-zizi, A., Milin, P., & Divjak, D. (2022). What makes the past perfect and the future progressive? Experiential coordinates for a learnable, context-based model of tense and aspect. *Cognitive Linguistics*, *33*, 251–289.

Roussakova, M., Sai, S., Bogomolova, S., Guerassimov, D., Tangisheva, T., & Zaika, N. (2002). On the mental representation of Russian aspect relations. In S. Benjaballah, W. U. Dressler, O. E. Pfeiffer, & M. D. Voeikova (Eds.), *Morphology 2000. Selected papers from the 9th Morphology Meeting, Vienna, 24–28 February 2000*, Amsterdam: Benjamins.

Rumelhart, D. E., McClelland, J. L., & PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1: Foundations*. Cambridge: MIT press.

Sasse, H.-J. (2002). Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology*, *6*, 199–271.

Shirai, Y., & Andersen, R. W. (1995). The acquisition of tense/aspect morphology: A prototype account. *Language*, *71*, 743–762.

Sims, A. D. (2015). *Inflectional defectiveness*. Cambridge: Cambridge University Press.

Soveri, A., Lehtonen, M., & Laine, M. J. (2007). Word frequency and morphological processing in Finnish revisited. *The Mental Lexicon*, *2*, 359–385.

Steinfeldt, E. (1963). *Russian word count: 2500 words most commonly used in modern literary Russian; guide for teachers of Russian*. Moscow, Progress: Publishers.

Stoll, S. (2001). The acquition of Russian aspect. *In:* University of California, B., Faculty of Arts. (ed.).

Stoll, S., Bickel, B., Lieven, E., Paudyal, N. P., Banjade, G., Bhatta, T. N., Gaenszle, M., Pettigrew, J., Rai, I. P., & Rai, M. (2012). Nouns and verbs in Chintang: Children's usage and surrounding adult speech. *Journal of Child Language*, *39*, 284–321.

Stump, G., & Finkel, R. A. (2013). *Morphological typology: From word to paradigm*. Cambridge: Cambridge University Press.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*, 49–100.

Thorndike, E. L., & Lorge, I. (1944). The teacher's word book of 30,000 words.

Townsend, C. (1975). In *Russian word-formation*, Columbus, Slavica.

Traficante, D., & Burani, C. (2003). Visual processing of Italian verbs and adjectives: The role of the inflectional family size. In H. R. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing*, Berlin: Mouton de Gruyter.

Varda, A. G. d., & Marelli, M. (2023). Data-driven cross-lingual syntax: An agreement study with massively multilingual models. *Computational Linguistics*, *49*, 261–299.

Vendler, Z. (1957). Verbs and times. *Philosophical Review*, *66*, 143–160.

Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. *WESCON Convention Record Part IV*, 96–104.

Widrow, B., & Lehr, M. A. (1990). 30 years of adaptive neural networks: Perceptron, madaline, and back-propagation. *Proceedings of the IEEE*, *78*, 1415–1442.

Wilmoth, S., & Mansfield, J. (2021). Inflectional predictability and prosodic morphology in Pitjantjatjara and Yankunytjatjara. *Morphology*, *31*, 355–381.

Wonnacott, E., Newport, E. L., & Tanenhaus, M. K. (2008). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*, *56*, 165–209.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Reading: Addison-Wesley.

Bickerton, D. (1981). *Roots of language*. Ann Arbor: Karoma publishers.

Hopper, P. J. (1982). *Tense-Aspect: Between Semantics and Pragmatics*. Amsterdam: John Benjamins.

Bardovi-Harlig, K. (2000). Tense and aspect in second language acquisition: Form, meaning, and use. *Language Learning: A Journal of Research in Language Studies*, *50*, 1.

Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics*, *14*, 161–219.

Divjak, D. (2017). The Role of Lexical Frequency in the Acceptability of Syntactic Variants: Evidence From that-Clauses in Polish. *Cognitive Science*, *41*, 354–382. https://doi.org/10.1111/cogs.12335.

Friedman, W. J. (1978). *Development of time concepts in children. Advances in child development and behavior*. Amsterdam: Elsevier.

Authors and Affiliations

Dagmar Divjak[1,2] · Irene Testini[1] · Petar Milin[1]

✉ D. Divjak
d.divjak@bham.ac.uk

I. Testini
i.testini@bham.ac.uk

P. Milin
p.milin@bham.ac.uk

[1] Department of Modern Languages, University of Birmingham, Edgbaston, Ashley Building, B15 2TT, Birmingham, UK

[2] Department of English Language & Linguistics, University of Birmingham, Birmingham, UK