Understanding 'many' through the lens of Ukrainian *багато*

Laura A. Janda[1] [iD] · **Yuliia Palii**[1]

Abstract

We reveal an ongoing language change in Ukrainian involving a construction with a subject comprised of the indefinite quantifier *багато* 'many' modifying a noun phrase in the Genitive Plural. Number agreement on the verb varies, allowing both Singular (in 69.1% of attestations) and Plural (in 30.9% of attestations). Based on statistical analysis of corpus data, we investigate the influence of the factors of year of creation, word order of subject and verb, and animacy of the subject on the choice of verb number. We find that, while all combinations of word order and animacy are robustly attested, VS word order and inanimate subjects tend to prefer Singular, whereas SV word order and animate subjects tend to prefer Plural. Since about the 1950s, the proportion of Plural has been increasing, overtaking Singular in the current decade. We propose that this Singular vs. Plural variation is motivated by the human embodied experience of construing a group of items as either a homogeneous mass (and therefore Singular) or a multiplicity of individuals (and therefore Plural). This proposal is supported by the identification of micro-constructions that prefer Singular and show reduced individuation of human beings.

1 Introduction

When you say "A lot of people came", do you conceive of the people as a unitary homogeneous mass or as distinct individuals? In Ukrainian (and in many other languages), this difference can be signaled by means of Singular vs. Plural number agreement on the verb. In Ukrainian this competition between Singular and Plural agreement can be observed for all quantifiers, only one of which is the focus of our study. We investigate the choice of verb number with a subject that consists of the Ukrainian indefinite quantifier *багато* 'much, many' and a Genitive Plural noun phrase, henceforth referred to as the "*bahato* construction". In this construction, Singular verb forms are found in 69.1% of examples, as opposed to Plural in 30.9%. Corpus data reveals an ongoing language change in Ukrainian, influenced also by the animacy of the noun and the order of the subject and verb, and we measure these effects by means of a logistic regression model.

✉ L.A. Janda
laura.janda@uit.no

Y. Palii
yuliia.palii@uit.no; julmay19@gmail.com

[1] ISK, HSL, UiT The Arctic University of Norway, 9037 Tromsø, Norway

Examples (1-a)–(1-d) from our dataset extracted from the GRAC corpus (see Sect. 4.1) illustrate variations in word order and the choice of verb number in the *bahato* construction with the same noun *люди* 'people' and the same verb *прийти* 'come'. In these examples we see both subject-before-verb (SV) word order ((1-a) and (1-b)), and verb-before-subject (VS) word order ((1-c) and (1-d)). For each word order, both Singular ((1-a) and (1-c)) and Plural ((1-b) and (1-d)) verb forms are observed.

(1)   a.   В акції брали участь лікарі, студенти, вчителі, **багато людей прийшло** з дітьми, – розповив Геннадій. (Світлана Талан «Оголений нерв», 2015)
'Doctors, students, teachers participated in the demonstration, **many people came** with their children – said Hennadij.'

       b.   Акція проходила мирно, на Майдані був концерт, **багато людей прийшли** з дітьми. (Газета «Свобода», 2019)
'The demonstration took place peacefully, there was a concert on the Majdan square, **many people came** with their children.'

       c.   Тоді **прийшло багато людей**, і гостям сподобалось. (Онлайн-ЗМІ «Маєш право знати - Бровари», 2014)
'**Many people came** then, and the guests liked it.'

       d.   На свято **прийшли багато людей**, всі навіть не помістилися у середині храму. (Інтернет-газета «Версії», 2017)
'**Many people came** for the holiday, there wasn't even room for all of them inside the church.'

The Singular versus Plural variation observed in the *bahato* construction is arguably of special interest to linguists because it presents a clash between grammatical form and semantic reference. While the indefinite quantifier is grammatically Neuter Singular, semantically *bahato* + Genitive Plural always refers to a multiplicity of items. Since both Singular and Plural verb forms are observed in this construction, this language variation raises a variety of questions:

- What is the relative frequency of Singular versus Plural verb number in the *bahato* construction?
- Has this relative frequency changed over time, and if so, how?
- How do word order and animacy influence the choice of verb number?

We motivate the theoretical basis for our study in Sect. 2 and engage previous scholarship on verb number variation in both Ukrainian and other languages in Sect. 3. Our data and the variables used in our analysis are described in Sect. 4. Section 5 presents our statistical analysis. Case studies of specific verbs and constructional variants are found in Sect. 6. We offer conclusions in Sect. 7.

2 Theoretical basis

Cognitive Linguistics (Lakoff, 1987; Geeraerts & Cuyckens, 2007; Langacker, 2008; Dąbrowska & Divjak, 2015; Dancygier, 2017) is the framework for this study of the *bahato* construction. Our study is further informed by relevant scholarship that acknowledges the interaction between syntactic and semantic agreement (Corbett, 2000: Chap. 6, Kibrik, 2019) and the role of definiteness or specificity in motivating semantic (and therefore Plural) agreement (cf. studies of this phenomenon in Serbian: Arsenijević, 2016; Milosavljević, 2018). Cognitive Linguistics takes the perspective that linguistic phenomena can be scalar,

meaning that we observe tendencies. For the *bahato* construction, this means that we will focus on the frequency distribution of Singular vs. Plural forms and the relationship between that distribution and various associated factors.

As a subfield of Cognitive Linguistics, Construction Grammar (Croft, 2001; Goldberg, 2006) approaches language as a system of interconnected constructions. "Constructions are defined to be conventional, learned form-function pairings at varying levels of complexity and abstraction" (Goldberg, 2013, p. 17; cf. similar definitions in Lakoff, 1987; Fillmore et al., 1988; Wierzbicka, 1988; Goldberg, 1995, 2006). The *bahato* construction we focus on in this article, comprised of a Subject quantified by *багато* and a verb, meets this definition. In the *bahato* construction, we find the lexical meaning of *багато* itself, along with the lexical meanings of the noun it modifies and the verb. The noun is marked as Genitive Plural, and the verb, in addition to tense, is marked for number, a value that can vary. Here we focus on the meaning of Singular vs. Plural on the verb.

The meaning of the grammatical distinction of Singular vs. Plural number might seem unremarkable or even trivial, but it invokes a complex of embodied human experiences. Johnson (1987) coined the term "image schemas", derived from Kant's use of the term schema to describe structures in human imagination. Johnson identifies several image schemas that organize meaning, such as CONTAINER (the experience of putting things in containers), PATH (the experience of movement along a trajectory), BALANCE (the experience of holding our bodies in relation to gravity), NEAR-FAR (the experience of relative distance). Among these is the MASS-COUNT image schema based on our experience of physical entities as masses (like sand and water) or countable objects (like cats and boats). Johnson (1987, p. 26) furthermore describes the human capacity to perform transformations on image schemas, such as:

> "*Multiplex to mass*. Imagine a group of several objects. Move away (in your mind) from the group until the cluster of individuals starts to become a single homogeneous mass. Now move back down to the point where the mass turns once again into a cluster."

Johnson's claim is that the same group of objects can be construed both as a mass and as a number of individuals. Lakoff (1987, pp. 427–428) connects the relationship between multiplex and mass to the use of quantifiers in language, pointing out that English quantifiers *all* and *most* can modify both masses (*all sand*, *most water*) and multiple individuals (*all cats*, *most boats*). Langacker (2008, pp. 130–140) applies this idea to linguistic understanding of nouns and the grammatical distinction between Singular and Plural. According to Langacker, Plural nouns function grammatically more like masses than like Singular nouns, adding the syntactic contexts in ((2-a)–(2-b)) to Lakoff's observations about certain English quantifiers. However, Langacker points out that Plurals also differ from masses in their behavior with other quantifiers and with demonstratives as in (2-c), and that Plurals, unlike masses, can combine with numerals.

(2)　　a.　*They're looking for *cat/sand/cats.*
　　　　b.　*a cat/*sand/*cats*
　　　　c.　*those/these/many/few/several/numerous cats vs. that/this/much/little sand*
　　　　d.　*three cats/*sand*

Langacker suggests that there are two ways that masses are portrayed in language: either as a "non-plural mass noun" (like *sand* and *water*), or as a "plural mass noun" (like *cats* and *boats*). These two portrayals, visualized in Fig. 1, correspond to different construals of a referent as either an amorphous undifferentiated whole marked as Singular, or as a bounded

Non-plural Mass Noun   Plural Mass Noun
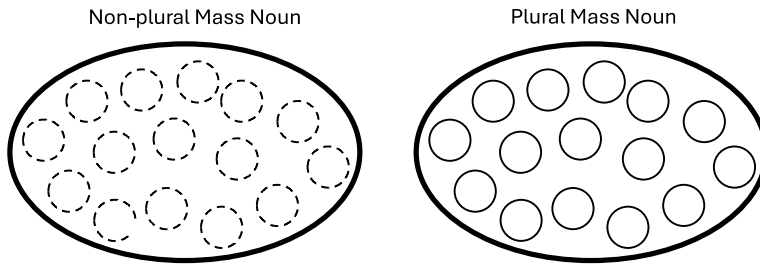
Fig. 1 Alternative construals of a group of entities as either mass or multiplex, adapted from Langacker (2008, p. 131)

group of individuals. The count-mass distinction is fluid and we, as human beings, can conceive of the same entity in alternate ways and "select the form whose meaning best suits our communicative intent". For example, we usually speak of sand as a mass and therefore Singular. However, we can also focus on individual, and therefore Plural, grains of sand.

The experience of a group of entities as either mass or multiplex is an example of "cognitive multistability" (Stadler & Kruse, 1995), similar in some ways to the experience of viewing an ambiguous image like a Necker cube or duck-rabbit illusion. We affirm that the option to choose Singular vs. Plural verb forms in the *bahato* construction is likewise motivated by cognitive multistability. This Singular vs. Plural choice is available also in other contexts, both in Ukrainian, as well as in other languages, for example in verb agreement with collective nouns, as in examples in (3-a)–(3-b) in Ukrainian and (4-a)–(4-b) in Spanish.

(3)   a.   *більшість **вважає**, що*
             *bil'šist' **vvažaje**, ščo*
             [majority consider.PRS.3**SG** that]
             'the majority **considers** that'
       b.   *більшість **вважають**, що*
             *bil'šist' **vvažajut'**, ščo*
             [majority consider.PRS.3**PL** that]
             'the majority **consider** that'

(4)   a.   *la mayoría **está** de acuerdo en*
             [the majority be.PRS.3**SG** of agreement in]
             'the majority **agrees** that'
       b.   *la mayoría **están** de acuerdo en*
             [the majority be.PRS.3**PL** of agreement in]
             'the majority **agree** that'

While Singular verb forms are more common in our data on the *bahato* construction (see Sect. 4), Plural forms are attested across the entire period covered by our data, including the earliest attestation we have of the *bahato* construction from 1819, shown here in example (5), where the use of Plural supports an interpretation of the predicate as signaling a continuous event referring to many individuals rather than a mere qualificative interpretation of a sum of suitors.

(5)   Я догадуюсь; тут живе одна бідна вдова з дочкою, то, мабуть, на Наталці возний засватався, бо до неї **багато женихів залицялись**. (Іван Котляревський «Наталка Полтавка», 1819)

> 'I guess so; here we have a widow living with her daughter and it seems that the coachman married Natalka, since **many suitors were courting** her.'

While variation between alternate linguistic forms is a necessary, though not sufficient, criterion for language change (Croft, 2000; Bybee, 2015), the question of whether a change is taking place in Ukrainian has not been previously addressed. We contribute the first study tracking the diachronic trajectory of an ongoing language change in Ukrainian motivated by alternative construals of multiplex realia. We observe in our data an increase in the use of Plural over time in the *bahato* construction.

3 'Many' in Ukrainian and in Slavic languages

We situate the *bahato* construction in the context of numeral syntax and agreement patterns in Ukrainian, and make comparisons with Russian and with other Slavic languages.

3.1 The bahato construction

Ukrainian has three types of quantifiers that can be used to describe quantities of more than one entity: cardinal numerals like *два/две* 'two' and *три* 'three', collective numerals like *двоє* 'two(some)' and *троє* 'three(some)', and indefinite quantifiers like *мало* 'little, few', *кілька* 'how many', and *багато* 'much, many'. In subject position, the cardinal numerals for '2', '3', and '4', plus 'both' and 'one and a half' – a group of numerals often termed 'paucal' – combine with noun phrases in the Nominative Plural (or Genitive Singular if the noun refers to paired items). All other numerals – cardinal numerals for '5' and above, collectives, and indefinite quantifiers – combine with noun phrases in the Genitive Plural (Arpolenko et al., 1980), including the indefinite numeral *bahato*. When used to quantify a subject, all quantifiers can be combined with both Singular and Plural verb forms. The examples in (6) show a subject consisting of the cardinal numeral *п'ять* 'five', the Genitive Plural form of *чоловік* 'person', and the verb *загинути* 'perish' in a past tense form. The verb form is Singular in (6-a), but plural in (6-b).

(6)   a.   Щонайменше **п'ять чоловік загинуло** та 16 отримали ушкодження в результаті вибуху у Владикавказі. (Газета «День», 2000)
'At least **five persons perished** and 16 were injured as a result of an explosion in Vladikavkaz.'

 b.   У Чехії, Польщі та Німеччині не менш як **п'ять чоловік загинули** внаслідок урагану "Герварт" (Онлайн-ЗМІ «LB.ua», 2017)
'In the Czech Republic, Poland, and Germany no fewer than **five persons perished** due to Storm Hervart'

Scholars of Ukrainian have engaged the topic of number agreement in verbs with quantified subjects since the early 20th century. Kurylo (1942, pp. 56–58, citing earlier observations by Potebnja) notes that the predicate verb in Plural is used in collocations with an animate subject, while collocations with a predicate verb in the impersonal form (Singular) answer the question How much? The latter impersonal constructions with a singular form are transitional between subject and non-subject constructions, and the transition is reflected in the ambiguity of the choice between Singular and Plural based on meaning. However, Antonenko-Davidovič's (1979, p. 136) manual of speech norms recommends that the verb

form should be in the Singular when the subject is governed by a "collective word" (his term for indefinite quantifiers such as *багато*) combined with nouns (adjectives) in genitive case.

Reflecting this tradition, Hryhor'jev et al. (2005, pp. 158–159) note several factors that influence the choice of number on the verb: a conceptual factor, a lexical factor, a syntactic factor, and a semantic factor. If conceptual focus is on the number of items the verb tends to be Singular, whereas focus on the items as separate individuals is associated with Plural verb forms. This is the same conceptual factor discussed in more detail in Sect. 2. The lexical factor has to do with the type of numeral, and Singular is most commonly used with an indefinite quantifier like *багато* 'much, many'. The syntactic factor has to do with the relative order of the subject and verb, with Singular preferred if the verb comes before the subject, and Plural preferred if the subject is before the verb. The semantic factor is animacy: inanimate nouns prefer Singular and animate nouns prefer Plural. Inanimate nouns referring to time have a particularly strong preference for Singular.

The conceptual factor pertains to the internal construal of a native speaker and is therefore not available for objective observation. The remaining three preferences stated by Hryhor'jev et al. (2005) are merely tendencies. It is easy to find examples that both confirm and defy these tendencies. Compare the examples in (1) and (7) and the visualization of factors in Table 1.

The examples in (1) show that both Singular and Plural verb forms occur with the animate noun *люди* 'people' in both SV and VS word order. All four examples in (7) feature the same inanimate noun: *реч* 'thing'. The examples in (7) show all possible combinations of Singular ((7-a) and (7-c)) and Plural ((7-b) and (7-d)) verb forms with both SV ((7-a) and (7-b)) and VS ((7-c) and (7-d)) word order for an inanimate noun. Collectively the examples in (1) and (7) present all eight possible combinations of verb number, word order, and animacy in the *bahato* construction.

(7)   a.   Оскільки збиралися похапцем, **багато речей залишилося** вдома, до якого кримські татари обов'язково повернуться. (Онлайн-ЗМІ «Zaxid.net», 2017)
           'Since they left in a hurry, **many things got left** at home, where the Crimean Tatars will definitely return.'

      b.   Із уже згаданих причин **багато речей залишилися** на папері. (Інтернет-газета «Дзеркало тижня», 2020)
           'For the reasons previously mentioned, **many things remained** on paper.'

      c.   Після СРСР нам **залишилося багато речей**, котрі нас і надалі тримають в спільному просторі... (Онлайн-ЗМІ «Zaxid.net», 2009)
           'After the USSR, we were **left** with **many things** that will continue to keep us in a common space.'

      d.   це досить суттєві гроші і таким способом **фінансувалися багато речей** в країні, включаючи політичну корупцію. (Журнал «НВ», 2018)
           'this is a substantial amount of money and this is how **many things** in the country **were financed**, including political corruption'

Table 1 identifies each example in (1) and (7) on the left, followed by the grammatical number of the verb form observed in each example. The lexical, syntactic, and semantic factors pertinent to each example are distributed across the two rightmost columns according to whether they should be expected to combine with Singular vs. Plural verb forms. From the perspective of Hryhor'jev et al. (2005), the indefinite quantifier *багато* 'many' is associated with Singular, as are VS word order and inanimate nouns, whereas SV word order and animate nouns are associated with Plural. Boldface is used to highlight the examples where two or all three factors suggest a stronger preference for Singular or Plural, and exclamation

Table 1 Comparison of factors associated with Singular vs. Plural verbs forms according to Hryhor'jev et al. (2005) with observed verb number in examples in (1) and (7). Boldface highlights predominant factors. The symbols (!) and (!!) indicate observed verb number that conflicts with some or all factors

Example	Observed verb number	Factors associated with Singular forms	Factors associated with Plural forms
(1-a)	Singular (!)	*багато*	**SV, animate**
(1-b)	Plural	*багато*	**SV, animate**
(1-c)	Singular	***багато***, VS	animate
(1-d)	Plural (!)	***багато***, VS	animate
(7-a)	Singular	***багато***, inanimate	SV
(7-b)	Plural (!)	***багато***, inanimate	SV
(7-c)	Singular	***багато***, VS, inanimate	
(7-d)	Plural (!!)	***багато***, VS, inanimate	

marks indicate examples where the number of the verb conflicts with that preference. In one example (1-a) we have a Singular verb despite the presence of two factors that would support the choice of Plural. In three examples, (1-d), (7-b), and (7-d) we find a Plural verb where we might expect a Singular, and this is particularly surprising for (7-d), where all three factors suggest a preference for Singular.

In sum, we see that the preferences for Singular vs. Plural verb forms identified by Hryhor'jev et al. (2005, pp. 160–161) cannot be stated as rules because they neither dictate nor exclude any combination of factors. Even an animate noun in SV word order can combine with a Singular verb form as in (1-a), and even an inanimate noun in VS word order can combine with a Plural verb form as in (7-d). The choice of either number seems to be always available, and further statistical analysis is needed to confirm the stated tendencies and determine which factors carry the most weight. The interplay between these factors has not been previously subjected to statistical analysis.

3.2 Comparison with Russian

Nesset and Janda (2023) undertook a statistical analysis of verb number agreement with quantified subjects in Russian. That study examined all cardinal numerals from '2' through '20', tens from '30' through '90', and hundreds up through '900', as well as collectives and indefinite quantifiers, including *много* 'many', the translation equivalent of Ukrainian *багато*. Nesset and Janda found that both word order and animacy play a role in the choice of verb number in Russian, with word order serving as a slightly stronger factor, and they found an interaction between the two factors. The directions of preferences, namely that SV word order and animate nouns tend to prefer Plural, are the same as those cited above from Hryhor'jev et al. (2005). A striking finding for Russian was that no discernible change from the baseline frequency of 55.96% Plural was found over a period of more than 200 years.

Although Nesset and Janda's study involved nearly 40,000 examples, only 105 of those involved *много* 'many', and this is too little data to establish whether the trends found for Russian numerals overall pertain specifically to *много*. We have therefore undertaken a targeted study of Russian *много* 'many', collecting 6,612 examples of both SV and VS word order from the Russian National Corpus over a period from 1742 to 2021. Of these, only 353 examples (5.3%) show Plural verb agreement, a striking difference from Ukrainian *багато* with 30.9% Plural agreement. This discrepancy is at least partly explained by the fact that

Ukrainian *багато* is a translation equivalent not just of Russian *много*, but also of Russian *многие* 'many', which combines predominantly with animate nouns and always agrees with a Plural verb form. A logistic regression analysis of our Russian *много* data reveals a highly significant and strong effect of word order (p < 2e−16, the functional equivalent of zero probability that this could happen by chance and coefficient of 2.5 increasing the chance of Plural for examples with SV word order), but no significant effect for the year of creation (p = 0.32, meaning that there is a 32% chance that the distribution over time is merely random). As we show below, by contrast Ukrainian is undergoing a language change in which Plural agreement is on the rise over time.

We confront corpus data on the Ukrainian *bahato* construction with a similar analysis that takes into account the year of creation as well as the word order and animacy of the noun. Like Nesset and Janda, we take into account the individual preferences (random effects) of the verbs, however there are several variables that appeared in that study but are not relevant for the *bahato* construction since we focus on only one quantifier, namely: quantifier type, quantifier frequency, and random effects of quantifiers.

We address the syntactic factor of word order and the semantic factor of animacy in our statistical model of data representing the *bahato* construction in Sect. 5 and go into more detail with specific verbs in the case studies in Sect. 6.

3.3 'Many' across the Slavic languages

A comprehensive analysis of expressions meaning 'many' in the Slavic languages goes beyond the scope of this article. We give merely a rather superficial survey to provide perspective for the semantic origins of Ukrainian *багато*. This is a survey of only some of the Slavic languages and their most common words for 'many'. We focus primarily on the languages in closest proximity to Ukrainian, namely Russian, Belarusian, Polish, and Slovak, and include Czech, Serbian, and Bulgarian for additional context.

Indefinite quantifiers other than 'many' tend to reflect Common Slavic origins fairly uniformly across the modern Slavic languages. 'Little, few' is *мало* in Ukrainian, Russian and Serbian, *мала* in Belarusian, *mało* in Polish, *málo* in both Slovak and Czech, and *малко* in Bulgarian. 'Some, several' shares a root with 'how many' and this relationship is transparent in most Slavic languages. In Ukrainian 'how many' is *скільки*, while 'some, several' is *декілька* or *кілька*. In the following languages the relationship is marked by the prefix *nĕ-*: Russian (*не*)*сколько*, Belarusian *коль
кі/некалькі*, Slovak (*nie*)*koľko*, Czech (*nĕ*)*kolik*, Serbian (*не*)*колико*, Bulgarian (*ня*)*колко*. Polish partially deviates from this pattern with *ile* 'how many', *kilka* 'some'.

As displayed in Table 2, the concept of 'many' has a diverse array of expressions and etymological sources across the Slavic languages. In all these languages except Bulgarian, the quantifiers combine with a noun phrase in the Genitive Plural.

The status of these words in terms of parts of speech varies according to the sources consulted. For example, Russian *много* is identified by some dictionaries as a numeral, and by others as an adverb, and by some as both a numeral and an adverb (cf. Ožegov & Švedova 1992).

Table 2 captures only some of the variety available in the various languages. Ukrainian has *чимало* as a less frequent synonym. *Mnogo* is possible, but uncommon in Polish. Russian can also use *полно*, cognate to Serbian *пуно* in the meaning 'many', both Czech and Polish can use *moc* (related to 'able') colloquially, and other examples could be added. Furthermore, it is not uncommon to find adjectival quantification in Slavic, as in Russian *многие* 'many (usually: people)', although this is not available in Ukrainian.

Table 2 Indefinite quantifiers meaning 'many' across a sample of Slavic languages. Glosses for source forms are approximate. Boldfacing indicates lexemes that corpora report to be of highest frequency. Languages are identified by their ISO codes. CSl = Common Slavic, Lith = Lithuanian, Goth = Gothic, Gmc = Germanic

	mŭnogo 'many' (CSl)	*velĭ-* 'great' (CSl)	*pĭlno* 'full' (CSl)	*bogato* 'rich' (Old Persian)	*daug* 'strong, useful' (Lith/Goth)	**gōđa* 'good' (Gmc)	? (unclear origin)
Ukr				***багато***			*шмат*
Rus	***много***						
Bel	*многа*			*багата*			***шмат***
Pol		***wiele/wielu***			*dużo*		
Slk	*mnoho*	***veľa***					
Ces	*mnoho*					***hodně***	
Srp	***много***		*пуно*				
Bul	***много***						

Ukrainian *багато* is related to Common Slavic *bogŭ* 'god', which is probably an early loanword from Iran via the Scythians: cf. Old Persian *baga-*, Sanskrit *bhaga-* 'dispenser, gracious lord' and *bhaj-* 'divide, distribute, share' (Buck, 1949, p. 1464, Boldyrjev, 1982). The other etymological sources for 'many' shown in Table 2 include, in addition to *mŭnogo* (cognate to Gothic *manags* 'much, many', cf. Buck, 1949, pp. 922–923), Slavic roots relating to 'great, big' and 'full', and later loanwords for 'strong, useful' and 'good'. The provenience of *шмат* in Belarusian and (colloquial) Ukrainian is unclear. This survey demonstrates how the concept of 'many' has diverged in its evolution in the Slavic language family.

4 Data and variables

In this section, we document the extraction and pre-processing of our data. We also define the variables and provide descriptive statistics about their distribution in our dataset. The aggregate set of clean data, as well as the R code used in our statistical analysis is publicly available at https://doi.org/10.18710/Y7VGQE (Janda & Palii, 2024).

4.1 Data and dependent variable verb number

The General Regionally Annotated Corpus of Ukrainian (GRAC, Shvedova et al., 2017–2024) was consulted to collect data for further analysis concerning the distribution of Singular vs. Plural verb forms in the target *bahato* construction. GRAC is a Sketch Engine corpus of over 1.8 billion words, representing texts from over 30,000 authors created between 1816 and 2023. This corpus is designed to serve as source material for linguistic research on Standard Ukrainian. Our data was collected during the month of February 2024.

We undertook four searches for the *bahato* construction where the NP is Genitive Plural, stratified by Verb number and by Word order (Subject-Verb = SV; Verb-Subject = VS), yielding data on the four possible variants: Sg-SV, Pl-SV, Sg-VS, Pl-VS. All searches were for subjects and verbs that were contiguous; our data did not include examples with intervening words (cf. Nesset & Janda, 2023 which was similarly limited to examples with contiguous subjects and verbs to facilitate preprocessing). We discovered that the use of a wildcard for verbs did not return examples for the Singular form *було* 'was', so two additional searches were undertaken to extract examples with that form in both SV and VS Word order. These

Verb number	Word order	# of examples	% of dataset
Sg	SV	3,550	12.46%
Pl	SV	7,348	25.79%
Sg	VS	16,150	56.68%
Pl	VS	1,443	5.06%
Total		28,491	100%

Table 3 *Bahato* construction data extracted from GRAC corpus, after removal of duplicates and irrelevant examples, grouped according to Verb number and Word order

searches returned an aggregate total of 42,084 examples along with the year of creation for each example ranging from 1819 to 2023. However, these results included both duplicates and irrelevant examples, in particular examples in which *багато* modifies a direct object rather than the subject of a verb. The nature of these types of noise in the data made it necessary to remove them by means of manual inspection. The portion of noise in the data was especially high among the VS-Pl results: of the 7,896 examples returned by Sketch Engine, 6,453 (82%) were duplicates or otherwise irrelevant, leaving only 1,443 relevant examples of the *bahato* construction. Altogether, 32.3% of examples returned from the searches were excluded, leaving 28,491 examples of clean data. Table 3 shows the distribution of clean data for the *bahato* construction. These data were aggregated for the statistical analysis in Sect. 5.

The overall distribution of number for verb forms in the *bahato* construction in this data is 69.1% Singular vs. 30.9% Plural.

4.2 Variables

In addition to the dependent variable of Verb number, we take into account both fixed effect (numerical and categorical) and random effect (open-ended items with individual preferences) variables. Our numerical fixed effect predictor variable is Year. There are two categorical predictor variables: Animacy and Word order. We have one random effect, called Verb lemma, coded as the infinitive form of each verb.

4.2.1 Year

The range of years represented in our data is highly skewed, with a median of 2013, meaning that 50% of the data cover a period of nearly two centuries from 1819 to 2013, whereas 50% cover just one decade from 2013 to 2023. In order to represent this variable appropriately in the model, we have scaled it so that it is centered on the mean and z-scored, meaning that each unit increase is equal to the standard deviation in the distribution, which is 23.5 years. The scaled variable is labeled "Year.sc" and is distributed as shown in Table 4, which can be used as a conversion table to interpret the analysis of Year.sc in Sect. 5.

4.2.2 Animacy

Given the fact that animacy is known to influence individualization and animacy was found to be a relevant factor in the analysis of the Russian quantified subject construction (Nesset & Janda, 2023), we manually annotated examples for animacy of the noun phrase quantified by *багато*. Our annotation revealed that human beings are referenced in 16,896 examples, whereas animals appear in 385 examples, and inanimates in 11,210 examples. For the purposes of our statistical analysis, we have combined humans and animals into a single animate category comprising 60.7% of the data.

Table 4 Comparison of values for Year and the scaled variable Year.sc

	Year	Year.sc
Minimum	1819	−7.9
1st quartile (25% of data is below this value)	1999	−0.2
Mean	2004	0.0
Median (50% of data is below this value)	2013	0.4
3rd quartile (75% of data is below this value)	2018	0.6
Maximum	2023	0.8

Table 5 The distribution of the 28,491 observations of the bahato construction according to Verb number, Word order, and Animacy

Verb number	VS Word order				SV Word order			
	Inanimate		Animate		Inanimate		Animate	
Singular	7,875	27.6%	8,275	29.0%	1,402	4.9%	2,148	7.5%
Plural	368	1.3%	1,075	3.8%	1,565	5.5%	5,783	20.3%

It should be noted that a binary distinction between animate vs. inanimate is a coarse measure since there are some inanimate nouns that refer to human beings via metonymy and can behave like animate nouns. For example, in our data the word *компанія* 'company' appears in some contexts nearly identical to those where we find *люди* 'people', as we see in examples in (8-a)–(8-b), both of which have the structure *багато* + Noun (Geinitive Plural) + *були змушені* 'were forced' + Infinitive.

(8)   a.   Пандемія стала важким ударом для індустрії, і **багато компаній були змушені** оголосити про банкрутство. (Журнал «НВ», 2020)
'The pandemic hit the industry hard and **many companies were forced to** declare bankruptcy.'

       b.   Через зростання вартості життя **багато людей були змушені** покинути житло в центрі міста та перебратись на окраїни. (Онлайн-ЗМІ «Хмарочос», 2019).
'Due to the rising cost of living, **many people were forced to** leave their homes in the city center and move to the suburbs.'

4.2.3 Word order

The overall distribution of Word order in our dataset is 17,593 examples of VS (61.7%), and 10,898 examples of SV (38.3%). A further breakdown of Word order by Animacy and Verb number is displayed in Table 5.

Table 5 shows that the three most common variants of the *bahato* construction are:

- VS, anim, sg (29.0% of the data), cf. example (1-c)
- VS, inanim, sg (27.6% of the data), cf. example (7-a)
- SV, anim, pl (20.3% of the data), cf. example (1-b)

Each of the remaining five combinations appear in less than 8% of the data, with the rarest being VS, inanim, Pl (1.3% of the data), cf. example (7-d). Note that this distribution comports with the predictions made in Table 1: none of the three most common variants

Table 6 Most common verb lemmas in our dataset

Verb lemma	Singular forms	Plural Forms	Total Frequency
бути 'be'	7,698 (93%)	599 (7%)	8,257
з'явитися 'appear'	648 (98%)	10 (2%)	658
загинути 'perish'	429 (82%)	92 (18%)	521
прийти 'come'	385 (78%)	108 (22%)	493
зібратися 'come together, gather'	444 (95%)	23 (5%)	467
залишитися 'stay, remain'	295 (67%)	148 (33%)	443

entails a conflicting prediction, and the rarest variant corresponds also to the variant with the strongest conflict – (7-d).

4.2.4 Verb lemma

We manually annotated each example in our dataset for the lemma (infinitive form) of the verb. There are a total of 2,311 unique verb lemmas in our dataset. Table 6 shows the frequency of the six most common verbs in our dataset, which collectively account for 38% of the total data.

Annotation for Verb lemma makes it possible for us to take into account the individual preferences of each verb in our statistical analysis, and this turns out to be quite important since, as we see in Table 4, there are strong differences in preferences. For example, *з'явитися* 'appear' prefers Singular in 98% of examples, while *залишитися* 'stay, remain' is close to the overall average with 67% Singular and 33% Plural.

4.3 Factors excluded from analysis

Due to limitations on technical feasibility and human resources, as well as the need to operationalize factors, several factors that might impact the distribution of Singular vs. Plural agreement in the *bahato* construction were not included in this study. Among excluded factors are the possible influence of intervening words, along with additional syntactic and semantic factors.

While annotation for subjective semantic factors such as the author's desire to focus on actions (and therefore Plural) as opposed to focusing on quantity (and therefore Singular) was not carried out, the differences in preferences of individual verb lemmas does reveal this factor, since verbs that prefer Singular focus on quantity, whereas verbs that prefer Plural tend to describe volitional activities of individuals (see Tables 9 and 10). Other factors that could not be taken into account include information structure, definiteness, topicality, territorial variation, and author preferences.

In order to address some of these limitations, in addition to our main dataset, we carried out two further studies targeting conjoined subjects and conjoined predicates, both of which are relatively infrequent phenomena in comparison with data in our main dataset.

It might be reasonable to expect that a conjoined subject would increase the probability of choosing Plural since the conjunction emphasizes the multitude of items, but our finding is that Singular remains the more common choice. We found 528 examples of conjoined subjects in the GRAC corpus, of which 323 (61.2%) showed Singular agreement as in (9-a), and 205 (38.8%) showed Plural agreement as in (9-b). This is not very different from the baseline distribution in our larger dataset, which has 69.1% Singular, and a chi-squared test

comparing the two datasets does not confirm a difference: while the difference is significant ($p < 0.001$), the Cramer's V effect size (0.023) is an order of magnitude too small to be considered meaningful.

(9) a. З критикою подібної рекламної кампанії **виступило багато активістів і активісток**, проте в акаунті закладу продовжили з'являтися пости із суперечливим змістом. (Журнал «НВ», 2019)
     '**Many activist men and women criticized** such an advertising campaign, but posts with controversial content continued to appear on the institution's account.'

     b. Після прийняття такого рішення з його критикою **виступили багато юристів і суддів**, визнавши що суддя Литвинова порушила норми Кримінально-процесуального кодексу і перевищила свої повноваження слідчого судді. (Онлайн-ЗМІ «LB.ua», 2014)
     'After the adoption of such a decision, **many lawyers and judges criticized** it, recognizing that Judge Litvynova violated the norms of the Criminal Procedure Code and abused her powers as an investigating judge.'

One might also expect that conjoined predicates could support the choice of Plural agreement since the conjunction would focus attention on a multitude of events. Our additional search of the GRAC corpus turned up 280 relevant examples. Indeed, it does seem that Plural is the preferred choice: 173 examples (61.8%) have Plural agreement as in (10-a), 94 (33.6%) have Singular agreement as in (10-b), and 13 examples (4.6%) have mixed agreement (usually with Singular for the first verb and Plural for the second one). However, in the larger picture it should be noted that conjoined predicates are a marginal phenomenon, accounting overall for less than 1% of usage in the *bahato* construction, and a chi-squared comparison with our larger dataset again shows a Cramer's V effect size (0.07) an order of magnitude too small to be considered meaningful.

(10) a. Дуже **багато депутатів виступало і говорило**, що цього року велика кількість фінансових ресурсів направлена в агропромисловий комплекс, безпосередньо в сільське господарство. (Стенограми засідань Верховної Ради України, 1997)
     '**Many deputies spoke and said** that that year a large amount of financial resources were directed to the agro-industrial complex, directly to agriculture.'

     b. Тут **багато депутатів виступали і говорили**, що основна причина, яка стримує розвиток фермерства, це те, що не завжди дають землю. (Стенограми засідань Верховної Ради України, 1991)
     '**Here, many deputies spoke and said**that the main reason that restrained the development of farming was that land was not always given.'

5 Statistical analysis

Our statistical model of the data on the *bahato* construction yields predictions about the effects of the variables Year, Animacy, and Word order on the choice of Singular vs. Plural verb forms. We present first the model and diagnostic measures of its performance, and then evaluate the impact of the predictor variables in more detail.

Table 7 Comparison of values observed in our dataset (rows) with the values predicted by our model (columns)

	Singular (predicted)	Plural (predicted)
Singular (observed)	18,339	1,361
Plural (observed)	1,804	6,987

Table 8 Results for fixed effects of mixed-effect logistic regression model predicting verb number from Year (as a scaled variable), Animacy, and Word order

	Estimate	Standard Error	z value	Pr ($>$|z|)	
(Intercept)	−2.15067	0.07040	−30.55	$< 2e{-}16$	***
Year.sc	0.93924	0.02988	31.44	$< 2e{-}16$	***
Animacyanim	0.62370	0.05320	11.72	$< 2e{-}16$	***
Word_orderSV	3.01155	0.04991	60.34	$< 2e{-}16$	***

5.1 Mixed-effects logistic regression model and diagnostics

We used the glmer() (generalized linear mixed model) function of R (version 4.3.3 (2024-02-29) – "Angel Food Cake") to produce a mixed-effects logistical regression model based on the 20,793 examples of the *bahato* construction in our data, using this formula:

$$\text{Verb\_number} \sim \text{Year.sc} + \text{Animacy} + \text{Word\_order} + (1|\text{Verb\_lemma})$$

The formula can be interpreted as: "The number of the verb is predicted in relation to the Year of the example, the Animacy of the quantified noun phrase, and the Word order of the example, taking into account Verb lemma as a random effect."

We initially tried a model that included also an interaction between Animacy and Word order, but discovered that the interaction was not significant (p-value = 0.23), so our model includes Animacy and Word order only as main effects.

The drop1() function showed that we should retain all the variables in our model. The C score value of 0.95 evaluates the fit of our model as excellent (Gries, 2021, pp. 335–336). All values from the Variance Inflation Factors are below 1.03, indicating that there are no problems with collinearity among the predictor variables. R-squared values for the model are good, accounting for 35%–65% of the variance in the data.

Our model correctly predicts the number of the verb for 88.9% of examples, as shown in the confusion matrix in Table 7, where observed values are in the rows and predicted values are in the columns. In the top row we see that 18,339 examples were both predicted and observed as Singular, but 1,361 examples that the model predicted as Plural were actually Singular in our data. The precision and recall values for both Singular and Plural are all 79% or higher.

Our statistical model was evaluated against a "null model" that would always choose the most common value for verb number, namely Singular, which has a baseline frequency of 69.1%. A binomial test shows that the probability that we could just by chance achieve a model that performs this much better than the baseline is zero.

Table 8 displays the results of our logistic regression model in terms of a prediction of Plural as opposed to Singular with reference to the Intercept. At the Intercept, the following values hold: Year.sc = 0.0 (= 2004), Animacy = inanimate, Word Order = VS.

The estimate is an effect size for prediction of Plural for verb number. Positive values indicate increased prediction of Plural relative to the Intercept, whereas negative values in-

dicate increased prediction of Singular. The standard error measures the differences between the regression predictions and observed values. All Standard error values are under 0.1, indicating a good fit of the model. The z values measure how extreme the estimate values are in comparison to a null model where the estimate is zero. Larger z values show that it is unlikely that one could obtain these estimate values by chance if the true value was zero. All z values exceed |11|, indicating more than 99% confidence level for the value of the estimate. The Pr ($>$|z|) column gives the p-value associated with each z value. The value listed as $< 2e-16$ (2 with the decimal point moved 16 places to the left) is the lowest value that R computes for this measure, essentially zero, meaning that these results are highly significant, also indicated by "***", signaling $p < 0.001$.

At the Intercept the estimate is -2.15, meaning that the model predicts a Singular verb form with an inanimate noun in the *bahato* construction with VS Word order in the year 2004. All other estimate values are adjustments to the predictions at the Intercept per unit of change for the given variable. The model's prediction of Plural increases with increased value for Year.sc, as well as with a change from inanimate to animate, and with a change from VS Word order to SV. It should be remembered that Year.sc ranges over nearly 9 units as opposed to Animacy and Word order with only two values each. This means that overall Year.sc has the strongest effect, with a total gain of almost 8.0 for the estimate. Next in strength comes Word order, with an estimate almost five times greater than Animacy. We address the predictions from the perspective of each predictor value in more detail below.

5.2 The effect of year

Figure 2 visualizes the effect of Year.sc on the choice of Singular vs. Plural for the number of the verb. The "rug" of ticks along the x-axis represents the density of data, which is fairly sparse before Year.sc $= -4$ (1910), but denser after Year.sc $= -2$ (1957). The y-axis is the prediction of verb number on a scale of 0 to 1, where predictions below 0.5 yield Singular, while predictions above 0.5 yield Plural. The line shows the prediction of the model, with a narrow ribbon showing the 95% confidence interval for the prediction. For nearly its entire range, Year.sc is associated with the choice of Singular, crossing over to Plural only toward its maximum value, where the prediction of Plural stands at 0.53.

5.3 The effect of animacy

Figure 3 plots the predicted values for Animacy. As in Fig. 2, the y-axis shows the prediction of number. A dotted line is included to show 50%, below which the model predicts more Singular and above which the model predicts more Plural. The width of the bars indicates the relative quantity of data, and the whiskers show the 95% confidence interval. The animate value was observed in more of the data and gives a higher prediction for the probability of Plural, but both animate and inanimate reflect that Singular is more expected than Plural.

5.4 The effect of word order

The effect of Word order is visualized in Fig. 4, with similar parameters as Fig. 3. We see a clear difference with VS word order preferring Singular, while SV word order prefers Plural.

5.5 The effect of verb lemma

Our model measures and ranks the preferences of all 2,311 verb lemmas in our data. Many of these lemmas are of low frequency, and therefore the confidence interval is wide, meaning

Year.sc effect plot
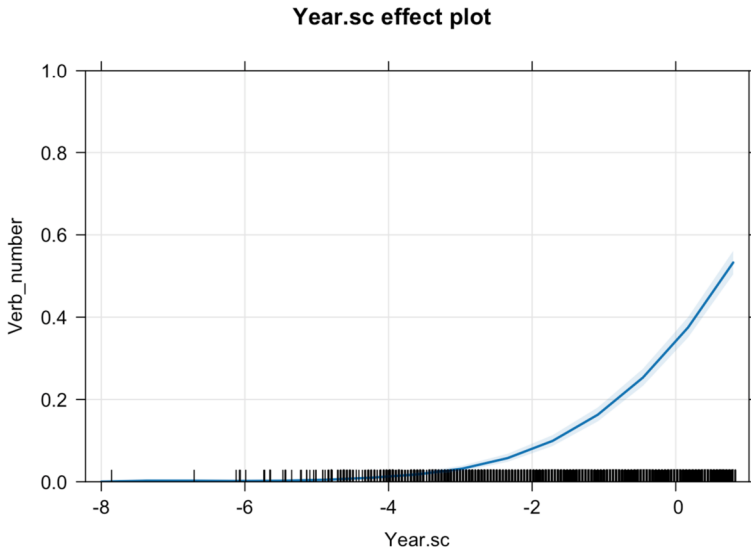
Fig. 2 The predicted values of Year.sc for Verb_number. See Table 4 to interpret Year.sc and note that the y-axis reports predictions of Plural, ranging from zero to 1 (100%)
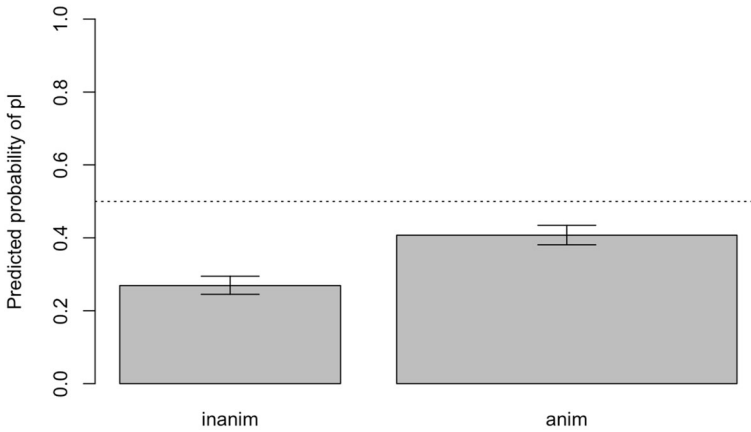
Fig. 3 The predicted values for Animacy

that the measurement is uncertain. Table 9 presents data only from verb lemmas that occur twenty or more times in our data, and therefore give more reliable estimates of preference. The left-hand part of Table 9 shows the ten lemmas with strong preference for Singular verb forms, ranked from the strongest to progressively less strong. The top two verbs are shaded to indicate that for these verbs 100% of examples have Singular verb forms. In the case of *минути* 'pass' this preference is stronger and more certain because we have 307 observations of this verb as compared to only 24 for *відбутися* 'happen'. The right-hand part of Table 9 shows the ten lemmas at the other extreme of the scale, those that most prefer Plural, and again the verb for which 100% of observations are Plural is shaded. The certainty
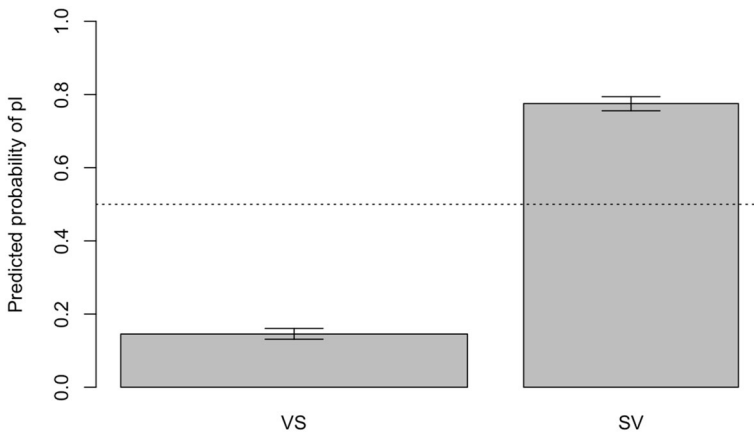
Fig. 4 The predicted values for Word order

Table 9 Verb lemmas that appear in twenty or more examples and are ranked by the model at the extremes of a distribution from preference for Singular to preference for Plural. Shading indicates verbs with 100% of observations of only Singular or only Plural

Top 10 high-frequency verbs with strongest preference for Singular		Top 10 high-frequency verbs with strongest preference for Plural	
минути 'pass'	307	*розкритикувати* 'criticize'	21
відбутися 'happen'	24	*засудити* 'condemn'	20
надійти 'tun up'	204	*ввести* 'bring in, introduce'	21
виникнути 'arise'	274	*заявити* 'declare'	43
виникати 'arise'	102	*очікувати* 'expect'	26
накопичитися 'accumulate'	140	*відзначати* 'mark, notice'	27
статися 'happen'	107	*відчути* 'feel'	30
надходити 'come'	84	*висловити* 'express'	50
полягти 'perish'	44	*отримати* 'get'	129
назбиратися 'gather, come together'	52	*пережити* 'experience'	23

of this preference is clearest for *отримати* 'get' with 129 examples, as opposed to *засудити* 'condemn' with only 20 examples.

It should be kept in mind that the model ranks verbs while taking into account also their behavior on other parameters. In other words, this is not simply a ranking of the percentage of Singular observations for each verb. There are many verbs other than *минути* 'pass' and *відбутися* 'happen' that occur exclusively with Singular forms in our data (see Table 10), but are ranked lower in individual preference by the model because of other factors, such as VS word order and inanimate nouns.

The density of observations is much stronger on the Singular end of this distribution: all ten verbs ranked from strongest at the Singular end were observed more than twenty times, with an average of 134 observations. In order to find ten verbs with twenty or more observations at the Plural end, it was necessary to look through 35 verbs, and the average number of observations for the verbs in the right-hand part of Table 9 is 39.

Table 10 Groups of verbs that strongly attract Singular forms in the bahato construction according to our statistical model

Group	Verb	N exx	% Sg	% VS	% inanim
PASSAGE OF TIME	минути 'pass'	307	100%	87%	100%
	пройти 'pass'	203	78%	81%	68%
HAPPEN	відбуватися 'happen'	307	100%	87%	100%
	відбутися 'happen'	24	100%	4%	100%
	виникнути 'arise'	274	99%	89%	100%
	виникати 'arise'	102	99%	81%	98%
	статися 'happen'	107	99%	78%	99%
EXIST	точитися 'exist'	46	100%	91%	100%
	існувати 'exist'	62	100%	95%	94%
	лежати 'exist'	53	98%	81%	62%
	бути 'be'	8257	93%	83%	54%
	знайтися 'exist'	57	100%	95%	19%
ACCUMULATE	нагромадитися 'accumulate'	23	100%	83%	100%
	накопичитися 'accumulate'	140	99%	89%	100%
	надійти 'show up'	204	100%	90%	97%
	надходити 'show up'	84	100%	89%	96%
	назбиратися 'gather'	52	100%	83%	94%
	постати 'appear'	34	97%	74%	97%
	з'явитися 'appear'	658	98%	97%	54%
	наїхати 'arrive'	33	100%	85%	12%
	розвестися 'accumulate'	40	100%	85%	10%
	зібратися 'gather'	467	95%	91%	6%
	збиратися 'gather'	66	97%	83%	3%
DIE	вимерти 'die out'	16	94%	31%	12%
	загинути 'die'	521	82%	65%	4%
	гинути 'die'	28	82%	54%	7%
	померти 'die'	96	77%	28%	3%
	полягти 'perish'	44	95%	50%	0%

These two groups of verbs differ in several relevant ways. The verbs that prefer Singular are all intransitive, describing things that merely happen; whereas the verbs that prefer Plural are all transitive, describing volitional actions. Although definiteness and topicality could not be included as factors in this study, their effect surfaces in these preferences among verb lemmas. The verbs at the extremes have corresponding preferences for Word order and Animacy. The top ten verbs preferring Singular tend to occur with VS Word order and, with inanimate nouns (except *полягти* 'perish'), whereas the top ten verbs preferring Plural tend to occur with SV Word order and/or animate nouns. We explore the behavior of individual verbs in more detail in Sect. 6.

6 Case studies

Close examination of the behavior of individual verb lemmas reveals several more specific versions of the *bahato* construction that could be called "micro-constructions" (a term introduced by Traugott, 2008a and 2008b). The verbs in our data that are associated mostly or exclusively with Plural forms tend to describe volitional actions that human beings undertake as individuals, however these verbs are thinly dispersed and disparate: they are of low frequency and hard to resolve further into semantic groups. By contrast, the verbs that strongly prefer Singular forms in the *bahato* construction tend to be of higher frequency and can be sorted into five groups, as shown in Table 10. The rightmost columns in Table 10 show the total number of examples of each verb in our dataset, and of those the percentage with a Singular verb form, the percentage with VS word order, and the percentage with an inanimate noun. For example, in the PASSAGE OF TIME group, the verb *минути* 'pass' appears in 307 examples, 100% of those have a Singular verb form, 87% are in VS order, and 100% of the nouns are inanimate. Note that the twenty-eight verbs in Table 10 are found in 12,305 examples, constituting 43% of our total dataset.

Within each semantic group in Table 10, verbs are sorted roughly according to how well their data conform to the pattern of Singular, VS, and inanimate. In addition, verbs that belong to Perfective-Imperfective pairs such as *відбутися-відбуватися* 'happen' are presented together. We examine each group in turn in more detail.

6.1 Verbs signaling passage of time

The primary verb in this group is *минути* 'pass', which signals passage of time in the *bahato* construction, with examples like (11). Nouns that name units of time such as *день* 'day', *місяць* 'month', *рік* 'year', are frequently found in this micro-construction. This micro-construction can be compared to the "measurement construction" in Russian identified by Nesset and Janda (2023, pp. 23–24) that likewise prefers Singular verb forms with VS word order and inanimate nouns referring to durations. It is likely that this micro-construction contributes to the observation by Hryhor'jev et al. (2005) that time expressions prefer Singular verb forms (see Sect. 3.1).

(11)   І хоча **минуло багато років**, вона про цю подію розповідала з великим
        хвилюванням. (Інтернет-газета «Високий замок», 2003)
        'And although **many years have passed**, she talked about this event with great excitement.'

The verb *пройти* 'pass' often behaves as a synonym for *минути* 'pass', but can also reference physical movement, in which case it is not participating in this micro-construction.

6.2 Verbs signaling happen

This micro-construction is expressed with the two aspectual pairs *відбутися* / *відбуватися* 'happen' and *виникнути* / *виникати* 'arise', and the verb *статися* 'happen'. All five verbs appear exclusively or almost exclusively with Singular forms and inanimate nouns. There is more variation in Word order, with four verbs ranging from 78% to 89% VS, but only 4% VS for the Perfective *відбутися* 'happen', itself rather infrequent. Example (12-a) illustrates the most common Word order (VS) for the Imperfective *відбуватися* 'happen', and (12-b) illustrates the opposite most common Word order for the corresponding Perfective *відбутися*. Together these two examples also show two of the most common nouns for all three 'happen' verbs: *подія* 'event', *зміна* 'change'.

(12)   a.   того дня у столиці **відбувалося багато подій** (Інтернет-газета «Україна молода», 2011)
'**many events took place** that day in the capital'

    b.   Відтоді **багато змін відбулося** в житті актриси і мами. (Інтернет-газета «Дзеркало тижня», 2017)
'Since then, **many changes have taken place** in the life of the actress and mother.'

    c.   Принаймні в автора цих рядків до режисера **виникло багато запитань.** (Інтернет-газета «Україна молода», 2012)
'The author of these lines at least **had many questions** for the director.' [literally: there arose many questions]

Example (12-c) illustrates the behavior of the Perfective *виникнути* 'arise' with one of the nouns typical for this aspectual pair; other common nouns for this pair are *дискусія* 'discussion' and *проблема* 'problem'.

6.3 Verbs signaling exist

The EXIST micro-construction is closely related to HAPPEN, especially in the case of the verb *точитися* 'exist (be circulating)' that combines with nouns similar to those found with *виникнути / виникати* 'arise', as in example (13-a). Other verbs in this group readily admit animate nouns, especially *лежати* 'exist (lie)', *бути* 'be', and *знайтися* 'exist (be found)'. While *бути* 'be' predominates in terms of frequency, it does not attract any particular types of nouns. The verb *лежати* 'exist (lie)' is mostly used in reference to people who are sick, hurt, or dead (corpses), as in (13-b), and the verb *знайтися* 'exist (be found)' is mostly used to describe situations where people are needed as participants or supporters, as in (13-c).

(13)   a.   Потім **точилося багато розмов** про матеріальну підтримку своєї збірної від вірменської діаспори. (Інтернет-газета «Високий замок», 2003)
'Then **there were many conversations** about the material support of the national team from the Armenian diaspora.'

    b.   У відділенні **лежало багато хворих** (Онлайн-ЗМІ «UNIAN.NET», 2020)
'**There were many patients** in the ward'

    c.   І що ви думаєте, **знайшлося багато охочих**? (Онлайн-ЗМІ «LB.ua», 2021)
'And what do you think, **were there many volunteers**?'

In the case of both *лежати* 'exist (lie)' and *знайтися* 'exist (be found)', when people are referred to, they have limited or subjugated agency. People in this micro-construction are not highly individuated, motivating the use of Singular verb forms.

6.4 Verbs signaling accumulate

This micro-construction is similar to EXIST, with the additional meaning of large quantity, often tinged with a negative connotation and emphasized by the meanings of prefixes that overlap somewhat with the meaning of *багато*. Like EXIST, this group of verbs includes both verbs that combine almost exclusively with inanimate nouns, as well as verbs that admit and even prefer animate nouns. Problems, questions, and complaints are often the things that appear in large quantities with the verbs preferring inanimate nouns, as we see in examples (14-a)–(14-c).

(14)   a.   столиця вже кілька років перебуває без міського голови і в місті **накопичилось багато проблем**. (Онлайн-ЗМІ «UNIAN.NET», 2014)
'the capital has been without a mayor for several years and **many problems have accumulated** in the city.'

       b.   До столичного відділення комітету **надійшло багато скарг**. (Онлайн-ЗМІ «LB.ua», 2021)
'**There appeared many complaints** addressed to the capital branch of the committee.'

       c.   Дуже вже **багато питань назбиралося** в нього до родича. (Валентин Тарасов «В темряві сонця», 2012)
'He **had a lot of questions** to his relative.'

The most numerous verb in this micro-construction, *з'явитися* 'appear', bridges the two subgroups, with nearly equal likelihood to appear with an inanimate noun referring to a document, as in (15-a) or with animate nouns referring to people close to an individual, as in (15-b).

(15)   a.   останнім часом **з'явилось багато повідомлень** про контрабанду нафти ІДІЛ до Туреччини. (Інтернет-газета «Європейська правда», 2015)
'recently **there have been many reports** about ISIS oil smuggling to Turkey.'

       b.   Там у нього **з'явилося багато друзів**, які його підтримували (Газета «7 днів», 2017)
'There **he had many friends** who supported him'

The verb *розвестися* 'accumulate' implies an undesirable accumulation, usually of vermin or people, that has perhaps gotten out of control, as illustrated in (16-a)–(16-b).

(16)   a.   Відвідувачі Гідропарку скаржаться, що там **розвелося багато щурів**. (Онлайн-ЗМІ «UNIAN.NET», 2013)
'Visitors to the Hydropark complain that **there are a lot of rats**.'

       b.   Дуже **багато опозиціонерів розвелось** в останній час, неправда, так? (Стенограми засідань Верховної Ради України, 2019)
'**There has been a great quantity of opposition members** lately, right?'

Other verbs that combine with animate nouns most often refer to people gathering at events, as in (17).

(17)   Попри дощову погоду, на щемливий захід із нагоди її відкриття зібралося багато людей. (Інтернет-газета «Україна молода», 2021)
'Despite the rainy weather, **many people gathered** at the poignant event on the occasion of the opening.'

6.5 Verbs signaling `die'

The majority of examples of this micro-construction involve people dying, as in (18-a). The few examples with inanimate nouns involve personifications and metonymies, as we see with the banks and types (of organisms) in (18-b)–(18-c).

(18)   a.   І **багато хлопців загинуло** на моїх очах. (Сергій Сущенко, «Спогади», 2018)
'And **many boys died** before my eyes.'

b. Дуже **багато банків померли** насправді ще 2009 року. (Журнал «НВ»,
    2015) – example with the inanimate noun
    '**A lot of banks died** back in 2009.'

c. в переломні моменти історії Землі масово **вимерло багато видів**
    організмів. Журнал «Наука і суспільство», 1987)
    '**many types** of organisms **died out** en masse during turning points in the
    Earth's history.'

6.6 Summary of micro-constructions

Verbs associated with five micro-constructions that strongly prefer Singular forms account
for 43% of the data in our study. Most verbs in these micro-constructions prefer VS Word
order, and differences are most palpable among the nouns they combine with. Two micro-
constructions are quite distinct: PASSAGE OF TIME, with nouns referring to durations, and DIE,
dominated by human beings. The remaining three micro-constructions are related to each
other. While HAPPEN combines almost exclusively with inanimate nouns, EXIST and ACCUMU-
LATE combine with both types of nouns: inanimate nouns refer primarily to communications
and problems, while animates are usually people who are identified in relation to a given
individual, idea, or event.

7 Conclusion

With respect to our research questions stated in Sect. 1, we have discovered the following:

- In the *bahato* construction the relative frequency in our data is 69.1% Singular vs. 30.9%
  Plural.
- This relative frequency has changed over time: Plural has increased and recently overtaken
  Singular.
- SV word order and animate nouns prefer Plural.
- Individual verbs have their own preferences for verb number.
- A handful of micro-constructions that prefer Singular play a large role.

Given what is known about the alternative construals of plurals as mass vs. multiplex
(Johnson, 1987; Lakoff, 1987; Langacker, 2008) and the existence of number variation based
on clashes between semantics and syntax (Corbett, 2000), it is perhaps unsurprising that
Ukrainian *багато* admits both Singular and Plural agreement. It also makes sense that both
Word order and Animacy should play a role in number agreement. Entities high on the Ani-
macy scale such as human beings and animals are more likely to be perceived as individuals
and are also more likely to appear as agentive subjects preceding the verb. When human
beings appear in micro-constructions with Singular verb forms, their agency is often attenu-
ated by death or disease or they are subordinated to another person or event, reducing their
individuation.

The rapid rise in frequency of Plural verb number over the past century is striking, par-
ticularly in comparison with the findings of Nesset and Janda (2023), where no appreciable
change in Singular vs. Plural verb number is observed in Russian over a period of more
than two centuries. It is hard to speculate on what might account for this difference between
Russian and Ukrainian. The Russian study was more comprehensive in terms of quantifiers,
including all numerals except '1'. It might be the case that various numerals in Russian are
undergoing shifts in opposing directions regarding preference for Singular vs. Plural verb

number, and that these shifts somehow cancel each other out, though this seems unlikely. It is important to note that Ukrainian *багато* has not one, but two translations in Russian: 1) the indefinite numeral *много* 'many' (included in the Nesset and Janda study) that has a strong preference for Singular, and 2) the adjective *многие* 'many' (usually referring to people, not included in the Nesset and Janda study) that requires Plural verb agreement. In other words, the fact that Ukrainian *багато* expresses the meanings of both the predominantly Singular Russian *много* and the exclusively Plural *многие* could account for the higher incidence of Plural with *багато*. However, this does not explain the fact that the frequency of Plural is increasing over time.

Declarations

Competing Interests　On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Antonenko-Davidovič, B. (1979). *Jak my hovorymo*. Smoloskyp. pp. 136.

Arpolenko, H. P., Horodens'ka, K. H., & Ščerbatjuk, H. X. (Eds.) (1980). *Čyslivnyk Ukrajins'koji Movy*. Naukova dumka.

Arsenijević, B. (2016) Srpski izrazi *mnogo* i *mnogi* u svetlu teorije slabih i jakih kvantifikatora. In B. Arsenijević & S. Halupka-Rešetar (Eds.), *Srpski Jezik u Savremenoj Lingvističkoj Teoriji*, 239–258. Univerzitet u Nišu.

Boldyrjev, R. V. (Ed.) (1982). *A–H.* = Vol. 1 of O. S. Mel'nyčuk (ed.) *Etymologičyj slovnyk ukrajinskoji movy*. AN URSR.

Buck, C. D. (1949). *A dictionary of selected synonyms in the principal indo-European languages*. University of Chicago Press.

Bybee, J. (2015). *Language change*. Cambridge University Press.

Corbett, G. G. (2000). *Number*. Cambridge University Press.

Croft, W. (2000). *Explaining language change. An evolutionary approach*. Longman.

Croft, W. (2001). *Radical construction grammar*. Oxford University Press.

Dancygier, B. (2017). Viewpoint phenomena in constructions and discourse. *Glossa: A Journal of General Linguistics*, *2*(1), 37.

Dąbrowska, E. & Divjak, D. (Eds.) (2015). *Handbook of cognitive linguistics*. De Gruyter Mouton.

Fillmore, C. J., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, *64*(3), 501–538.

Geeraerts, D. & Cuyckens, H. (Eds.) (2007). *The Oxford handbook of cognitive linguistics*. Oxford University Press.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work*. Oxford University Press.

Goldberg, A. E. (2013). Constructionist approaches. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford handbook of construction grammar*, 14–31, Oxford University Press.

Gries, S. Th. (2021). *Statistics for linguistics with R*. De Gruyter.

Hryhor'jev, O. M., Doloman, S. Je., Lysenko, Ju. V., Makarenko, Je. A., Mamalyta, A. I., Marčuk, N. J., Pazjak, O. M., Ponomariv, O. D., Rizun, V. V., Furduj, M. I., & Ševčnko, L. Ju. (Eds.) (2005). *Sučasna ukrajins'ka mova*. (3rd ed.). Lybid'.

Janda, L. A., & Palii, Y. (2024). Replication Data for: Understanding 'many' through the lens of Ukrainian *багато*. https://doi.org/10.18710/Y7VGQE.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. University of Chicago Press.

Kibrik, A. A. (2019). Rethinking agreement: Cognition-to-form mapping. *Cognitive Linguistics*, *30*(1), 37–83.

Kurylo, O. (1942). Uvahy do sučasnoji ukrajins'koji literaturnoji movy. *Krakiv: Ukrajins'ke Vydavnyctvo.*, 56–58.

Lakoff, G. (1987). *Women, fire, and dangerous things*. The University of Chicago Press.

Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.

Milosavljević, S. (2018). Kongruencija glagolskog predikata sa subjaktskim sintagmama konstituisanim osnovnim brojem. *Zbornik Matice Srpske za Filologiju i Lingvistiku*, 43–56.

Nesset, T., & Janda, L. A. (2023). A network of allostructions: quantified subject constructions in Russian. *Cognitive Linguistics*, *34*(1), 67–97. https://doi.org/10.1515/cog-2021-0117.

Shvedova, M., von Waldenfels, R., Yarygin, S., Rysin, A., Starko, V., Nikolajenko, T. et al. (2017–2024). *GRAC: General Regionally Annotated Corpus of Ukrainian*. Electronic resource: Kyiv, Lviv, Jena. Available at uacorpus.org.

Stadler, M., & Kruse, P. (1995). The function of meaning in cognitive order formation. In P. Kruse & M. Stadler (Eds.), *Ambiguity in mind and nature* (pp. 13–21). Springer.

Traugott, E. C. (2008). The grammaticalization of NP of NP patterns. In A. Bergs & G. Diewald (Eds.), *Constructions and language change*, 23–45. Mouton de Gruyter.

Traugott, E. C. (2008). Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English. In R. Eckhardt, G. Jager, & T. Veenstra (Eds.), *Variation, selection, development: Probing the evolutionary model of language change*, 219–250. Mouton de Gruyter.

Wierzbicka, A. (1988). *The semantics of grammar*. John Benjamins.