Daniela Katunar
Faculty of Humanities and Social Sciences, University of Zagreb
dkatunar@ffzg.unizg.hr

Weather domain in Croatian: a corpus–based overview of precipitation and non–precipitation expressions

This study investigates weather expressions in Croatian, focusing on precipitation and non–precipitation weather phenomena. It highlights the complexity of the weather domain as described by extant syntactic, semantic and lexical studies showing variation both inter– and intralinguistically, and as the focus of future investigations. Croatian is examined with respect to proposed typologies of weather expressions, built around the notion of coding weather occurrences with the predicate or the argument of the sentence, or both, as well as investigations of lexicalization patterns of weather verbs with regards to Figure, Path and Manner. Though such classifications are based on cross–linguistic comparisons, their applications for systematizing intralinguistic variability are investigated on Croatian data. Since descriptions of weather expressions in Croatian have mostly focused on a subset of the most common weather verbs, *verba meteorologica*, a corpus–based approach is used to expand the dataset with different types of syntactic structures and predicates used in describing weather events, complemented by lexicographic sources. Weather expressions in Croatian are analyzed according to the semantic properties of dynamic and static weather phenomena, contextual grounding of a weather expression, manner of occurrence and the notions of primary and secondary weather expressions. Based on the overview of lexicographic and corpus data, this paper sketches an outline for future systematic study of the weather domain in Croatian.

---

1. Introduction

Weather can be viewed as a universal human experience since it is a domain intricately related to the environment we live in and the impact it has on our cultural practices and beliefs. It is oftentimes a silent companion to our daily activities, a very salient obstacle to them when it "becomes bad", a topic of daily news reports or small talk interactions, and sometimes a violent unpredictable force to be reckoned with. Therefore, it is not surprising that linguistic investigations of the weather domain have pointed out its relevance in understanding the way languages name and construe the world around us. Weather phenomena have been proposed as potential lexical universals, e.g., *rain* is featured in Swadesh's basic vocabulary list (Swadesh 1972; cf. Goddard 2001). The peculiar syntactic status of weather verbs and weather expressions has often been the focus of studies related to weather

predicates, with studies trying to tackle the lack of a clear subject in sentences such as *it's raining* (e.g., Bolinger 1973; Jackendoff 1985; Belaj and Tanacković Faletar 2015)[1], or impersonal "subjectless" expressions such as It. *piove*.3SG, Cro. *kiši*.3SG 'it's raining' (Ruwet 1986; Belaj 2007; Bleotu 2012). Studies also focused on the argument structure representations of weather verbs in terms of describing weather verbs as unaccusative or unergative (if their subject is to be understood as a Patient or an Agent, Bleotu 2012: 61), with many syntactic tests pointing to a tendency towards the unaccusative hypothesis in most, but not all cases (Ruwet 1989), or alternative construals of weather events as basis for their variation (Levin and Krejci 2019). The nature of such discussions has usually focused on the observation pointed out in Ruwet (1986: 206) that "the strange character of weather expressions in general is due to a conflict between the analytical nature of syntax and the global way in which we experience atmospheric phenomena", in the sense that it is difficult to distinguish the event itself from its participants (i.e. to distinguish 'rain' from 'raining'). This is reified in Langacker's (1991: 366) description of rain, snow, heat, cold, fog etc. as phenomena that lack the clear–cut division between setting, process and discrete participants. In terms of participants and events, Meulleman and Paykin (2016) discuss possible conflations of Figure, Path or Manner in weather verbs based on Talmy's (1985, 2000) notion of lexicalization patterns and conclude that the weather verb category presents challenges to extant classifications in terms of V–framed and S–framed languages due to its heterogeneity within and among different languages.[2] Recently, classifications and typologies of weather expressions and languages have been proposed based on the syntactic (or argument structure) type languages employ as a canonical weather expression in order to present a systematization of the "bare" elements needed to describe a weather event in a language, e.g., the predicate type, It. *piove*.3SG, the argument type, Ru. *идёт снег* 'it is snowing; lit. go.3SG snow', or both, e.g., *le vent souffle* 'wind blows' (Eriksen et al. 2010; Eriksen et al. 2012).

In light of such studies, one of our main goals is to expand extant data from various languages with data from Croatian. As to our knowledge, few studies have dealt with Croatian weather expressions (cf. Belaj 2007; Belaj and Tanacković Faletar 2015)[3], and descriptions of weather expressions in Croatian do not typically go beyond discussion of the properties of impersonal verbs. Thus Croatian, at face value, seems like a typical predicate type language, having impersonal verbs which

---

1   Depending on the author, *it* can be analyzed as an expletive subject, or an element that denotes ambient conditions, see Levin and Krejci (2019).

2   Talmy's (1985, 2000) division relies on distinguishing typologically relevant tendencies of languages as to what element of a Motion (or Location) event is lexicalized by the verb, and what is lexicalized via satellites, e.g., prefixes, prepositions and cases in Croatian, compare Spanish *La bottella entró a la cueva (flotando)* 'lit. the bottle entered the cave (floating) (Path)' and Cro. *Boca je u–plutala u pećinu*.Acc 'the bottle floated into the cave (Manner)'. When it comes to *verba meteorologica*, the question is whether a weather verb conflates 'rain – Figure', 'falling – Path' or 'dispersed type of motion of particles typical of rain – Manner'. More on this in section 4.

3   Most descriptions can be found in Croatian grammars, and they do not take an in–depth approach to an analysis of weather verbs or expressions per se.

code weather events, e.g., *kiši*.3SG 'it's raining'. However, as we will show, this does not hold for other precipitation phenomena, compare Engl. *it's hailing*, Cro. **tuči*. As Meulleman and Paykin (2016) point out, oftentimes studies of the argument structure of weather verbs focus only on a few examples, or just one – rain – thus promoting it to the status of the most prototypical weather verb, possibly characteristic of the whole class. In line with their observation, we believe a study that expands its investigation to include other precipitation phenomena (e.g., sleet, frost), other hydrometeors (e.g., fog) or complex weather events (such as storms) may contribute to the description of differences *within* a class of phenomena, such as weather phenomena and events. Also, unlike extant studies, we will base our observations on corpus data from the *Croatian web corpus* (hrWaC) (Ljubešić and Klubička 2014), a methodology expected to highlight usage–based intralinguistic variability of what could be considered competing weather expressions[4], e.g., *kiši*.3SG and *pada kiša* 'lit. fall.3SG rain', and expand the list of weather verbs and weather expressions with, for example, weather verbs which code manner of precipitation, e.g., *rominjati* 'drizzle with a slight hum'. This aligns with our second goal, to sketch out the semantics of weather in Croatian in the way that is it represented in the lexicalization of weather events and phenomena and provide guidelines for future research.

2. Theoretical background

2.1 Weather domain and weather expressions

At the onset of this study, it is important to define what exactly is meant by the terms *weather domain* and *weather expression*. In Croatian, the noun *vrijeme* is polysemous and can refer to the concepts of 'time' and 'weather'.[5] Some derivatives show specialization only for the temporal meaning, e.g., *vremenit* 'temporal; timely; old', *vremešan* 'elderly, old', while the atmospherical meaning is contextually distinguished in expressions such as *Kakvo je vrijeme?* 'What is the weather like?', *lijepo vrijeme* 'nice weather', *ružno vrijeme* 'bad (lit. ugly) weather' or collocations such as *vremenske prilike* 'weather conditions', *vremenska nepogoda* 'any type of natural weather disaster'[6] or *vremenska prognoza* 'weather report'. The morphological antonym, *nevrijeme*, is specialized to mean 'severe weather' and is often synonymously used with *oluja* 'storm', or more generally includes any type of a dynamic and adverse weather occurrence specified by a weather–related adjective, e.g., *grmljavinsko nevrijeme* 'thunderstorm', *olujno nevrijeme* 'stormy bad weather with strong winds', *snježno nevrijeme* 'a snowstorm'.

---

4   We will expand upon the notion in the following sections of the paper.
5   Etymologically, it is related to the IE. root *wert– 'to spin' and is diachronically connected to the notions of motions and movement (Gluhak 1993). Compare Bulg. *vrème*, Mac. *vreme*, Slov. *vreme*, Serb. *vreme* (Gluhak 1993).
6   Cf. Polish *pogoda* 'weather; good weather' (Andrason 2019), Cro. **pogoda*.

Struna (Croatian Special Field Terminology) defines weather as an "atmospherical state and phenomena in a certain place and a period not longer than a few months", while the meteorological definition in the dictionary *Croatian Language Portal* defines weather as a "state of the atmosphere in a given moment which is defined by values of the meteorological elements and weather phenomena". The meteorological elements and phenomena refer to experientially familiar notions such as precipitation, winds, clouds, humidity, air temperature and the like. The weather domain is therefore one the basic features of the environment that surrounds us and we are constantly immersed in it. Particular weather phenomena are culturally salient and intricately connected to our knowledge of the world, for example, passing of the seasons (e.g., *Sveta Kata, snijeg na vrata* 'St Kate[7], snow at the gate'), regional climate features (e.g., *bura* 'bora, an Adriatic strong northern wind'), and effects that weather phenomena have on our mood and behavior (e.g., *južina* 'a weather state when a SE wind (*jugo*) blows and people feel uncomfortable and drowsy'). The weather lexicon referring to weather phenomena can be divided along the lines of Eriksen et al. (2012: 17) as referring to a) precipitation phenomena, e.g., *kiša* 'rain', *snijeg* 'snow', *tuča, grad* 'hail' etc., b) non–precipitation phenomena, e.g., *vjetar* 'wind', *grom* 'thunder', *munja* 'lightning', c) temperature, atmosphere and light conditions, e.g., *hladnoća* 'the cold', *vrućina* 'the heat', *dan* 'day', *danje svjetlo* 'daylight', *mrak* 'dark', *vlaga* 'humidity' and d) sunshine, *sunce* 'sun'. To this we can add complex weather events[8], e.g., *oluja*, *nevrijeme* 'storm', which include the co–occurrence of various types of precipitation and non–precipitation phenomena (clouds, rain, winds, thunder).

Weather expressions, in the broadest sense, are ways in which we talk about the weather, and pertain to language structures available to speakers in describing weather phenomena and weather events. These structures constitute the weather vocabulary of a language on the one hand, i.e., naming various types of weather phenomena and processes, and on the other, syntactic structures available to speakers in describing a weather event. Both can serve as a research focus in discussing the inter– and intralinguistic variability of weather expressions. One example of this would be building a (folk) taxonomy of terms for particular weather phenomena, and observing the differences in the number of terms and lexical gaps for a particular phenomenon (for example, various terms for snow, its hyponyms and meronymy, e.g., *snowflake*, and dialectal variants).[9] Another perspective, one

---

7   The holiday of St Catherine in the Roman Catholic calendar, 25th of November.
8   Cf. Andrason (2019: 71, 93–94), where the term *mixed events* is used for two meteorological phenomena which co–occur simultaneously, as coded in syntactic constructions, e.g., *snow is blowing*. To distinguish his syntactic perspective from the nominal descriptions of phenomena lexicalizing co–occurrence of two or more meteorological phenomena (e.g., a lexeme like *storm* can include a plethora of these and be conceptualized as a Gestalt at the same time), we use the term complex events and phenomena (cf. Meulleman and Paykin 2016).
9   For example, Babić (1982) list 22 terms for snow in Croatian, although most of them are attested in hrWaC only once, and many are dialectally restricted or archaic, e.g., *rodinjak* 'springtime snow', *škvrljinjak* 'February snow (Slavonia region)'.

that most commonly utilizes the term *weather expression*[10], investigates weather expressions as argument structures of a weather event, with a focus on weather verbs and weather predicates. As Eriksen et al. (2010) point out, weather expressions are interesting from the point of view of argument structure, as they code events without proper participants in various ways, as in 1.a–i).

1)

Latin

    a) *pluit*. (Kienpointner 2016: 57)
        rain.3SG
        'It is raining'

    b) *Iuppiter pluit*. (Kienpointner 2016: 60; cf. Eriksen et al. 2010)
        Jupiter rain.3SG
        'Jupiter rains/Jupiter sends rain'

Polish

    c) *Pada   deszcz*. (Andrason 2019: 72)
        fall.3SG rain.NOM.SG
        'It is raining'

    d) *Kapa–ło   /la–ło*. (Andrason 2019: 74)
        drip.PAST.3SG.N / pour.PAST.3SG.N
        'it was raining a little / intensively'

Icelandic (Eythórsson and Sæunn Sigurðardóttir 2016: 92)

    e) *Vindinn   hvessir*.
        the–wind.ACC gets–windy
        'It gets windy'

    f) *Það / Hann rignir mikið í dag*
        it / he   rains much today
        'it rains a lot today'

Palestinian Arabic (Givón 2001: 119)

    g) *id–dunya ti–shti*
        the–world 3SG.MASC–rain
        'it is raining (lit. the world is raining)'

Turkish (Kienpointner 2016: 61)

    h) *Yağmur yağıyor*.
        rain.NOM is raining
        'It is raining'

    i) *Kar   yağıyor.*
        snow.NOM is raining
        'It is snowing'

---

10   Or meteorological expression, cf. Eriksen et al. (2010).

As the examples show, languages can code weather events in various ways, by omitting the subject altogether (as in a) and d)), by inserting expletive elements (f), locational background entities (g)[11] or deities (b) in lieu of a referential subject, or by the use of cognate subjects (h). Thus, when looking at the weather lexicon in general we can divide our attention between the taxonomy of weather terms and naming of weather phenomena and the argument structures present when describing a weather event. In this study we focused on both, in the sense of examining uses of weather terms (rain, fog, frost etc.) in the corpus, but in the service of investigating types of predicates and verbs which occur with specific weather phenomena.

2.2 Classifications of weather expressions

Beginning with Ruwet (1986) and continuing in recent investigations of weather expressions (Eriksen et al. 2010, 2012; Andrason 2019), several classifications of weather expressions and their typologies have been proposed. Ruwet (1986: 203–204) lists six solutions languages have at their disposal when describing a weather event:

a) a null subject and a semantically dominant predicate (It. *piove* 'rain.3SG', Fr. *il pleut* 'it.expletive rain.3SG'),

b) purely nominal sentences (e.g., *Thunder!* for 'it's thundering'),

c) a lexical subject and an empty or almost empty verb (e.g., Basque *urra da* 'rain is', Fr. *il y a du vent* 'there is wind'),

d) the verb duplicates the semantic content of the subject in various ways (Fr. *le vent souffle* 'wind is blowing', Eng. *thunder is roaring*),

e) reduplication of subject and predicate (e.g., Fr. *le tonnerre tonne* 'thunder is thundering', Tur. *yağmur yağıyor* 'rain rains'[12])

f) a much more analytical representation of a weather event, conceivable but probably rare (e.g., *water is descending from the sky*).

More recently, a formal typology of weather expressions has been proposed based on data from various languages[13] (Eriksen at al. 2010, 2012), which divides weather expressions into three main types depending on the element that is used to convey the weather event:

a) the predicate type, e.g., It. *piove*, Ger. *es regnet* 'it is raining', Pol. *jest chłodno* 'it is cold; lit. be.AUX.3SG cold',

b) the argument type, e.g., Ru. *идёт снег* 'it is snowing; lit. go.3SG snow'[14], Eng. *there is a storm*,

c) the argument–predicate type, e.g., Toqabaqita *thato e thato* 'the sun is shining' (cognate type), Eng. *wind blows* (split type).

---

11   A term taken from Eriksen et al. (2010). Givón (2001: 119) calls them dummy subjects.

12   Cf. Kienpointner (2016).

13   25 languages from various language families (e.g., Turkic, Sino–Tibetan, Uralic, Indoeuropean, Austronesian, Niger–Congo, East Bougainville, Korean). Croatian was not included in the reported data.

14   To avoid different Latin transcriptions various authors use of Russian examples we will use Cyrillic notation.

The authors introduce subtypes based on parts of speech variation (e.g., verbal or adverbial predicate, as in a), types of constructions used (atransitive[15], as in a), intransitive or existential, as in b), and further subdivide the argument–predicate type into the *cognate type* (same morphological form for the argument and the predicate) and the *split type*, where the argument and the predicate express different facets of the same event, as in c). Similarly, a meteo–scale has been proposed by Kienpointner (2016), which classifies weather expressions as realizing either the Phenomenon pole, i.e., by means of pure verb V, Lat. *pluit*, an Entity Pole by means of a noun + "weak" verb, Ru. *идёт дождь* 'goes rain', and an area in between, Eng. *it's raining*, Ger. *der wind weht* 'the wind blows'.

Such classifications are built around what the authors consider to be basic ways to say what is happening weather–wise at a given moment, with a *basic weather expression* as described by Kienpointner (2016) or *canonical* uses of meteorological constructions as described by Eriksen et al. (2012), though the two may differ in definition somewhat. According to Kienpointner (2016), basic weather expressions are unmarked, more frequent, usually shorter, syntactically less complex and semantically simpler representations of meteorological phenomena then their marked counterparts (e.g., Ger. *es regnet* 'it's raining' vs. *der Regen fällt* 'rain is falling'). Eriksen et al. (2010, 2012), on the other hand, use the term *canonical uses* of weather expressions/verbs to exclude non–canonical examples of weather verbs which may differ from canonical uses in the sense that they add participants which are not integral to the weather event itself. As can be seen from the classifications of weather expressions, not all classifications correlate completely with one another, but they do point to a basic tripartite way of coding weather events by means of semantically assigning the bulk of the weather information on either the predicate or the argument of the sentence, or both.

2.3 Weather verbs in Croatian

Since *verba meteorologica*, weather verbs such as *kišiti* 'to rain', *sniježiti* 'to snow', *grmjeti* 'to thunder', *sijevati* 'to flash (of lightning)' received the most attention in extant syntactic studies of Croatian (Samardžija 1986; Katičić 2002; Belaj 2007; Belaj and Tanacković Faletar) they form a logical starting point. From the point of view of valency, Samardžija (1988) and Šojat (2008) describe them as *non–valent* (or *avalent* in Tesnière's terms) and treat them as one the four possible classes of verbs according to their valency (along with one valent, two valent and three valent verbs). In Croatian grammars, weather verbs are provided as typical examples of

---

15   We take over the term *atransitive* as used by Eriksen et al. (2010, 2012) to refer to constructions not requiring a subject position to be filled at all, and it can be used synonymously with *avalent* verbs (see below). From a syntactic perspective this speaks to the acceptability of such examples without a subject or object, which are coded in Croatian by the nominative and accusative case, respectively. From a semantic perspective, there are other "peripheral" semantic roles typical of atransitive sentences which do speak to a broader argument structure context of atransitive constructions, such as Location and Time, being filled by adverbials, e.g., *U Zagrebu kiši* 'it is raining in Zagreb', *danas je hladno* 'it is cold today'.

impersonal verbs (Cro. *bezlični glagoli*), lacking (or not requiring) a subject, and are distinguished from verbs which can undergo impersonalization on the sentence level. Pranjković (2020: 12) regards *verba meteorologica* as prototypical verbs of happening, i.e., a class of verbs which denotes involuntary or non–agentive processes in nature or people, and their subject, if realized, is redundant. Regarding weather verbs, Barić et al. (1997: 427) point out that they can be personal or impersonal, depending on whether the subject is realized in the sentence or not; if the subject is realized it is not an agent (cf. Katičić 2002). Silić and Pranjković (2007: 316–317) describe impersonal, ('unanalyzed', single member) sentences as those sentences which occur either with impersonal verbs, e.g., *grmi* 'it is thundering' or an auxiliary verb (copula) + adverbial, e.g., *danas je vedro* 'it is sunny today', *ondje je bilo baš zagušljivo* 'it was very stuffy there'. Omission of the subject, according to the authors, can be motivated by redundancy or the fact that the subject is non–existent or cannot be named for some reason. One test to determine the presence of the impersonal verb is based on gender agreement of the active adjectival participle (Barić et. al. 1997: 427), as in 2).

2)
a) *danas je   puhao   vjetar*
   today AUX.be blow.3SG.M wind.NOM.SG.M
   'wind was blowing today'
b) *danas je   puhalo.*
   today AUX.be blow.3SG.N
   'it was windy today'

Belaj (2007) criticizes the 'either – or' stance towards impersonal verbs and argues that a distinction can be made between personal–impersonal (a semantic and pragmatic category) and subject–subjectless (a syntactic category) and introduces the notion of desubjectivization which can be observed in common weather verbs, among other examples. Introducing the notion of semantic and morphological redundancy in his study (e.g., *kišiti* 'to rain' – *kiša* 'rain') also allows for insights into variability between weather verbs, most prominently the insight that some verbs may allow for a greater degree of participant awareness and thus the possibility for participants to occur is higher (e.g., *padati* 'to fall/rain', compared to *kišiti* 'to rain'). We agree with Belaj's stance, based on Langacker's (1991) description of weather constructions, that the semantics and pragmatics of weather expressions (and more broadly impersonal verbs) motivate their syntactic realizations.[16] Indeed, one point this study aims to expand upon is what is a "higher possibility" in terms of corpus–based examples and lexicalization of particular weather features, and what

---

16   See Belaj and Tanacković Faletar (2015) for a reiteration and expansion of the analysis in the framework of Cognitive Grammar, with a focus on the role of the medial demonstrative *to* in allowing speakers alternate conceptualizations of weather events with respect to the setting–subject category, e.g. *jel' to padalo*.3SG.N/ *padala*.3SG.F [kiša.3.SGF] 'did it rain'.

possible factors may influence participant awareness when it comes to such examples, as the extant studies do not base their insights on a corpus–based approach and data. As for morphological redundancy, Croatian does not use the cognate type (subject and verb with the same root, e.g., *kiša kiši* 'rain is raining, as confirmed by only 5 examples in the corpus, mostly with a stylistic function, but uses of the split type (e.g., *vjetar puše* 'wind blows') are not uncommon (more below).

There are several questions different from that of the syntactic nature of impersonal verbs that are not addressed in previous studies of Croatian, as to our knowledge, notably a) focus on the weather domain in general, with studies focusing either on parts of the noun vocabulary and etymologies of weather terms on the one hand[17], or weather verbs as part of the broader category of impersonal verbs on the other, b) apart from the few main (non–)precipitation types (snow, rain, thunder, lightning), what do different precipitation (and non–precipitation) phenomena use as the typical way of expressing the occurrence of a specific weather phenomenon (e.g., *pada tuča / *tuči* 'it's hailing'), c) a comprehensive compiled list both of weather phenomena and weather predicates, d) an examination of the main lexicalization strategies and conceptualizations of weather events and phenomena. Therefore, one of our main goals is to present this data in the present study.

3. Weather and corpus data

In this study we limit our attention to precipitation and 'non–precipitation' phenomena (cf. Eriksen et al. 2012), and leave temperature, atmospheric conditions, and stages of day for future investigations. We base our observations on data from the *Croatian Web Corpus* (hrWaC), the largest general corpus of Croatian with over 1 billion tokens, containing data from various standard and colloquial varieties of Croatian. The corpus approach consisted of a two–way analysis of weather phenomena. On the one hand, we investigated weather verbs discussed in extant literature, e.g., *kišiti* 'to rain', and on the other, we investigated verb collocates of weather phenomena in their noun forms, e.g., *kiša* 'rain'.[18] The latter approach allowed us to expand the list of weather verb candidates with various verbs with specialized meanings, e.g., *rominjati* 'drizzle with a slight hum', *bridjeti* 'to blow with intensity, searing the skin with cold', which are not included in abovementioned reference grammars as examples of *verba meteorologica*. This, in turn, allowed us to investigate corpus uses of such verbs and investigate what constructions (atransitive, intransitive) are typical of them, thus allowing us to examine their usage tendencies when it comes to the predicate or argument–predicate types discussed

---

17 Cf. Poje (2002), Gluhak (2004), Belaj (2007), Belaj and Tanacković Faletar (2015).
18 The starting point was to compile a list based on extant literature, reference grammars, lexicographic sources, as well as meteorological sources which list, e.g., precipitation types (cf. Penzar and Penzar 1992). When it comes to the list of precipitation and non–precipitation phenomena, we do not consider it exhaustive, but representative of various phenomena.

above. Moreover, this approach allowed us to investigate which precipitation and non–precipitation phenomena have verbs dedicated to describing the occurrence of the phenomenon, and which do not (e.g., *mraz* 'frost'), and what syntactic constructions are used instead to denote the occurrence of the phenomenon. This is related to our interest in investigating differences *within* a class of phenomena and serves as a basis to propose some generalizations of semantic features of weather phenomena, as reflected in their uses and syntactic constructions serving as primary ways of denoting the occurrence of weather phenomena.

Some caveats of corpus research have to be mentioned at the onset. One is that weather discourse is often contextually dependent, and therefore written corpus data may differ from that of spoken language, though we do not find great differences in examples from, e.g., forums, blogs and the like contained in the corpus. Another caveat is that most of the lexemes examined, especially verbs, are polysemous units (see section 4 below), and therefore quantitative data was supplemented with qualitative procedures, word sense disambiguation and sampling in some cases. However, corpus data are relevant as they provide a basis for the overall usage tendencies of weather expressions, attestations of syntactic types and their frequency, as well as contextual data itself.

Finally, when it comes to coding strategies of weather events and phenomena, Croatian has a variety of ways to express what is happening weather–wise, as do other languages, and one can find examples of all of these coding strategies in corpus examples (cf. Eriksen et. al 2010, 2012), with fine–grained distinctions usually a typical part of intralinguistic variability, as in 3):

3)

a) the predicate type
Kiši. (impersonal verb, atransitive)
rain.3SG
'It is raining'

Jutro *je kišno i* *tmurno*. (intransitive, adjectival predicate)
morning.NOM AUX.be rainy and gloomy
'The morning is rainy and gloomy'

Lije *ko iz* *kabla*. (simile idiom)
pour.3SG like from bucket
'It is pouring like from a bucket'

b) argument–predicate type
Pada kiša. (intransitive)
fall.3SG rain.NOM.SG
'it is raining / rain is falling'

Kiša *tiho* *rominja*. (intransitive, manner of raining)
rain.NOM.SG silently drizzle.3SG
'Rain is drizzling silently'

c) the argument type

U Zagrebu je   kiša. (intransitive, basic locative/existential construction)
in Zagreb.LOC AUX.be rain.NOM
'There is rain in Zagreb'

Bit će kiše   (uskoro). (existential construction, future tense)
be AUX.have rain.GEN (soon)
'There will be rain (soon)'

Dolaze jesenske kiše. (intransitive, motion construction)
arrive.3PL autumn.NOM.PL rain.NOM.PL
'Autumn rains are coming'

Kiša   je   stala. (intrasitive, phase verb)
rain.NOM.SG AUX.be stop.3SG
'The rain stopped'

Not all of these examples are equally relevant when placing Croatian in the above–mentioned typologies, as they are not all "basic" ways of denoting the occurrence of a weather phenomenon. In order to systematize the approach to weather–related data in the corpus (and the lexicon), we propose a tentative classification of expressions as follows, for the purposes of distinguishing future avenues of research:

a)   structures that are pertinent to correlating Croatian data to proposed typologies of weather expressions (basic weather expressions), e.g., atransitive *verba meteorologica* or adverbial predicates (cf. Eriksen et al. 2010),

b)   typical and frequent ways to describe some aspect of its occurrence, e.g., phases *počela/stala je kiša* 'rain began/stopped',

c)   verbs (and other parts of speech) forming a specialized weather–related lexicon, e.g., of snow: *okopnjeti* 'to thaw (only of ice and snow)', *prtiti (snijeg)* 'to make a path in the snow by walking or showeling', *zamesti* 'cover up (of snow); snow in' (compare *zamet* 'snowdrift made by wind'). We will refer to b) and c) jointly as secondary weather expressions in this paper.

d)   morphosemantic relations, such as *kisnuti* 'to get rained upon', *prokišnjavati* 'to leak (as in a leaky roof)' < *kiša* 'rain', relevant in a morphologically rich language such as Croatian.[19]

e)   Investigations of ways weather is conceptualized in a language and culture, but which do not necessarily refer to weather conditions proper, e.g., conceptualizing natural forces as causes or inanimate agents (or effectors), e.g., *turiste je iz zemlje potjerala i kiša* 'rain, as well, drove the tourists out of the country'. Idiomatic and metaphorical uses of weather expressions, which use weather as the source domain for describing various target domains, are relevant for this investigation, e.g., *izgledao je pokisnuto* 'he looked sad (lit. rained upon)', *pljuštale su pohvale* 'praises were pouring'.

---

19   Cf. Raffaelli and Kerovec (2008).

While all of the above–mentioned examples can be viewed as relevant when describing the weather domain in Croatian (in terms of what weather is, how we categorize weather phenomena, how we conceptualize weather and how we use it to conceptualize other domains), not all of those can fit within the limits of this study. We mention them because one caveat of corpus–based research of weather is that one finds a mix of these categories with varying degrees of frequency, and for this reason in our analysis we focus on weather expressions proper, verbs (or verb senses) which denote the occurrence of weather phenomena, and point and treat as secondary those examples which are frequent or have specialized weather–related meanings and thus we considered them important to mention from a descriptive point of view. We are leaving morphosemantic, idiomatic or metaphorical uses aside for future study.

4. Precipitation and 'non–precipitation' phenomena in Croatian

It must be said at the onset of this section that weather vocabulary in Croatian shows an overlap between scientific and folk taxonomies. Many terms are used both in specialized terminologies and everyday use (e.g., rain, snow, fog), while some terms are restricted to one or the other type of language use (e.g., hydrometeors). For the collection of weather terms in this study we focused mostly on the general Croatian vocabulary found in lexicographic sources, notably the *Croatian Language Portal*, *Dictionary of Croatian language* (Šonje 2000) and the *Great Dictionary of the Croatian Standard Language (VRH)* (2015). We also consulted meteorological and terminological sources such as the *Meteorological glossary and multilingual dictionary* (Gelo et al. 2005), an extensive and useful resource which contains a list of archaic and regional terms as well, the Croatian version of the *International cloud atlas* (Katušin 2007) and *Struna*. This provided us with a basis for a comprehensive overview of the main precipitation and non–precipitation terms to be used in the corpus investigation, though we do not consider the list to be exhaustive, particularly because of regional varieties which may be underrepresented in lexicographic and corpus data. In the first part of these sections, we will present the list of these terms, as they were used in our corpus research to expand the list of potential weather verbs which co–occur with them and those which have verbal counterparts, and to investigate their occurrence in various syntactic structures pertaining to the description of weather events. We also find such a list useful for future investigations of the weather vocabulary in Croatian and its lexicalization patterns.

In Croatian, the superordinate category of precipitation is lexicalized as *oborina*, a term recommended by meteorologists, while the synonym *padalina* is in use in geographic and philological literature (Penzar and Penzar 1992), as well as everyday language, and sources vary as to their listing and recommendations of use as synonyms.[20] A broader category which precipitation phenomena belong to are

---

20   A third term, *padavina*, is also listed as a synonym, but is in limited use (freq 0.31).

hydrometeors[21], which also include, e.g., *magla* 'fog', *sumaglica* 'mist', *virga* 'virga', *snježna vijavica* 'drifting and blowing snow', *morski dim* 'sea spray' and so forth (Gelo et al. 2005). We included some of them in the list below because of high frequencies and experiential similarity to precipitation (e.g., fog). The notion of precipitation, as described in the *Glossary*, can be used in the broader or narrower sense, the latter including only precipitation which falls from the clouds (rain, snow), and the former adding precipitation condensing near the ground or ground objects (frost, rime), though most sources use it in the broader sense. This experientially based division can be tied to the profiling of some precipitation as dynamic, and some as static (see below). Frequencies of weather terms vary, as shown in Table 1.[22]

HYDROMETEORS	*hidrometeori*	frequency	
PRECIPITATION	*oborina*	5.12	
	padalina	3.17	
RAIN	*kiša*	57.75	
	dažd (archaic)	0.14	
Rain – types	*pljusak* 'rainshower'	4.01	
	škropac 'rainshower (regional); light short rain'	0.01	
	prolom (provala) oblaka 'cloud burst'	0.45/(0.01)	
	kišica.DIM 'light rain'	0.89	
	[[Adj] *kiša*], e.g., *blatna/monsunska/ledena kiša* 'mud/monsoon/freezing rain'	10.88	
SNOW[23]	*snijeg*	49.02	
Snow – types	*pršić* 'powder snow'	0.08	
	sniježić.DIM, *sniježak*.DIM 'light snow'	0.01	
	[[Adj] *snijeg*], e.g., *vječni/mokri/žuti snijeg* 'everlasting/wet/yellow snow', *zrnati snijeg* 'snow grains'	8.46	

---

21   A term typical of the scientific taxonomy, but not of everyday use.

22   Of course, we have to point out that these token frequencies do include, for example, metaphorical extensions of some terms, e.g., *vihor rata* 'the whirlwind of war', collocations such as *kiša meteora* 'meteor shower; lit. rain of meteors', or collocations denoting subtypes of precipitation, e.g., *prhki snijeg* 'crisp snow; also *pršić*), *umjetni snijeg* 'artificial snow', *umjetna magla* 'artificial fog (as in a fog machine)'. Some frequencies were attained by manually or semi–automatically filtering results which included erroneous lemmatizations, e.g., *rosa* 'dew' – *Rosa* 'personal name' – *rose* 'rose wine'. Some culturally salient terms may skew the frequencies of some lexemes. For instance, around half of the concordance examples of *mraz* 'frost' occur in the Croatian name of Santa Claus, *Djed Mraz* 'lit. Grandfather Frost'. Nevertheless, corpus data provide some insight into the saliency of particular precipitation phenomena and their dominant place in a folk taxonymy of weather terms.

23   The snow vocabulary is more extensive than listed in the table, but many terms refer to snow on the ground, e.g., *bljuzga* 'slush, sludge', thus different syntactic structures than those denoting the falling of snow are found, e.g., *hodati po bljuzgi* 'walking on slush'. Babić (1982) lists regional varieties of snow, in a type of folk taxonomy, subsuming precipitation such as sleet as type of snow. He also lists terms such as *cijelac* 'virgin snow', *čvrstac* 'rough, thick snow', *sušac* 'dry snow', cf. *suhi snijeg*, *rodinja*k 'springtime snow', *škvrljinjak* 'February snow (Slavonia region)'. Most of these are attested in the corpus only once or are not attested at all.

SLEET	*susnježica* 'rain with snow, sleet'	0.37
	ledena zrna/zrnca 'ice pellets, sleet'	0.01
DRIZZLE	*rosulja*	0.11
	izmaglica 'fog precipitation, drizzle, also mist'	0.96
HAIL[24]	*tuča*	1.28
	grad	0.03
GRAUPEL (SOFT HAIL, SNOW PELLETS)	*solika*	0.06
	krupa	0.15
SMALL HAIL	*sugradica*	0.02
	sutuča	0
DEW	*rosa*	≈ 4
RIME	*inje*	0.36
FROST	*mraz*	8.43
	slana[25]	0
DIAMOND DUST	*ledene iglice*	0.01
CLEAR ICE	*poledica*	0,86
Hydrometeors – other	*magla* 'fog'	17.33
	sumaglica 'mist'	0.33
Non–precipitation (electrometeors and wind)		
LIGHTNING	*munja*	6.04
THUNDER	*grom* 'thunderbolt, thunder'	6
	grmljavina, (grmljava)[26] 'thunder'	2.96
WIND	*vjetar*	49.41
Wind – intensity	*povjetarac* 'breeze'	1.89
	vjetrić.DIM, *vjetrićak*.DIM 'light wind'	0.57, 0
	lahor 'light wind, light air'	1.16
	orkan 'force 12 wind'	0.34
	vihor 'whirl wind'	2.38

Table 1. Weather terms investigated and their frequencies.

---

24   Both terms for hail have extensive homography in the corpus, *tuča, tučnjav*a 'fight', *grad* 'city'. The frequencies presented in the table are conservative estimates after filtering for homographs. However, *tuča* 'hail' is the dominant term, as in random samples of 500 lines for both words, 46% of examples use *tuča* denoting 'hail', and 0% have *grad*.

25   Attested only 2 times as precipitation, homograph with *slan* 'salty'.

26   Attested in 1 example.

As Table 1 shows, the most frequent weather phenomena are rain, snow, fog, wind, thunder and lightning. Based on word formation patterns, some terms can be considered primary (*kiša* 'rain', *snijeg* 'snow', *tuča* 'hail') and some secondary (*su–snjež–ica* 'sleet' < *snijeg* 'snow', *su–magl–ica* 'mist' < *magla* 'fog', *ledena kiša* 'icy + rain'). Rain, snow and wind have the most extensive vocabulary when it comes to lexicalizing intensity, or when it comes to distinguishing subtypes of these phenomena, where the multi–word pattern [Adj] + *precipitation*[27] is used frequently. Wind has an elaborate vocabulary devoted to subtypes (hyponyms).[28] The overall hyperonymy–hyponymy relation (e.g., wind – types of wind) does play a role in frequencies observed with realizing the precipitation with or without an argument in many of the corpus examples, as we will illustrate below.

Turning our attention to weather verbs, a similar procedure was followed, in the sense that reference grammars and studies on impersonal verbs (Barić et. al 1995; Katičić 2002; Silić and Pranjković 2007; Belaj 2007; Belaj and Tanacković Faletar 2015) were consulted as to the list of *verba meteorologica*, yielding a list of 11 weather verbs, though none of the studies deal with an in–depth corpus or lexical semantic analysis of all of these verbs comparatively, most focusing on a few main examples (rain, snow and falling). Additional steps were performed in order to gather as much weather verb candidates for investigation, a) consulting lexicographic sources as to possible verbs containing the same root as a weather phenomenon, e.g., *škropiti* 'to fall in small drops, drizzle' (*škropac* 'light short rain'), *bljuznuti* 'to pour (shortly)' (*bljuzga* 'slush') (Šonje 2000), b) identifying verbs which collocate[29] with weather terms in the corpus, e.g., *rominjati* 'to drizzle', *fijukati* 'to blow, hum' and denote the occurrence of a weather event with additional seman-

---

27 Table 1 contains type frequencies for that pattern.

28 Over 30 terms (e.g., *bura* 'bora', *maestral* 'mistral') listed in the sources.

29 The way to collect and identify verb candidates was based on two steps. One, we looked at Word Sketch collocates of weather phenomena, e.g., *kiša* 'rain', *kišica* 'rain.DIM' and others from Table 1. Secondly, based on the list gathered by such procedure (and lexicographic entries of verbs), we reversed the procedure and looked at collocates of verb lemmas, based on logDice collocation measures provided by the Sketch Engine search platform with the verb as KWIC +/–3. Based on this list we were able to gain an overview and attestation of those weather terms which co–occur with specific verbs. Overall V frequencies are reported for concordance lemma search (V infinitive), default POS any, and they provided a basis for sampling as well. Like some nouns in Table 1, some verbs, notably *pršiti* 'to spray, to seep', *rositi* 'to dew', *piriti* 'to blow lightly' had extensive erroneous lemmatization and POS tagging resulting in homographs such as *Pršo* (surname), *rose* (wine) and *pirit* (crystal), so frequencies in Table 2 are estimations after semi–automatically filtering results and manually checking results. Relative frequencies of the [V+specific precipitation lemma] uses obtained by looking at weather collocates of verbs are reported in column 3 of Table 2. This procedure allowed us to gain plethora of examples from the corpus for our qualitative analysis, as well as the insight into basic tendencies of some verbs patterning with only some precipitation phenomena, and not others. Though this part of the procedure is focused on collecting data mainly from intransitive uses of weather verbs, it is informative as to the predominance of atransitive uses of some verbs (none or few noun collocates attested). For reference, only the verbs *kišiti* 'to rain', *sniježiti* 'to snow', *pljuštati* 'to pour', *lijevati* 'to pour', *padati* 'to fall', *grmiti* 'to thunder', *sijevati, bljeskati* 'to flash' and *puhati* 'to blow' have frequencies in the impersonal form of the adjectival participle >0.01, thus indicating atransitive uses, though this does not mean that contextually other verbs are not acceptable (more below). Furthermore, this procedure was continuously supplemented with qualitative analysis of examples, sampling, acceptability tests in different contexts based on our own observations and so forth. In this sense, this is a corpus–based, and not a corpus–driven approach.

tic features we will discuss below. Step b) was very relevant to address the issue of verb polysemy in gathering usage examples, in the sense that many of the verbs in the list appear frequently with meanings outside of the weather domain, e.g., *struja/melodija/miris/zrak struji* 'electricity/melody/scent/air is flowing' – *vjetar struji* 'wind is flowing (lit. streaming)'. For example, *padati* 'to fall.IPF' is listed in *VRH* (2015) as having the following senses: I. 1. lose balance, 2. a. descend or approach, b. drop, 3. a. descend to the ground, b. cover the ground, c. to set, d. to lean, 4. hang in relaxed way, 5. a. to stop giving resistance, b. to lose life in combat, c. to be dethroned, d. to be conquered, e. reach the end of one's existence, f. to drop, lower [in quantity or value], g. to be the end or consequence of something, h. to be under the influence of sb/smth, 6. to remain on smth, 7. (synsemantic verb), 8. a. to be blamed, b. to enter some state, II. to get a failing grade. Looking at some of the top 20 collocates of *padati* in the corpus[30], it is clear that weather terms are interspersed among other collocates indicating different senses, e.g., *kiša* 'rain', *snijeg* 'snow' (sense 3), *nesvijest* 'unconsciousness'– *padati u nesvijest* 'to become unconscious' (sense 7), *granata* 'grenade' – *granate su padale* 'grenades were falling' (sense 2), *cijena* 'price' – *cijene padaju* 'prices are dropping' (sense 5.f). Thus, only one (or few) senses are attested as denoting the occurrence of a weather phenomenon, and these were the focus of this study, as one of its primary goals is to identify verbs which are used in weather expressions, denote the occurrence of a (non–)precipitation phenomenon and compare the semantic features pertinent to lexicalization of weather events among this set of verbs.[31] This means that we focused on examining weather terms collocates for each verb, and this step provided us with information as to what weather phenomena are attested as co–occurring with certain verbs, and which phenomena are not. This allowed us to obtain comprehensive samples of verb uses in the weather domain, on which we based our qualitative analysis and generalizations. The list of investigated verbs is presented in Table 2.

Verb	overall V frequency	frequency of V + (non–) precipitation
kišiti 'to rain'	0.74	Ø
daždjeti 'to rain, archaic'	0.04	Ø
pljuštati 'to pour.IPF'	1.47	*kiša* 'rain' (0.32); *kišurina* 'rain. AUG' (0.01)
pljusnuti 'to slap; to pour.PF (shortly)'	0.88	*kiša* 'rain' (0.04); *pljusak* 'rain-shower' (<0.01)
sipiti 'to seep'	0.88	*kiša* 'rain' (0.09), *kišica* 'rain. DIM' (0.06), *snijeg* 'snow' (0.03)

---

30   logDice measures are 10.63, 9.47, 7.60, 7.40 and 6.82, respectively.
31   Comparing meteorological verb senses to other senses would expand the limits of this study significantly, so we reserve this topic for future investigations.

rominjati 'to drizzle with a hum, to roam'	0.21	*kiša* 'rain' (0.08), *kišica* 'rain.DIM' (0.05)
lijevati, liti 'to pour (of rain, but also liquids)'	3.63	*kiša* 'rain' (0.14)
škropiti 'to sprinkle'	0.23	*kiša* 'rain' (0.01); *kišica* 'rain.DIM' (<0.01)
bljuznuti 'to pour; to spit'	0.01	n/a
sniježiti 'to snow'	0.23	Ø
pršiti 'to spray, to seep'	0.21	*snijeg* 'snow' (0.02); *snježak* 'snow.DIM' (<0.01)
lepršati 'to flutter, to flap'	0.8	*snijeg* 'snow' (0.01) (cf. *pahulja* 'snowflake' (0.02))
rositi 'to drizzle, to dew (become dewy/wet with drops)'	0.7	*kiša* 'rain', *kišica* 'rain.DIM' (<0.01)
padati.IPF 'to fall'	92.68	*kiša* 'rain' (7.28), *kišica* 'rain.DIM' (0.13), *kišurina* 'rain.AUG' (0.02), *snijeg* 'snow' (3.08); *tuča* 'hail' (0.15) // *susnježica* 'sleet', *pljusak* 'rainshower', *rosa* 'dew', *oborina* 'precipitation' (≈ 0.05) // *rosulja* 'drizzle', *solika* 'graupel' (0.01) // *magla* 'fog' (<0.01) // *sumaglica* 'mist', *mraz* 'frost', *inje* 'rime', *poledica* 'clear ice' (0)
pasti.PF 'to fall'	193.32	*kiša* 'rain' (3.04), *snijeg* 'snow' (2.38) // *tuča* 'hail' (0.04), *pljusak* 'rain shower' (0.07), *rosa* 'dew' (0.03) // *magla* 'fog' (0.01) // *susnježica* 'sleet', *solika* 'graupel', *mraz* 'frost', *inje* 'rime' (<0.01) // *poledica* 'clear ice' (0)
ispadati se 'to fall out (completely)'	0.11	*kiša* 'rain' (0.02)
zapadati 'to fall (behind something)'	1.56	*snijeg* 'snow' (0.04)
napadati (se) 'to fall (sative)'	34.82	*snijeg* 'snow' (0.71), *kiša* 'rain' (0.04)
paduckati.DIM 'to fall lightly'	0.03	*kišica* 'rain.DIM' (0.01), *kiša* 'rain', *snijeg* 'snow' (<0.01)
grmiti/grmjeti 'to thunder'	1.91/0.29	*grom* 'thunder' (<0.01)/(Ø)

tutnjiti 'to rumble'	0.41	*grmljavina* 'thunder (collective)' (0.01), *grom* 'thunder' (<0.01) (0.02)
sijevati 'to flash'	1.04	*munja* 'lightning' (0.14)
bljeskati 'to flash'	1.02	*munja* 'lightning' (0.01)
puhati 'to blow'	12.36	*vjetar* 'wind' (3.06)
bridjeti/briditi 'to sear with burning sensation'	0.29/0.02	*vjetar* 'wind' (<0.01)
piriti 'to blow (lightly)'	<0.26	*vjetar* 'wind' (0.05), *vjetrić* 'wind.DIM' (0.02), *povjetarac* 'breeze' (0.01), *lahor* 'light breeze' (<0.01)
pirkati.DIM 'to blow lightly, to breeze'	0.04	*vjetar* 'wind', *vjetrić* 'wind.DIM' (0.01)
brijati 'to sear with cold'	12.17	*vjetar* 'wind' (0.04), *bura* 'bora' (0.03)
šumiti 'to hum'	0.84	*vjetar* 'wind' (0.03)
fijukati 'to whistle'	0.22	*vjetar* 'wind' (0.04), *bura* 'bora' (0.02)
zavijati 'to howl'	1.83	*vjetar* 'wind' (0.05), *bura* 'bora' (0.01), *mećava* 'blizzard' (<0.01)
viti (se), vijati 'to spin, whirl'	<1	*vjetar* 'wind' (0.01)
kovitlati se 'to whirl'	0.32	*vjetar* 'wind' (<0.01)
strujiti 'to stream'	<0.5	*vjetar* 'wind' (<0.01)

Table 2. Verbs attested as occurring in weather expressions related to precipitation and non–precipitation events. Overall V frequency refers to the relative frequency of a V lemma in the corpus, while the V with (non–)precipitation refers to the relative frequency of a V + a particular precipitation lemma (+/– 3 KWIC), e.g., *padati+kiša* 'fall + rain'; some with very low frequencies compared to others. Ø refers to verbs which appear in the atransitive constructions, so the co–occurrence with precipitation noun lemmas is not applicable or negligible (≈0).

Reviewing examples from overall and weather domain uses of verbs indicates that the verbs investigated are polysemous[32], i.e., used within and outside the weather domain, as confirmed by their lexicographic entries as well. For example,

---

32 When we talk about polysemy, we use it in the sense of lexical polysemy proper, semantic extensions of verb meanings based on metonymy, metaphor, generalization and specialization, not syntactic polysemy (e.g., atransitive vs. intransitive uses of the verb with the same atmospheric meaning). On a constructional account both types of polysemy could be systematically examined, but this is not the focus of the present paper.

pljusnuti is used much more often in the corpus in the meaning of 'slap' than 'pour (shortly)', *lijevati* has a common sense of casting molten metals and its strongest collocate is iron. What the comparison of corpus and lexicographic data also shows is that a) some weather verbs are productive metaphorically, as verbs with their basic lexicographic meanings listed as relating to weather appear in other contexts, e.g., *munja je sijevnula* 'lightning flashed', *šaka je sijevnula* 'a fist flashed', *kontra je sijevnula* 'a counterattack flashed (sport)', and atmospheric source meanings need not be the most frequent in the corpus, b) the category of weather predicates is enriched by verbs whose basic meanings may not be listed as weather–related, but their use in weather contexts contains semantic features which profile a specific aspect of one weather event in terms of manner of occurrence, some in borderline uses, e.g., *lepršati* 'to flutter', *strujiti* 'to stream'.

Based on our investigation of these lexical units and syntactic structures, in the following sections we will point out several main observations:

a) Within the category of precipitation, a division can be made along the lines of dynamic and static precipitation based on the syntactic structures used.

b) *Padati/pasti* 'to fall' has the role of a generalized precipitation verb, though its role differs with respect to dynamic and static precipitation.

c) When it comes to verbs, they lexicalize different aspects of weather event[33] depending on the (non–)precipitation type (cf. Meulleman and Paykin 2016), such as: i) Figure (*kišiti* 'to rain'), ii) Path (*padati* 'to fall') and iii) Manner (*rominjati* 'drizzle lightly with a hum'). The latter can be divided further based on semantic features such as sound, intensity or manner of motion serving as sources for the construal of manner features. Quantity of precipitation is coded morphosyntactically by two means – the partitive genitive construction and prefixation.

d) These lexicalization patterns play a role in the prevalence of the predicate or the argument–predicate type for some of the verbs, along with factors such as contextual grounding, ellipsis and hyponymy–hyperonymy relations of weather terms. In particular, the split type can be regarded as the basis for the productive use of Croatian weather verb vocabulary with respect to manner, as it divides the semantic features of a weather phenomenon across subject and verb, e.g., intensity, *kišica sipi* 'rain.DIM (lightly) seeps', *vjetrić pirka* 'wind.DIM (lightly) breezes'.

4.1 Dynamic and static phenomena

When it comes to divisions of activities and states, weather verbs (*thunder*, *rain*, *snow*) have been described as activities in terms of their aspectual and actionality

---

33   As used by Talmy (2000), cf. Meulleman and Paykin (2016).

properties (e.g., Dowty 1979; Van Valin and LaPolla 1997; cf. Vendler 1957).[34] This notion of activity is tied to the notion of dynamicity as exhibited by precipitation and non–precipitation phenomena. Eriksen et al. (2012: 9) divide dynamic and static weather events based on whether a given event involves a perceptible on–going or momentary activity or not. In their division of dynamic and static events precipitation and non–precipitation events are dynamic, while temperature conditions, for example, are static. Looking at the prototypical Croatian examples (rain, snow, hail), this division holds, though not across all precipitation phenomena. We therefore adjust this notion for the analysis of Croatian data by distinguishing dynamic and static precipitation, but the division may be different in other languages.[35]

The differences in types of precipitation have to do with some precipitation perceived as having a downward motion (rain, snow), and some not having a clear point of origin (fog, rime). Therefore verbs of being (basic locative construction with the *be* copula, *nastajati* 'to become', appearing (*stvarati se* 'to appear, to be made') or covering (*prekriti*.PF/*prekrivati*.IPF 'to cover') are quite common with precipitation phenomena without a clear point of origin (fog can both rise or come down, or just appear), spatial extension (covering) or static properties (rime, dew or frost usually just lie on the ground / shrubbery) (cf. Meulleman and Paykin 2016). They are profiled predominantly as static events, or events without a clear starting and end point. For example, the acceptability of sentences with end–phase meanings such as *Kiša je stala/*Kiša se razišla* 'the rain stopped'/*the rain dissipated' and **Magla je stala/Magla se razišla* '*the fog stopped/the fog dissipated' or motion verbs, *diže se magla / magla se spustila na grad* 'fog is rising / fog descended upon the city' and **kiša se dignula* 'the rain rose up (unacceptable)' / *kiša se spustila* 'the rain descended' point to different spatio–temporal properties of these weather events. Precipitation phenomena most commonly profiled as static are fog, frost, dew, clear ice and rime, as in 4).

4) *Mraz / inje je na tlu i granama drveća.*
   'Frost / rime is on the ground and tree branches.'

Fog is usually denoted as a static event in primary expressions, either in the argument or predicate type, e.g., *na vrhu je bilo hladno, maglovito, sa snijegom* 'it was cold, foggy, with snow at the top', *jutro je maglovito* 'the morning is foggy' (adjectival predicate), *maglovito jutro* 'a foggy morning', adjective 'foggy', argument type with copula, *magla je, bura i dalje puše* 'it is foggy, bora is still blowing', *magla je na stanici* 'fog is at

---

34   In comparison to English, Croatian has a complex lexical aspect system realized via suffixes or prefixes, e.g., *grmiti* 'to thunder.IPF' – *zagrmiti* 'to thunder.PF (once)', *sijevati* 'to flash.IPF' – *sijevnuti* 'to flash.PF (once)'. Oftentimes it not only a perfective – imperfective distinction that which is lexicalized by affixation, but other features as well, e.g. *na–padati se* 'to fall.PF (sative)'. This makes derivation a relevant aspect of lexicalization descriptions in Croatian.

35   Diachronically, this division can change for Croatian as well, in some examples. As Gluhak (2004) points out in his etymological study on dew and rain in Croatian, the notion of dew falling from the sky was widespread previously.

the station', while secondary expressions typically refer to the covering properties of fog, e.g. *magla je obavila / prekrila grad* 'the fog engulfed / covered the city'.[36]

As Table 2 shows, the use of *pasti*.PF/*padati*.IPF is widespread with different precipitation types, though it is not attested with all of them, and frequencies fall drastically between dynamic and static precipitation, especially in the imperfective *padati*.[37] We are inclined to posit, from the perspective of the use of *pasti*.PF/*padati*.IPF as a generalized precipitation verb, two verb senses being used with precipitation. In the first sense, with dynamic precipitation, the verb retains downward motion and activity as the meaning (e.g., *pada kiša* 'rain is falling'), while in the other sense it denotes an appearance of the phenomenon.[38] The latter requires an argument, e.g., *pada noć* 'night is falling' / **pada*. Thus, when we talk about motion as being coded by *pasti*.PF/*padati*.IPF, we talk about one of the senses with dynamic precipitation uses, the other sense showing generalization or bleaching as the process behind the semantic shift[39], generalization occurring within and outside the weather domain, e.g., *pala je magla* 'fog fell', *pala je noć* 'night fell', compare *pale su uvrede* 'insults fell (there were insults made)'.

4.2 Weather verbs and lexicalization of weather events

When it comes to the question of what is lexicalized by weather verbs in Croatian, a study by Meulleman and Paykin (2016) critically examines the approaches put forth by Talmy (1985, 2000), Jackendoff (1985) and others as to the main feature lexicalized being Figure (rain – to rain), Path (rain – downward motion) or something else. In their analysis they investigate whether what is being lexicalized is in fact Manner and put forth some convincing arguments as to why Manner should be considered an important factor in weather event lexicalization.[40] One aspect they examine are metaphorical meanings of verbs, e.g., *he stormed out of the room*, the other manner distribution across V–framed and S–framed languages for weather verbs. As the authors conclude, weather verbs differ from regular motion verbs, and they present challenges when one wants to apply a uniform description of their conceptual structure, though metaphorical meanings do show regularities.

---

36 Some verbal derivatives of *magla* 'fog' analogous to the verbal predicate type are attested, but they show a semantic shift, thus not referring to the weather event itself, e.g., **Grad se magli / Prozor se magli* '*The city is fogging up' / 'the window is fogging up' < *magliti se, zamagljivati se* 'to be fogging up'. A few instances (15 for fog, 7 for dew and frost) co–occur with the verb *pasti* 'to fall' and denote the appearance of the precipitation, e.g., *pao je prvi mraz* 'the first frost fell'. This speaks to the process of the verb becoming a generalized weather predicate, a verb of happening (compare *pala je noć* 'night fell'), though still retaining a motion component in most of the examples.

37 Croatian, like other Slavic languages, has a prolific lexical aspect system morphologically coded on verbs via prefixes and suffixes (cf. Šojat 2008; Polančec 2020), so this also ties into notions of activities and states.

38 Compare *kiša je padala pola sata* 'it rained for half an hour', **mraz je padao pola sata* 'frost was falling for half an hour', and *danas je pao prvi mraz* 'today the first frost fell (appeared)'.

39 Cf. Raffaelli and Kerovec (2008).

40 Croatian is generally an S–framed language.

What we find as a relevant insight in their analysis is the way manner can be construed in terms of weather event lexicalization. Croatian verbs which specify intensity or sound properties (see Table 2), and can be viewed as hyponyms to the main weather verbs (e.g., *kišiti* 'to rain' > *rominjati* 'to drizzle with a hum') illustrate this question, as manner properties serve as sources for the weather verb vocabulary extension (e.g., *kovitlati se* 'to whirl', *bridjeti* 'to sear the skin with a burning sensation, as in cold or heal; to blow in such a manner'). We will describe these in detail in the following sections. Also, drawing on the predicate or argument–predicate type distinction as described above, they are more common in the split type in the corpus, and we stated above, and the split type could be the productive basis for their lexicalization. Their predicate type uses rely, in some cases heavily, on the analogy to canonical expressions for rain, snow and wind, and contextual grounding of the utterance itself, e.g., *vani škropi* 'it is sprinkling outside', *jučer je ?škropilo*. SG.N/*kišilo* 'it was ?sprinkling[41]/raining yesterday'. Those that seem more felicitous in the impersonal form are a) verbs with a cognate subject, *pljuštati* 'to pour' < *pljusak* 'downpour' and b) verbs specialized for use (with their basic meaning) in the weather domain, e.g., *rominjati* 'to drizzle with a hum'. Metonymy–based[42] verb senses attested as borderline weather predicates (in terms of frequency) were attested only in the argument–predicate type, e.g., *snijeg je lepršao (/ ?lepršalo je) oko odmarališta* 'snow was fluttering (/?it was fluttering) around the resort'.

When it comes to the two main precipitation verbs, *kišiti* 'to rain' and *sniježiti* 'to snow', they are not productive metaphorically and are specialized to refer only to rain and snow, respectively, unlike, e.g., English examples, cf. *it was raining bullets – *kišili su metci*, but some of the other (manner–coding) verbs are, e.g., *metci su pljuštali* 'bullets were pouring'. Since Croatian avoids cognate subjects in atmospheric meanings only when there is morphological redundancy present, otherwise uses predicate and argument–predicate coding quite frequently and interchangeably may indicate these verbs lexicalize substance as figure or emission of substance (cf. Levin and Krejci 2019) more so than manner. This is in line with Belaj's (2007) analysis as to the level of participant awareness playing an important role in the personal–impersonal distinction for Croatian impersonal verbs.

Another important factor to take into account is that Croatian, unlike e.g., English, can be said to have two competing basic weather expressions for rain and snow, as reflected by the corpus data and everyday use. That is, while in English examples *it's raining* and *rain is falling* one can identify the latter as a marked expression, corpus data and use do not seem to support such a strong distinction for Croatian, where *padati*.IPF/*pasti*.PF 'to fall', lexicalizing involuntary motion along

---

41 Conceivable when contextually grounded with a question, e.g., Was it raining heavily yesterday? – No, it sprinkled a bit.

42 Metonymy is the basis for some of the verbs coding manner of motion attested, as snowflakes, air and dust are the things that flutter, stream or whirl. For predicational metonymy see, e.g., Brdar–Szabó and Brdar (2003).

a downward Path[43], is the generalized pattern across dynamic precipitation phenomena and is used in motion constructions, e.g., *kiša mi je padala po glavi* 'rain fell on my head'/*?kišilo mi je po glavi* 'it rained on my head'. In the most stripped down forms, the two syntactic patterns can be used as synonymous and competing patterns, *kišiti* and *padati*.IPF/*pasti*.PF +*kiša* 'to rain', *sniježiti* and *padati*.IPF/*pasti*.PF +*snijeg* 'to snow', and both can be viewed as "basic" ways to denote rain or snow, one being syntactically simpler and the other being more frequent and used across precipitation phenomena (cf. Kienpointner 2016, section 2 above), both not adding an additional argument apart from precipitation.[44] *Fall+precipitation* pattern is more frequent[45] for both snow and rain, compare *kišiti* (0,74) – *padati*.IPF/*pasti*.PF +*kiša* (1,91) 'to rain', *sniježiti* (0,23) and *padati*.IPF/*pasti*.PF + *snijeg* (1,84) 'to snow'.

When used with less frequent types of precipitation (e.g., *rosulja* 'drizzle', *susnježica* 'sleet') *padati*.IPF/*pasti*.PF + *precipitation* is the main way of denoting the occurrence of a phenomenon attested in the corpus, compare *Kiši. / Pada kiša.* 'it is raining / rain is falling', *Sniježi. / Pada snijeg* 'it is snowing / snow is falling', **Tuči. / Pada tuča.* 'Hail is falling' **Rosulji. / Pada rosulja* 'Drizzle is falling', **Susnježi / Pada susnježica.* 'Sleet is falling', but also **kišila je tuča/rosulja/susnježica* 'hail/drizzle/sleet was falling/*raining'.[46]

While the more analytical nature of this pattern provides speakers with a somewhat different function, e.g., malleability as to modification[47] (or, for example, introducing rain as the topic) and thus may account for its higher frequency, we do find this to be a competing pattern with *kišiti* and *sniježiti*, i.e., we do not find it secondary as to the typology of weather expressions, especially since *pada*.3SG can be used with an ellipsis of the argument (rain or snow), most commonly in spoken language. It is acceptable and conceivable that one can say *baš jako pada* 'it is falling intensely' and think of, for example, hail or sleet, but this would require contextual grounding of that specific precipitation type for the example to attain that inter-

---

43   Compare the derivative *padalina* 'precipitation', lit. 'that which falls'.

44   Its use in Croatian meteorological expressions differs from data reported for some of the Slavic languages. For example, Russian uses the argument type with the generic motion verb 'go' + precipitation (*идёт дождь* 'it's raining') while Polish has the specialized form *pašć* 'fall–precipitation' reserved for the weather domain, and uses the form *spašć* 'fall' for other types of 'falling' (cf. Andrason 2019; Croatian *spasti* lexicalizes a transition from a higher to lower position and is often used metaphorically as 'decline; degrade').

45   Frequency is reported taking into account the relatively free word order in Croatian, not only the string [V+precipitation]. The Word Sketch result is presented above, a more conservative estimate, while the co–occurrences attained by looking at collocations measures as described above present the pattern as much more frequent, *padati* +*kiša* – 7,26, *pasti*+*kiša* – 2,35 / *padati* +*snijeg* – 3,08, *pasti*+*snijeg* – 1,9.

46   This could be in line with the observation that Figure is not a productive lexicalization pattern with precipitation, but Manner and Path are, though this is a tentative observation at best. Archaic terms, *daždjeti* 'to rain' and *rositi* 'to dew, rain', could be offered as counterexamples, though they are not used today in any substantial amount.

47   Cf. Meulleman and Paykin (2016). Also, compare common corpus examples, *pada* 'fall' + *obilna kiša* 'abundant rain' (quantity), *jaka/slaba/umjerena kiša* 'heavy/moderate/light rain' (intensity), *mjestimična kiša* 'sporadic rain; lit. in some places' (distribution/location), *povremena/stalna/jesenska/ljetna/dugotrajna kiša* 'occasional/continous/autumn/summer/long lasting rain' (temporal properties), *sitna/gusta/olujna kiša* 'tiny/thick/stormy rain' (manner), *dosadna/naporna/neumoljiva kiša* 'boring/tedious/unrelenting rain' (experiencer–based properties).

pretation, in the sense that it is experientially observed in the context of the utterance or introduced earlier in the discourse. That would indicate that the verb itself is ambiguous as to the type of precipitation occurring, and that the distribution of semantic information is typically distributed across the argument and the predicate, with an ellipsis of the argument motivated by the context of the utterance.

Moreover, quantity of precipitation is coded morphosyntactically via prefixation of *padati*.IPF/*pasti*.PF, *is–padati se* 'to rain sufficiently, completely', cf. **iskišiti se*, *napadati* 'fall in larger quantities and stay on the ground' and *zapasti/zapadati* 'fall sufficiently, covering the surroundings', cf. *zasniježiti* 'to start snowing'. A special construction is the impersonal form of the verbs *(na/za)padati* and the precipitation noun in the genitive case (cf. Menac 1986)[48], with an impersonal form of the verb and the participant in the genitive position, e.g., *palo je 25 cm snijega* '25 cm of snow fell'/**sniježilo je 25 cm snijega*. This genitive construction is typical of Motion – Path encoding, e.g., *palo mi je 2 metra snijega na auto* '2 meters of snow fell on my car', cf. *ušlo mi je 30 litara vode u stan* 'lit. 30 liters of water entered my apartment'.

Like *pasti/padati* 'to fall', the verb *puhati* 'to blow' alternates between the atransitive and intrasitive uses, e.g., *od jutra jako puše* 'it is blowing (it is windy) since morning' and *od jutra jako puše vjetar* 'wind is blowing since morning', and this is also dependent on the context of the utterance. Unlike *pasti/padati* 'to fall', its basic meaning is related to air currents, i.e., the weather domain[49], though numerous uses of *puhati+vjetar* in the corpus indicate that the argument is not redundant, as is the case with morphological redundancy. With *puhati*, hyponyms denoting different, culturally salient types of wind typically co–occur to specify the type of wind in question, compare the list of weather collocates with *puhati* in Table 3 below. This may point to a way that lexical semantic relations of weather phenomena (e.g., hyperonymy/hyponymy) influence tendencies in syntactic patterning.

Verb	Subject
puhati 'to blow'	*vjetar* 'wind' (hyperonym)
	hyponyms: *bura* 'bora, a strong northern Adriatic wind', *jugo* 'jugo, a south–east warm wind', *maestral* 'mistral, sea breeze', *povjetarac* 'breeze', *vjetrić* 'wind.DIM' *sjeverac* 'a cold northern wind' (compare *sjever* 'North'), *burin* 'a night breeze', *tramontana* 'tramontana'

Table 3. List of the top wind–hyponym collocates with *puhati* 'to blow'.

---

48   This is not a construction subtype commonly attested for other predicate–type verbs, e.g., *?kišilo je puno kiše* 'it rained a lot of rain', *?puhalo je puno vjetra* 'it blew a lot of wind'. As for the classification of such a genitive use among the plethora of genitive uses in Slavic (cf. Silić and Pranjković 2007), and in line with Menac (1986), we would view it as a partitive genitive, as quantity is often indicated, e.g., *malo* 'little', *puno* 'a lot'.

49   And onomatopoeically relates to the domain of sound (Gluhak 1993). Another meaning adds a causative component, 'to create an air stream by blowing', e.g., *puhati balone* 'to blow balloons', or idiomatic meanings, e.g., *puhati na hladno* 'lit. to blow into cold; to have an excessive reaction to something that hasn't happened'. Its perfective variants are *puhnuti* 'to blow (lightly and shortly)' and the inchoative *zapuhati* 'to start blowing'.

4.3 Weather predicates – specifying the manner of a weather event

As we stated in the Introduction, one of the goals of our investigation was to expand the data on Croatian weather verbs and examine different tendencies in the lexicalization of weather events. Therefore, we extracted verbs which denote weather events, are not among the commonly discussed *verba meterologica*, but further specify some property of the weather event in question, e.g., *rominjati* 'to drizzle with a slight hum'. These verbs lend support for the Manner lexicalization of weather events, and they are often intricately connected with only one type of weather phenomenon. This manner component can be exploited in motion constructions proper (e.g., *pljusnula je na pločnik* 'she splashed onto the pavement'), though in this section we will focus on analyzing potential regularities of sources of manner construal, and regard manner as manner of occurrence of a weather phenomenon based on some intrinsic property.

When it comes to **rain**, two of the verbs, *rominjati* 'drizzle with a slight hum' and *sipiti* 'lightly and silently fall in small droplets' denote rainfall of small intensity, and two denote rainfall of great intensity, *pljuštati* 'to pour.IPF' and *lijevati* 'to pour. IPF'. *Pljuštati* 'to pour' seems to be more specialized for rain, as its primary meaning is in the weather domain, and its nominal counterpart is *pljusak* 'rainshower, downpour'. *Pljuštati* is the one verb that lexicalizes the participant clearly in the form of the verb, and is common in the atransitive constructions as well, e.g., *vani pljušti* 'it is pouring outside', compare *kiša pljušti* 'rain is pouring', **pljusak pljušti* 'downpour is pouring'. The perfective from *pljusnuti* is also attested in some examples in the corpus, but the verb also has the meaning 'to hit; slap someone' and is not frequently used in the meteorological meaning. All forms utilizing the root *pljus–* are described in the Croatian Language Portal as onomatopoeically motivated by the sound of splashing on the surface of water. The primary meaning of *lijevati* 'to pour', on the other hand, is related to the motion of liquids in a stream, compare *rijeke se slijevaju u more* 'rivers flow into the sea'.

As for **snow**, there are a few predicates denoting the manner of snowing, *sipiti* 'to seep, drizzle', *pršiti* 'fall slowly in small particles (snowflakes)' (compare *pršić* 'a dry granular snow') and *lepršati* 'flutter' (a borderline snow predicate attested in the corpus, as it denotes a type of motion typical of wind, e.g., flutter of the flag, wings of butterflies, flutter of snowflakes).

5)

a) *Vani sipi   sitan   snijeg.*
   outside drizzle.3SG tiny.NOM snow.NOM
   'A tiny (granular) snow is drizzling outside'

b) *Po tom gradu neprestance prši   bijeli topli snijeg.*
   on this city.LOC constantly fall–in–small particles.3SG white warm snow.NOM
   'White warm snow is constantly falling over that town'

c) *U gorskim područjima leprša snijeg.*
    in hillside area.LOC.PL flutter.3SG snow.NOM
    'Snow is fluttering in hillside areas'

Wind seems to be the most prolific weather phenomenon when it comes to these types of verbs. Such verbs appear with the argument present (*vjetar* 'wind', *povjetarac* 'breeze' etc.) and they denote various aspects of wind. The first is intensity, with verbs such as *pirkati/piriti* 'to breeze' denoting wind of weak intensity, and verbs such as *bridjeti* 'to blow with intensity, searing the skin with cold' and *brijati* 'to blow strongly (of cold wind)' combining the cold feeling of the wind with intensity. Second is manner of motion, with random rotational movement of the wind being profiled by *kovitlati se* 'to whirl', *viti* 'to flutter, whirl' and linear movement with *strujiti* 'to stream (of air movement)'. These verbs are more commonly used with the wind being the cause or setting of the movement of other entities, e.g. *vjetar kovitla šareno lišće* 'wind is whirling the colorful leaves', *pahuljice se kovitlaju na vjetru* 'snowflakes are whirling in the wind' rather than denoting motion of the wind proper, e.g., *vjetar se već kovitla u daljini* 'the wind is already whirling in the distance', therefore it is a borderline use of the verbs denoting motion of the wind. A third aspect commonly denoted is motivated by the sound of the wind, with verbs such as *šumiti* 'to hum, whirr', *hujiti* 'to blow, hum', *fijukati* 'to whistle (of wind)[50], *zavijati* 'to howl' typical of these uses, as in 12).

6) *Vjetar puše / šumi / fijuče / zavija kroz guste hrastove krošnje.*
    'The wind is blowing / humming / whirring / howling through the thick oak
     tree tops'

Sound is etymologically connected to the meaning of **thunder** as well (*grmjeti*[51] 'to thunder', cf. Gr. *khrómos, khrómē* 'a hum, storm', Gluhak 1993), with the other common verb explicitly denoting a rapid intense sound, *tutnjiti* 'to boom, to rumble, to thunder' (compare *tutanj* 'a booming sound, as of thunder or cannons), *gromovi tutnje* 'thunders are booming', *oluja je protutnjala gradom* 'the storm thundered through the town'. Apart from the noun *grom*, denoting a singular event, a collective meaning is present in the derivative *grmljavina* 'a series of thunders', most commonly used with *tutnjiti* 'to boom, to thunder', or verbs such as *odjekivati* 'to echo', but not common with *grmjeti* 'to thunder'.

As for **lighting**, there are two synonymous verbs connected to it, *sijevati*.IPF/ *sijevnuti*.PF 'to be flashing / to flash' and *bljeskati*.IPF/*bljesnuti*.PF 'to be flashing / to flash'. *Sijevati* 'to flash' is the primary verb most frequently used with lightning, denoting a short and intense burst of light, while *bljeskati* 'to flash' primarily denotes

---

50 Whistling made by humans is described as *fućkati* 'to whistle'.

51 Or the alternate form *grmiti*. The derived form *zagrmiti* 'to start thudering; to thunder once' has the inchoative meaning (similarly as *zapuhati* 'to start blowing') and can be used in the atransitive form, *zagrmilo je vani* 'it thundered outside'. *Pregrmjeti* 'lit. to thunder over smth; to weather smth' is used only metaphorically.

reflection of light and is more commonly used with other entities that emit light, e.g., *bljeskaju božične lampice* 'Christmas lights are flashing/twinkling'. Primacy of *sijevati* 'to flash' over *bljeskati* 'to flash' is further confirmed in the frequent common expression *grmi i sijeva* 'it is thundering and flashing', but not ?*grmi i bljeska* 'it is thundering and flashing'.[52] However, the nominal construction *bljesak munje* 'a flash of lighting' points to clear connection between the synonym *bljeskati* and the concept of lightning.[53]

We summarize our observations regarding the lexicalization of manner features in Table 4. Croatian data show that manner as a semantic feature is complex and lends itself to potential future comparisons across languages as to the different or similar sources of the manner features, with regards to investigating cross–linguistic commonalities in conceptualizations of weather events, as shown below.

sound	*rominjati* 'drizzle with a slight hum' *pljuštati* 'to pour.Imperf'	RAIN
	zavijati 'to howl' *šumiti* 'to hum, whirr', *fijukati* 'to whistle (of wind)'	WIND
	tutnjiti 'to boom, to thunder'	THUNDER
motion	*sipiti* 'lightly and silently fall in small droplets' *lijevati* 'to pour.Imperf (intensely flow)'	RAIN
	sipiti 'lightly and silently fall in small granules' *pršiti* 'fall slowly in small particles (snowflakes)' *lepršati* 'flutter' (borderline)	SNOW
	viti 'to flutter, whirl' *strujiti* 'to stream (of air movement)' *kovitlati (se)* 'to whirl' (borderline)	WIND
short and intense emission of light	*bljeskati/bljesnuti* 'to flash'	LIGHTNING
experiencer–based properties (tactile temperature)	*brijati* 'to blow strongly (of cold wind)', *bridjeti* 'to blow with intensity, searing the skin with cold'	WIND

Table 4. Prominent features of weather events as lexicalized by verbs denoting manner.

---

52   50 occurrences in the corpus versus 1.
53   A nominal form of *sijevati, sijev* 'a flash' is reported in the lexicographic data but is not common in use and is attested only 3 times in hrWaC.

5. Secondary weather expressions and complex weather events

In the end we will briefly mention two kinds of secondary weather expressions. One has to do with expressions that, from the point of view of the typology of weather expressions, are not the primary expressions used to denote the occurrence of a weather event but may be typical ways to describe some aspect of its occurrence, e.g., *počela je kiša* 'rain began'. To make it clear, if this were the only way to say 'it's raining' in Croatian then we would treat this as a primary, and not a secondary weather expression. As it stands, we introduce this term to point out those expressions that "build upon" the weather event in various ways and are common in corpus data. Motion and phase verbs, *doći* 'come', *proći* 'pass', *početi* 'begin', *(pre)stati* 'stop', are typically used with dynamic precipitation phenomena, e.g., *dolazi kiša / prošao je snijeg* 'rain is coming / the snow passed', and they are all of the argument type. As for complex weather phenomena, they are, on the other hand, the primary way of denoting their occurrence, e.g., motion verbs, *dolazi / približava se oluja* 'a storm is coming/approaching', *oluja nas je zaobišla* 'the storm passed us by', alongside existential constructions, e.g., *vani je nevrijeme / prolom oblaka* 'there is a storm / cloudburst outside'.

The other type of secondary weather expressions has to do with specialized weather meanings of some of the verbs related to a weather phenomenon. They also build upon weather events in various ways. We will illustrate this with the example of snow. Secondary weather expressions show a vocabulary with highly specialized meanings related to properties of snow, either in intransitive or transitive constructions, e.g., *zabijeliti* 'to whiten', *bijeliti se* 'to be white', *okopnjeti* 'to thaw (only of ice and snow)', *prtiti (snijeg)* 'to make a path in the snow by walking or shoveling', *zamesti* 'cover up (of snow); snow in' (compare *zamet* 'snowdrift made by wind')[54]. Such expressions, in other words, lexicalize particulars aspects of weather phenomena and represent our knowledge of them, so they offer future avenues for systematic investigations. A complex weather event, *mećava* 'blizzard, strong wind and snow' is attested with one of these secondary weather expressions, e.g., *zamela nas je mećava* 'a blizzard snowed us in', but usually appears in the argument type with non–weather related predicates, e.g., *počela je mećava* 'a blizzard started', *vani je mećava* 'there is a blizzard outside'.

Finally, simultaneous (mixed) weather events are most commonly coded by coordination, e.g., *grmi i sijeva* 'it is thundering and flashing', *pada kiša i puše vjetar* 'rain is falling and wind is blowing', or by the preposition *s(a)* 'with', and there is some variation in the corpus when coding multiple simultaneous weather events, e.g., *kiši s tučom – pada kiša s tučom – padaju kiša i tuča* 'it is raining with hail' – 'rain with hail is falling' – 'rain and hail are falling'. Systematic examination of variations of such encodings, however, lies outside of the scope of this study.

---

54 A lot of these have metaphorical uses outside of the weather domain, e.g. *(o)kopniti* 'to weaken, to lose weight', *zamesti svaki trag* 'to cover up any trace of smth', but with the exception of *(za)bijeliti (se)* their basic meanings are glossed as related to snow.

Conclusion

The goal of this study was, in some ways, twofold. On the one hand, the goal was to describe weather expressions, particularly weather verbs, with respect to lexicalization of weather events in Croatian. As the data show, weather verbs and weather expressions form a more complex category than what is commonly used to refer to *verba meteorologica* in Croatian, i.e., impersonal forms of verbs denoting weather events. Moreover, the purpose of this study was to place Croatian data in the context of proposed typologies of weather expressions, with data showing a dynamic relationship between the predicate type and the argument–predicate type in use. This implies that the strategies speakers use when describing a weather event are more nuanced than typically reported, and that the (verbal) predicate type is not the most dominant one across all weather phenomena, pertaining instead to only a few prominent examples. Though aimed at cross–linguistic comparisons, our analysis shows that the proposed tripartite typology of weather expressions is useful when applied to the description of language–specific data, as it can provide some systematization of Croatian weather data variation, the prototypical examples of syntactic types and different functions they perform, as well as provide some insights into the bases of lexicalization features of Croatian weather verbs. A corpus–based overview showed that, apart from the most frequent examples from extant descriptions denoting the process of rain, snow, wind etc. (*kišiti* 'to rain', *sniježiti* 'to snow', *puhati* 'to blow', there are many other verbs used to denote particular aspects of a weather phenomenon, some showing regularities across weather phenomena, e.g., lexicalizing sound properties of the event. An analysis of such manner features of weather events opens new avenues for linguistic comparisons of their sources. Some examples of lower frequency verbs used as borderline cases may indicate how particular verbs are incorporated into weather event structures in future systematic investigations. Other types of expressions may not be seen as equally typologically relevant, but are still, in our opinion, relevant in the semantic description of the weather domain. For this reason, we introduced the notion of a secondary weather expression, to include examples which show a specialized weather lexicon. This goes to show that there are salient properties of weather events which can be used to further investigate cross–linguistic regularities in the lexicalization of weather events, but also that lexicalization and syntactic patterns can vary intralinguistically and offer different ways to describe a weather phenomenon. A systematic study of the weather domain was a goal in and of itself. Its purpose was to sketch out future guidelines in the study of weather in Croatian, and by introducing a distinction between weather expressions and weather–related expressions we aim to address other aspects of the weather domain, such as metaphorical meanings, nominal weather expressions or force–dynamic aspects of the weather domain in future research.

References

Andrason, Alexander (2019). Weather in Polish: a contribution to the typology of meteorological constructions. *Studia Linguistica* 73(1): 66–105, https://doi.org/10.1111/stul.12091

Babić, Stjepan (1982). Tko ima više naziva za snijeg: Eskimi ili mi?. *Jezik: časopis za kulturu hrvatskoga književnog jezika* 30(2): 60–62

Barić, Eugenija, Mijo Lončarić, Dragica Malić, Slavko Pavešić, Mirko Peti, Vesna Zečević and Marija Znika (1997). *Hrvatska gramatika*. Zagreb: Školska knjiga

Belaj, Branimir (2007). Konceptualnosemantički aspekti prototipnih struktura s bezličnim i obezličenim glagolima i njihove sintaktičke implikacije. Kuna, Branko, ur. *Sintaktičke kategorije. Zbornik radova znanstvenoga skupa s međunarodnim sudjelovanjem Hrvatski sintaktički dani*. Osijek – Zagreb: Filozofski fakultet – Institut za hrvatski jezik i jezikoslovlje, 21–50

Belaj, Branimir i Goran Tanacković Faletar (2015). O prezentativno–demonstrativnim funkcijama medijalnoga oblika TO u konstrukcijama s bezličnim glagolima. *Fluminensia* 2: 61–81

Bleotu, Adina Camelia (2012). Why does IT always rain on me? On weather verbs. *Proceedings of the First Central European conference in linguistics for postgraduate students*. Budapest: Pázmáy Péter Catholic University, 59–81

Boas, Franz (1911). *Introduction. Handbook of American Indian Languages, Vol.1*. Bureau of American Ethnology, Bulletin 40. Washington: Government Print Office. (Smithsonian Institution, Bureau of American Ethnology)

Bolinger, Dwight (1973). Ambient "it" is meaningful too. *Journal of Linguistics* 9(2): 261–270, https://doi.org/10.1017/S0022226700003789

Brdar–Szabó, Rita and Mario Brdar (2003). The manner for activity metonymy across domains and languages. *Jezikoslovlje* 4(1): 43–69

Dowty, David R. (1979). *Word Meaning and Montague Grammar. The Semantics of Verbs and Times in Generative Semantics and in Montague PTQ*. Dordrecht: D. Reidel Publishing Company, https://doi.org/10.1007/978–94–009–9473–7

Eriksen, Pål Kristian, Seppo Kittilä and Leena Kolehmainen (2010). The linguistics of weather: Cross–linguistic patterns of meteorological expressions. *Studies in Language* 34(3): 565–601, https://doi.org/10.1075/sl.34.3.03eri

Eriksen, Pål Kristian, Seppo Kittilä and Leena Kolehmainen (2012). Weather and language. *Language and Linguistics Compass* 6(6): 383–402, https://doi.org/10.1002/lnc3.341

Eyþórsson, Þórhallur and Sigríður Sæunn Sigurðardóttir (2016). A brief history of Icelandic weather verbs: Syntax, semantics and argument structure. *Working Papers in Scandinavian Syntax* 96: 91–125

Gelo, Branko et al. (2005). *Meteorološki pojmovnik i višejezični rječnik*. Zagreb: DHMZ

Givón, Talmy (2001). *Syntax: an introduction* (Vol. 1). Amsterdam/Philadelphia: John Benjamins Publishing. https://doi.org/10.1075/z.syns

Gluhak, Alemko (1993). *Hrvatski etimološki rječnik*. Zagreb: August Cesarec

Gluhak, Alemko (2004). Praslavenska riječ *dъždъ 'kiša' i 'rosno nebo'. *Filologija* 43: 11–30

Goddard, Cliff (2001). Lexico–Semantic Universals: A Critical Overview. *Linguistic Typology* 5(1): 1–65, https://doi.org/10.1515/lity.5.1.1

Jackendoff, Ray (1985). *Semantics and Cognition*. Cambridge, MA: MIT Press

Jojić, Ljiljana (ed.) (2015). *Veliki rječnik hrvatskoga standardnog jezika*. Zagreb: Školska knjiga

Katičić, Radoslav (2002). *Sintaksa hrvatskoga književnog jezika, 3. izdanje*. Zagreb: Nakladni zavod Globus

Katušin, Zvonimir (ed.) (2007). *Međunarodni atlas oblaka* (Vol. 1 and 2). Zagreb: DHMZ

Kienpointner, Manfred (2016). Weather verbs in Latin, German, and other languages. Contrastive and typological remarks. *Pallas. Revue d'études antiques* 102: 57–67, https://doi.org/10.4000/pallas.3561

Langacker, Ronald W. (1991). *Foundations of Cognitive Grammar*, vol. 2. Stanford, CA: Stanford University Press, https://doi.org/10.1515/9780804764469

Levin, Beth and Bonnie Krejci (2019). Talking about the weather: Two construals of precipitation events in English. *Glossa: a journal of general linguistics* 4(1): 58. 1–29, https://doi.org/10.5334/gjgl.794

Ljubešić, Nikola and Filip Klubička (2014). {bs, hr, sr} wac–web corpora of Bosnian, Croatian and Serbian. Bildhauer, Felix and Roland Schäfer, eds. *Proceedings of the 9th Web as 22.1 (2021): 5–36 35 Corpus Workshop (WaC–9)*, 29–35. Gothenburg: Association for Computational Linguistics, doi: 10.3115/v1/W14–0405

Menac, Antica (1986). Partitivni genitiv u ruskom i hrvatskom književnom jeziku. *Filologija* 14: 183–196

Meulleman, Machteld and Katia Paykin (2016). Weather verbs sifted through a motion sieve. *Contrastive Linguistics* 5: 58–67

Penzar, Ivan and Branka Penzar (1992). Riječ oborina i njezino značenje u meteorologiji. *Hrvatski meteorološki časopis* 27(27): 53–61

Poje, Dražen (2002). Meteorološko nazivlje u hrvatskim rječnicima. *Hrvatski meteorološki časopis* 37(37): 37–55

Polančec, Jurica (2020). *A Typology of Aspect–Actionality Interactions*. PhD dissertation. Zadar: Sveučilište u Zadru

Pranjković, Ivo (2020). Gramatička svojstva glagola zbivanja. Palić, Ismail, ed. *Sarajevski filološki susreti 5* (vol.1). Sarajevo: Bosansko filološko društvo

Raffaelli, Ida and Barbara Kerovec (2008). Morphosemantic fields in the analysis of Croatian vocabulary. *Jezikoslovlje* 9(1–2): 141–169

Ruwet, N. (1986). On Weather Verbs. *Papers from the Regional Meetings*, Chicago Linguistic Society. 22.1. P. 195–215

Ruwet, N. (1989). Weather verbs and the unaccusative hypothesis. Kirschner, Carl and Janet Ann DeCesaris, eds. *Studies in Romance Linguistics: Selected Proceedings from the XVII Linguistic Symposium on Romance Languages*. Amsterdam/Philadephia: John Benjamins Publishing, 313–345, https://doi.org/10.1075/cilt.60.20ruw

Samardžija, Marko (1986). *Valentnost glagola u suvremenom hrvatskom jeziku*. PhD dissertation. Zagreb: Filozofski fakultet Sveučilišta u Zagrebu

Silić, Josip and Ivo Pranjković (2007). *Gramatika hrvatskog jezika za gimnazije i visoka učilišta*. Zagreb: Školska knjiga

Swadesh, Morris (1972 [1960]). What is glottochronology? Swadesh, Morris and Joel Sherzer, eds. *The Origin and Diversification of Languages*. London: Routledge and Kegan Paul, 271–284

Šojat, Krešimir (2008). *Sintaktički i semantički opis glagolskih valencija u hrvatskom jeziku.* PhD dissertation. Zagreb: Filozofski fakultet Sveučilišta u Zagrebu

Šonje, Jure (2000). *Rječnik hrvatskog jezika*. Zagreb: Leksikografski zavod Miroslav Krleža i Školska knjiga

Talmy, Leonard (1985). Lexicalization patterns: Semantic structure in lexical forms. *Language typology and syntactic description* 3(99): 36–149

Talmy, Leonard (2000). *Toward a cognitive semantics* (Vol. 2). Cambridge, MA: MIT Press, https://doi.org/10.7551/mitpress/6848.001.0001

Van Valin, Robert D. and Randy La Polla (1997) *Syntax: structure, meaning, and function.* Cambridge: Cambridge University Press, https://doi.org/10.1017/CBO9781139166799

Vendler, Zeno (1957). Verbs and times. *The philosophical review*, 143–160, https://doi.org/10.2307/2182371

Electronic sources

Croatian Web Corpus (hrWaC)

http://nlp.ffzg.hr/resources/corpora/hrwac/

Hrvatsko strukovno nazivlje – Struna

http://struna.ihjj.hr/

Hrvatski jezični portal

https://hjp.znanje.hr/

Domena meteorološkog vremena u hrvatskom: korpusno utemeljen opis oborinskih i neoborinskih meteoroloških izraza

U ovoj se studiji istražuju meteorološki izrazi u hrvatskom, s fokusom na oborine i neoborinske meteorološke pojave. Domena meteorološkog vremena u hrvatskom smješta se u kontekst postojećih istraživanja *verba meteorologica* u hrvatskim jezikoslovnim opisima s jedne strane, te u kontekst tipoloških istraživanja sintaktičkih struktura kojima se opisuju meteorološki događaji s druge. Kako bi se proširila opisana jezična građa u vezi s domenom meteorološkog vremena u hrvatskom u radu je odabran korpusno utemeljen pristup, te se raspravlja o prednostima i izazovima takvog pristupa, i predlaže način na koji se građa može klasificirati za potrebe analize na temelju primarnih, sekundarnih te metaforičkih značenja meteoroloških izraza. Na temelju korpusnih potvrda i leksikografskih podataka, kategoriji meteoroloških glagola (*kišiti*, *sniježiti* i dr.) pridodani su glagoli koji označuju način zbivanja, npr. *rominjati*, kao i druge sintaktičke strukture kojima se može označiti meteorološki događaj, a izlaze izvan uporabe bezličnih glagola, npr. *maglovito je*, *mraz je na tlu*. Unutarjezična raznolikost meteoroloških izraza u hrvatskom nadalje se analizira prema semantičkim svojstvima dinamičnih i statičnih meteoroloških pojava, leksikalizacijskim obilježjima koja su kodirana meteorološkim glagolima (Lik, Put i Način), kontekstnom usidrenju meteorološkog izraza, načinu zbivanja te pojmovima primarnih i sekundarnih meteoroloških izraza. Na temelju leksikografskih i korpusnih podataka, u ovom se radu predlaže nacrt za daljnje sustavno proučavanje domene meteorološkog vremena, posebice s obzirom na leksikalizaciju meteoroloških događaja i pojavnosti.

Keywords: weather verbs, semantic typology, lexicalization, Croatian
Ključne riječi: meteorološki glagoli, semantička tipologija, leksikalizacija, hrvatski jezik