Exhaustivity and homogeneity effects with distributive-share markers

Bosnić, Ana; Demirdache, Hamida; Spenader, Jennifer

Published in:
Syntax

DOI:
10.1111/synt.12223

DOI: 10.1111/synt.12223

REGULAR ARTICLE

Exhaustivity and homogeneity effects with distributive-share markers: Experimental evidence from Serbian *po*

Ana Bosnić[1,2] | **Hamida Demirdache**[2] | **Jennifer Spenader**[1]

[1]University of Groningen

[2]University of Nantes/French National Center for Scientific Research

Correspondence
Ana Bosnić, University College Groningen, Hoendiepskade 23/24, 9718 BG Groningen, the Netherlands
Email: a.bosnic@rug.nl

Abstract
There are two competing approaches to the semantics of distributive-share markers: they are either universal distributive quantifiers over events or are merely event-plurality markers. To address this debate, we present new conclusions based on novel experiments with Serbian transitive sentences in which the distributive-share marker *po* was attached to the direct object. The first two experiments investigated *exhaustivity* effects in transitive sentences with *po*, while the third experiment probed *homogeneity* effects across three types of negative transitive sentences: with *po* marking the object, with the distributive-key quantifier *svaki* ('every') in subject position, and with neither. If *po* is a universal quantifier, then it should enforce exhaustive distribution over a distributive key and remove homogeneity effects in negative sentences with a definite subject. If instead *po* is an event-plurality marker with no universal quantificational force, then it should neither enforce exhaustive distribution nor remove homogeneity effects in negative sentences with a definite subject. We conclude that there are two populations of Serbian speakers with systematic patterns of interpretation: one population interprets *po* as a universal quantifier and one

population interprets *po* as an event-plurality marker. We conjecture that this population split might reflect an ongoing diachronic change in the semantic import of the distributive-share marker *po*.

KEYWORDS
distributive-share markers, event plurality, exhaustivity, homogeneity, maximality

1 | INTRODUCTION

We experimentally tested two competing approaches to distributive-share markers by exploring exhaustivity effects and homogeneity effects in Serbian transitive sentences containing the distributive-share marker *po*. This work builds on recent investigations of distributive-share markers in Serbian and Korean presented in Bosnić, Spenader & Demirdache 2020 as well as Knežević & Demirdache 2017, 2018. While there is one line of research that argues in favor of analyzing distributive-share markers as universal quantifiers distributing over event aspects (Zimmermann 2002, Balusu 2006, Balusu & Jayaseelan 2013, Champollion 2016b, Bosnić, Spenader & Demirdache 2020), another contends that distributive-share markers are event-plurality markers and not universal quantifiers (Cable 2014, Knežević 2015, Knežević & Demirdache 2017, 2018). Recent experimental research with Serbian and Korean (Bosnić, Spenader & Demirdache 2020) found that with intransitive verbs, distributive-share markers have exhaustivity requirements that support a universal-quantifier analysis.

The current study builds on this earlier work, turning from intransitives to transitives. We report here the findings of three experiments designed to empirically assess the predictions of the two competing theories. To the best of our knowledge this is the first experimental work investigating either exhaustivity effects or homogeneity effects for transitive sentences with distributive-share markers, whether in Serbian or any other language.[1]

Experiments 1 and 2 were designed to test for *exhaustivity* effects in transitive sentences with the distributive-share marker *po*. If *po* is a universal quantifier, then it should enforce exhaustive distribution over a distributive key. If instead *po* is an event-plurality marker with no universal quantificational force, then it should allow nonexhaustive distribution. Experiment 3 (building on Križ & Chemla 2015) was designed to test for *homogeneity* effects across three types of negative transitive sentences: with *po* marking the object (and a definite plural in subject position), with the distributive-key quantifier *svaki* ('every') in subject position, and with neither marker (just the definite plural subject). If *po* is a universal quantifier, then it should remove homogeneity effects in negative sentences with a definite subject. If instead *po* is an event-plurality marker with no universal quantificational force, then it should not remove homogeneity effects in negative sentences with a definite subject.

Experiment 1 revealed, and experiment 3 confirmed, the existence of two populations with systematic and distinct response patterns: (i) speakers who interpret *po* as a universal

[1] More generally, there are only a handful of experimental studies of distributive-share markers (mostly acquisition studies; É. Kiss et al. 2013, Knežević 2015, É. Kiss & Zétényi 2018, Knežević & Demirdache 2017, 2018, and the already-mentioned Bosnić, Spenader & Demirdache 2020).

quantifier over events that allows distribution over either atomic entities or nonatomic entities, whether the latter are entities bigger than atoms (pluralities) or entities from noncount domains (time/space); and (ii) speakers who interpret *po* as an event-plurality marker. We speculate that the population split discovered here might be evidence for an ongoing diachronic process of semantic weakening of *po*, from a distributor with universal force to a marker of event plurality.

The article is organized as follows. Section 2 sets the theoretical background by laying out the phenomenon of distributive-share marking, the properties of the Serbian marker *po*, the competing theoretical approaches, and our previous experimental findings for intransitive sentences in Serbian (Bosnić, Spenader & Demirdache 2020). Section 3 is devoted to experiment 1, which probed exhaustivity effects in transitive sentences with *po* marking the direct object. Section 4 presents a follow-up to the exhaustivity experiment, testing whether controlling for the definiteness of bare plural subjects in *po* sentences could explain away the population split uncovered in experiment 1. Section 5 is devoted to experiment 3, which probed homogeneity effects in negative transitive sentences with *po*, with the universal quantifier *svaki*, and with neither. Section 6 recapitulates our empirical and theoretical findings, addressing challenges and open issues, and outlines future research directions.

2 | BACKGROUND

2.1 | Two theoretical approaches to distributive-share markers

Let us start with English: as illustrated in (1a), distributive markers such as *each* in English overtly force distributive readings of ambiguous sentences such as (1b).

(1) a. **Each** boy is holding two balloons.
→ Two balloons per boy
b. The boys are holding two balloons.
→ Two balloons in total (or two balloons per boy as a less preferred option)

For a distributive reading to hold, there has to be a pairing of two arguments, the *distributive key* and the *distributive share* (in the terminology of Gil 1982, Choe 1987, Gil 1995). The distributive key is the plural argument denoting the set over which distribution is taking place (in (1a), it is *boy*, the restriction of the universal quantifier), while the distributive share is the argument denoting what is being distributed (*two balloons*). In some languages, the distributive-key argument need not be explicit: distribution may take place over covert arguments, such as events and event aspects (space or time) (Gil 1995, Schwarzschild 1996, Zimmermann 2002, Cabredo Hofherr & Laca 2012, Champollion 2016a, 2016b).

Crosslinguistically, distributive markers can (morpho)syntactically attach either to the distributive key (English *each*, Dutch *elke*, Spanish *cada*, among others) or to the distributive-share argument (Korean *-ssik*, Japanese *-zutsu*, *po* in Slavic languages, German *jeweils*, Hungarian number reduplication, among others).[2] Thus compare (1a) with the Serbian and Korean sentences in (2). Unlike in (1a), the distributive markers in (2) syntactically attach to the distributive-share argument in the sentence (*two balloons*/*dva balona/phwungsenul twukay*), and the distributive key is left unmarked.

(2)	a.	Dečac-i	drž-e	[**po**	dva	balon-a].		Serbian
		boy-PL.NOM	hold-PRS.3PL	DISTR	two	balloon-PAUCAL.ACC		
		'(The) boys are holding DISTR two balloons.'						
	b.	Sonyen-tul-i	[phwungsen-ul	twu-kay-**ssik**]	tul-ko	iss-ta.		Korean
		boy-PL-NOM	balloon-ACC	two-CLF-DISTR	hold-PROG	be-DECL		
		'(The) boys are holding DISTR two balloons.'						

Distributive-share markers differ not only syntactically from distributive-key markers but also semantically. The general consensus is that distributive-share markers such as *po* and *-ssik* offer a broader spectrum of possible interpretations, yielding distribution over individuals but also allowing distribution over events (Lasersohn 1998, Oh 2006). Thus, the sentences in (2), for instance, straightforwardly allow a *participant-* or *individual-distributive* reading where sets of two balloons are distributed over boys (partitioned atomically: i.e., each boy is individually relevant); but crucially, unlike (1a), they also allow *event-distributive readings* where the event of balloon holding is broken down/partitioned into a contextually determined number of subevents involving (at least) one boy holding two balloons. These (sub)events of boys carrying sets of two balloons are partitioned/distributed temporally or spatially. On the temporal event-distributive reading, boys are carrying, individually or together, sets of two balloons at different times. On the spatial event-distributive reading, boys are carrying two balloons individually or together at different locations but at the same time. While the temporal reading allows the same or different boys to be carrying the same or different sets of two balloons on each occasion, the spatial reading requires the balloon-holding events to take place simultaneously at different locations, and hence there will have to be different boys per event (Gil 1990, Oh 2006, Knežević 2015).

Distributive-key markers such as *each* are considered to be distributive universal quantifiers, and, according to Gil 1995, this status is a typological universal. As universal quantifiers, distributive-key markers require all members of the set interpreted as the distributive key to exhaustively participate in the described event; for example, for (1a), each boy in the contextually given set must participate in holding two balloons.

The issue, however, of whether or not distributive-share markers also involve universal quantification is a matter of debate in the literature. One line of research, going back to Gil 1995, argues in favor of universal quantification (but over events as well as individuals: Faller 2001, Zimmermann 2002, Balusu 2006, Balusu & Jayaseelan 2013, Champollion 2016b, Bosnić,

[2]There are also markers that form a syntactic constituent (at least on the surface) with the distributive share while semantically associating with the distributive key, such as binominal *each* in English, illustrated in (i), and *chacun* in French. As such, these have also been referred to in the literature as distributive-share markers.

(i) The boys are holding [two balloons **each**].

Note, however, that unlike other distributive-share markers (e.g., *po* in Serbian), binominal *each* does not allow event-distributive readings. Under Zimmermann 2002's classification, binominal *each* and *chacun* are determiner distance-distributive markers. See Safir & Stowell 1988, Oh 2001, Zimmermann 2002, Oh 2006, Champollion 2012, 2016b, and references therein for discussion.

Spenader & Demirdache 2020). The other line of research argues that distributive-share markers simply signal event plurality, suggesting a pluractional analysis (Lasersohn 1995, Matthewson 2000, Müller & Negrão 2012, Cable 2014, Knežević 2015, Cabredo Hofherr et al. 2018); alternatively, it is argued by McKercher & Kim 2000 for Korean -*ssik*, as well as Faller 2001 for Quechua -*nka*, that these markers are not associated with distributivity at all but are merely group-forming devices (serving to form multiple groups from the denotation of the noun they combine with).[3]

The main aim of the current article is to contribute further critical experimental data and arguments to this debate. Are distributive-share markers universal quantifiers or are they event-plurality markers?

We focus on the distributive-share marker *po* in Serbian, found in many Slavic languages and taken by many authors to be a universal quantifier over events (see Zimmermann 2002, Champollion 2012, Przepiórkowski 2015, Champollion 2016b, among others). Knežević 2015 rejects the universal-quantification analysis by claiming that *po* lacks a core property of universal distributive quantifiers: *po* does not require its distributive key to be exhaustively distributed over by members of the distributive share. To illustrate, consider the examples in (3) in the four contexts given in figure 1.

(3)	a.	The children are carrying one suitcase **each**.					English
	b.	Dec-a	nos-e	**po**	jedan	kofer-∅.	Serbian
		children-NOM	carry-PRS.3PL	DISTR	one	suitcase-ACC	
		‘(The) children are carrying DISTR one suitcase.’					

The English sentence in (3a) can only be used to describe the scenario that is atomic (involving individual, one-to-one pairing) and exhausted. The sentence is judged false on the other three scenarios. This is the case because *each* requires that its distributive key be partitioned into

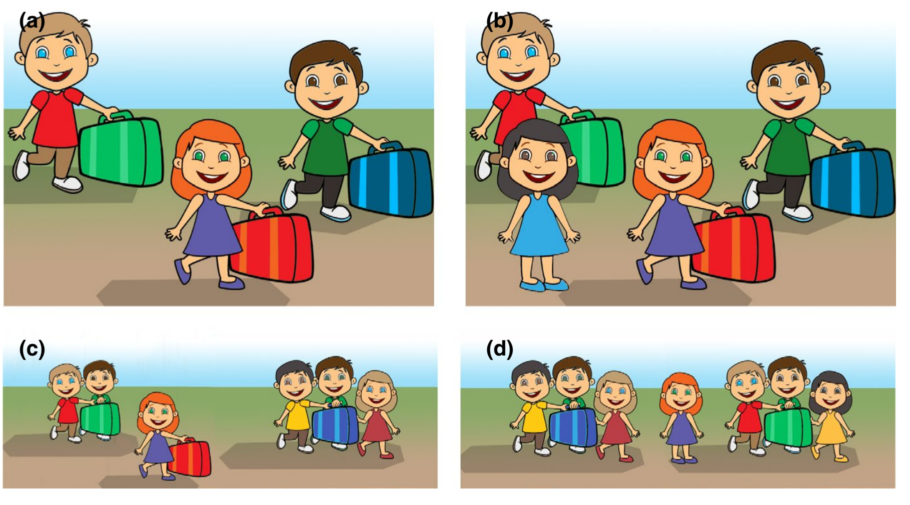

FIGURE 1 Visual representation of (non)atomic and (non)exhaustive readings for the sentences in (3). (a) Atomic and exhausted. (b) Atomic and nonexhausted. (c) Nonatomic and exhausted. (d) Nonatomic and nonexhausted

[3]Specifically, on McKercher & Kim’s group-forming analysis, the marker “combines with the meaning of a quantifier phrase and returns a higher type denotation which means Groups of *n*” (p. 243). Importantly, this analysis, just like the pluractional analysis (e.g., Knežević 2015), does not predict exhaustivity requirements.

individual atomic parts and also requires, as all universal quantifiers do, that distribution be exhaustive: each and every member/part of the distributive key must be distributed over (i.e., carry a suitcase). In contrast, the Serbian sentence (3b), according to Knežević's claims, can be used to describe all four scenarios, exhausted or not, atomic or nonatomic.[4] Knežević concludes that exhaustivity and atomicity are both irrelevant to the truth conditions of *po*. Crucially, *po* cannot be analyzed as involving universal quantification since it lacks this defining property of a universal quantifier: the meaning of a sentence with *po* does not enforce exhaustive/universal distribution over a distributive key (forming the restriction of the quantifier). Knežević's proposal is that *po* and similar distributive-share makers simply signal plurality of events (see also Matthewson 2000, Müller & Negrão 2012, Cable 2014).

Importantly, the core issue for determining whether the semantics of distributive-share marking involves universal quantification is exhaustivity, not atomicity. This is so (i) because even a universal quantifier such as *every* allows so-called partially distributive readings, unlike *each*, which obligatorily distributes down to individuals (e.g., *John photographed every /*each student but not separately*),[5] and (ii) because distributive-share markers also allow spatiotemporal distributive readings, and time is a continuous noncount domain and thus cannot be partitioned atomically. Positing a universal quantifier in the definition of distributivity operators serves to ensure exhaustivity, that is, distribution over every part of the plurality to which the operator applies (without necessarily ensuring atomicity: take for instance Schwarzschild 1996's cover-based distributivity operator Part, which allows nonatomic distribution).[6]

Knežević's analysis contrasts sharply with a universal-quantification analysis of distributive-share markers, proposed, for example, in Balusu 2006 and Balusu & Jayaseelan 2013 for reduplicated numerals in Telugu. These works take an opposite stance to that defended by Knežević for Serbian *po* by assuming that the semantics of distributivity with reduplicated numerals does indeed involve a relation between a distributive key and a distributive share, with exhaustivity as a diagnostic for identifying the distributive key. To quote Balusu & Jayaseelan:

---

[4]False situations for the *po* sentence in (3b) under Knežević's semantics would include a single collective event in which three children are together carrying one suitcase. Such a single collective situation would be judged false because the semantics of *po* requires that there be at least two such subevents of children carrying one suitcase (irrespective of whether the children act collectively or individually). Any situation where there is more than one suitcase per subevent is also predicted to be judged false since *po* requires that there be exactly *n* entities per subevent, where *n* stands for the numeral that *po* attaches to (here 'one').

[5]Note that Serbian has a universal quantifier *svaki* that corresponds to English *every* and can, moreover, co-occur with *po*. See Knežević & Demirdache 2018 for experimental evidence on *svaki–po* sentences. Adopting Knežević 2015's analysis of *po* as an event-plurality marker, Knežević & Demirdache argue that co-occurrence of *svaki* with *po* enforces exhaustivity and atomicity (yields a meaning akin to that of *each*). See section 6 for further discussion.

[6]Link's original atomic version of the D operator is given in (i), and Schwarzschild's cover-based Part version is in (ii). Both ensure universal/exhaustive distribution but differ in their "granularity," in Champollion 2016b's terms; that is, they differ with respect to "the size of the entities over which we distribute" (p. 4).

(i) $[\![\mathrm{D}]\!] = \lambda P_{et}\lambda x\forall y[y \leq x \wedge \mathrm{Atom}(y) \rightarrow P(y)]$
(Link 1987)

(ii) $[\![\mathrm{Part}_C]\!] = \lambda P_{et}\lambda x\forall y[y \leq x \wedge C(y) \rightarrow P(y)]$
(Schwarzschild 1996)

> Notice that the members of the distributive share need not be exhaustively used up when being distributed, whereas the members of the sorting key need to be exhaustively used up in being distributed over. This is a key difference between the distributive share and the sorting key that we will be making use of to diagnose whether a particular DP contributes a sorting key or is rather a distributive share that is distributed over some covert key.
>
> (Balusu & Jayaseelan 2013: 10)

This view assumes that a distributive operator is essentially a universal quantifier, with a distributive key, that is, the quantifier's restriction, and a distributive share, that is, the quantifier's scope.

These opposing claims call for experimental testing—all the more so because there are other, typologically diverse languages whose distributive-share markers appear to share the range of interpretations that *po* has: Japanese, Korean, Tlingit, Quechua, Telugu. This experimental testing is precisely what we sought to do in Bosnić, Spenader & Demirdache 2020 by conducting a series of picture-verification tasks with adult native speakers of Serbian and Korean. Our goal was to find out how these two theoretical approaches, universal quantification and event plurality, apply to empirical data in two different distributive-share languages. We present the findings in the next subsection.

2.2 | Experimentally testing exhaustivity: previous findings

In Bosnić, Spenader & Demirdache 2020 we experimentally probed *spatial* event distribution, using intransitive sentences in Serbian and Korean. Taking as our theoretical premise that a key diagnostic for universal quantification is the exhaustivity requirement on the distributive key, we sought to experimentally determine whether distribution over the members of different presumed distributive keys has to be exhaustive or not. Since the test sentences were intransitive, the distributive-share marker (*po* in Serbian, *-ssik* in Korean) would have to be attached to the only available DP in the sentence, the subject (see the examples below in (4) and (5)). This meant that the distributive key would have to be an implicit spatiotemporal argument of the verb and that the only available reading would be an event-distributive reading over covert spatiotemporal locations.

The materials were designed to check for exhaustivity requirements on different putative implicit spatial distributive keys, comparing the predictions of the universal-quantification and event-plurality accounts. Figure 2 shows four main conditions from two experiments; the corresponding test sentences are in (4) and (5). Figures were the three conditions under experiment 1 (53 Serbian and 26 Korean participants), and figure 2d was the crucial condition of the follow-up experiment 2 (31 Serbian and 30 Korean participants).

(4)	Pleš-e	[**po**	jedan	majmun-∅].		Serbian
	dance-PRS.3SG	DISTR	one	monkey-NOM		
	'DISTR one monkey is dancing (at different locations).'					
(5)	[Wenswungi-ka	han-mali-**ssik**]	chwum-ul	chwu-ko	iss-ta.	Korean
	monkey-NOM	one-CLF-DISTR	dance-ACC	dance-PROG	be-DECL	
	'DISTR one monkey is dancing (at different locations).'					

In experiment 1, designated spatial locations were experimentally set up as potential distributive keys. Pictures were presented in a frame story about photos that children took when they visited

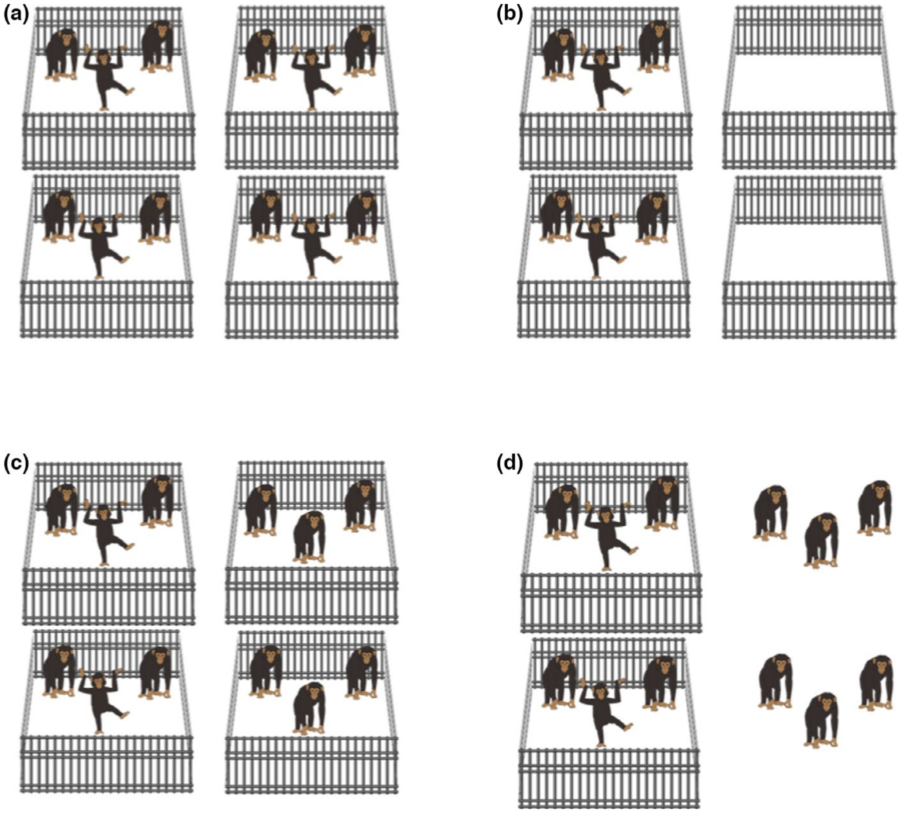

FIGURE 2 The four main conditions in the Bosnić, Spenader & Demirdache 2020 experiments, for the intransitive sentences in (4) and (5). (a) Exhausted spatial units. (b) Nonexhausted spatial units (empty cages). (c) Nonexhausted spatial units (nondancing monkeys). (d) Exhausted spatial units (nondancing monkeys outside cages)

a zoo. Each picture showed different animals, with spaces visually delimited by cages or caves. Although the visual context and preceding discourse context aimed to make the cages/caves relevant spatial locations and thus serve as the (implicit) distributive key, these spaces were never explicitly mentioned, thus forcing participants to determine the implicit distributive key themselves.

On an event-plurality account, there is no distributive key that must be exhausted. All that matters, for the truth conditions to be satisfied, is that there be at least two events of a monkey dancing. Thus, if participants consider *po* to be merely an event-plurality marker, they should say yes (i.e., the sentence can be used to describe the picture) to all of the pictures in figure 2 since they indeed all satisfy the condition that there be at least two events of a monkey dancing. According to the universal-quantification account, on the other hand, the visual input of the experimental stimuli gives three possibilities for a distributive key. The first possibility is that (i) salient spatial locations, namely cages, serve as the distributive key. The results for experiment 1 revealed that both figure 2a, the baseline condition where the presumed distributive key (cages) is exhausted since each cage contains a dancing monkey, and figure 2b, where the distributive key is nonexhausted since two of the cages do not contain monkeys, were overwhelmingly accepted. Crucially, however, figure 2c, where the distributive key is also nonexhausted but here because two of the cages contain nondancing monkeys, was rejected (the no responses were almost 90%). Rejection of figure 2c suggests that there is indeed some form of exhaustivity requirement after all. Experiment 2 was thus designed to check the two other possible distributive keys that the visual input provides, by adding figure 2d, a critical condition: are the participants exhausting (ii) groups/triplets of monkeys in cages or just (iii) groups/triplets of monkeys? The results—rejection of figure 2d about 75% of the time—clearly favored option (iii): each visually salient group of monkeys, whether it is caged or cageless, must be exhausted. Speakers generally judged the test sentence as true if every group of monkeys contained a dancing monkey. There is thus an exhaustivity requirement on groups of monkeys.

In Bosnić, Spenader & Demirdache 2020 we took these intriguing findings to provide evidence that the semantics of spatial event distribution involves a covert spatial distributive key. The groups/triplets of monkeys made salient in the visual context serve to divide the distributive key into chunks of space that must be exhaustively distributed over. This conclusion argues in favor of a universal-quantification analysis where *po* is a universal distributive quantifier that can distribute over atomic as well as nonatomic entities—whether the latter belong to noncount domains, such as time, or are entities bigger than atoms/singular individuals, namely groups/pluralities of atomic individuals.[7] We follow the literature in assuming that the ability to distribute over nonatomic domains— typically time and space but also groups, as argued in Bosnić, Spenader & Demirdache 2020—is characteristic of event-distributive operators.

To conclude, note that this proposal fits well with Champollion 2016b's semantic parameters for distributivity operators, according to which variation can be along a parameter of granularity specifying the "size of the entities over which we distribute: for example, atoms or amounts of space or time" (p. 4). If the granularity parameter is set to atoms, as is the case for *each*, then distribution over spatiotemporal dimensions is excluded: since time is a continuous and noncount dimension, there are no atoms to distribute over. If granularity is left unspecified, as would be the case for *po*, not only will distribution over noncount dimensions like time be possible but also, as assumed in Bosnić, Spenader & Demirdache 2020, over individuals that are bigger than (but made up of) atoms/singular individuals, that is, groups or plural individuals.

3 | EXPERIMENT 1: OVERT SUBJECT ARGUMENTS AS DISTRIBUTIVE KEYS

The empirical evidence adduced so far that the semantics of distributive-share markers involves a distributive key that must be exhausted comes from the interpretation of intransitive sentences alone. As mentioned, however, intransitive sentences such as (4)/(5) are special in that their sole argument has to be interpreted as the distributive share (since it is the only argument to which *po* can attach), which means in turn that the distributive key will have to be a covert argument that has to be inferred from the linguistic and visual input.

But the distributive key can also be an overt selected argument of the verb, as long as there is another argument available for *po* to attach to and mark as the distributive share. In fact, *po* can attach to any argument in a Serbian sentence. Thus, in a simple transitive sentence, *po* can, in

[7]For details on the experimental design, analysis of the results, and theoretical implications of the findings, see Bosnić, Spenader & Demirdache 2020. In a nutshell, our account of the interpretation of *po* in (4)/(5) builds on Zimmermann 2002's analysis of the German distance-distributive quantifier *jeweils*: *po* is argued to be a locative preposition with universal quantificational force with a meaning akin to *per*, establishing a distributive relation between individual events (of monkeys dancing) and plural nonatomic individuals (triplets of monkeys) that serve as the distributive key, requiring that there be one dancing monkey per each triplet of monkeys.

principle, be on the subject, object, or both arguments. The same appears to hold for distributive-share markers in Japanese, Korean, Telugu, Quechua, or Tlingit.

We restrict our attention here to transitive sentences where *po* attaches to the direct object:

(6) Dva čovek-a nos-e **po** tri kofer-a.
two man-PAUCAL carry-PRS.3PL DISTR three suitcase-PAUCAL.ACC
'(The) two men are carrying DISTR three suitcases.'

There are various possible readings for this sentence (out of context). The most conventional reading is the so-called participant-distributive reading that distributive-key markers such as *each* yield, paraphrasable as '(The) two men each carried a set of three suitcases'. Distribution on this reading is over atomic/singular individuals.

The sentence in (6) also allows spatiotemporal distributive readings, where multiple events of two men carrying three suitcases are distributed over different spatiotemporal locations. If distribution is temporal, the only constraint is that the suitcase-carrying events do not happen at the same time. This yields scenarios where either the same set or different sets of two men are carrying the same or different sets of three suitcases on more than one occasion. If, however, distribution is spatial, then the suitcase-carrying events are happening simultaneously in more than one place. These readings come for free, assuming the granularity parameter for *po* is left unspecified, as mentioned in section 2.2: distribution can be over entities of different sizes, such as atoms or amounts of space and time.

The critical question we address in this article is whether or not the findings for intransitive sentences extend to transitive sentences such as (6), where the object is marked as the distributive share by *po*, thus allowing the unmarked subject to serve as an overt distributive key and yielding distribution over singular individuals. If indeed the semantics of *po* does involve universal quantification, then again we should find exhaustivity effects, just as with intransitive sentences, but this time on the plural subject argument serving as the distributive key, that is, as the overt restriction of the universal quantifier.

Recall, however, the conclusion from the previous findings: *po* can distribute over either atomic individuals or nonatomic individuals such as groups/pluralities, again as expected on the assumption that the granularity parameter for *po* is unspecified. This allows the set of individuals denoted by the subject argument to be partitioned nonatomically into groups/plural individuals. But we should nonetheless find that these plural individuals must be exhaustively distributed over. If no such exhaustivity requirement is found, then the arguments for the universal-quantification approach will be called into question, and the obvious question will be how to reconcile the findings for transitive sentences with those for intransitive sentences.

Experiment 1 was designed to explore these issues with native speakers of Serbian.

3.1 | Method and procedure

3.1.1 | Participants

A total of 70 adult native speakers of Serbian (49 female and 21 male; mean age 28.03, minimum 15, maximum 59) completed an online questionnaire that took less than 10 minutes. All participants were self-reportedly monolingual Serbians.

3.1.2 | Design and procedure

We used a 2 × 3 factorial design and a picture-verification task. The independent variables were Groups (i.e., how many groups depicted), with two levels, one and four, and Picture, with three levels, A = exhausted, B = nonexhausted, and C = nonexhausted with a different object. We had five observations per condition, five controls, and 20 filler items: 55 items in all. The experiment was preceded by a short context, explaining that the computer would present visual scenarios accompanied by sentences meant to describe them. Participants were asked to decide, for each scenario presented, if it was possible to describe it with the sentence presented.

3.1.3 | Items

We used variations of one type of test sentence, with the structure '[Subject]s are [verb]ing *po* one [object]':

(7)	Majmun-i	drž-e	**po**	jedan	kišobran-∅.
	monkey-PL.NOM	hold-PRS.3PL	DISTR	one	umbrella-ACC

'(The) monkeys are holding DISTR one umbrella.'

There were 30 unique situations, using 10 distinct verbs and 15 different agents (the same animals used in the experiment in Bosnić, Spenader & Demirdache 2020). The items were distributed in six lists, and each participant saw only one condition per item. The independent variables were manipulations of the exhaustivity of the situations (Picture) and of the grouping of the animals in the pictures (one big group versus four smaller groups: Groups), resulting in six combinations, illustrated by figure 3.

Under the three conditions labeled as Groups = one, the (putative) distributive key is *partitioned atomically*. The members of the set to be distributed over are thus singular individuals:

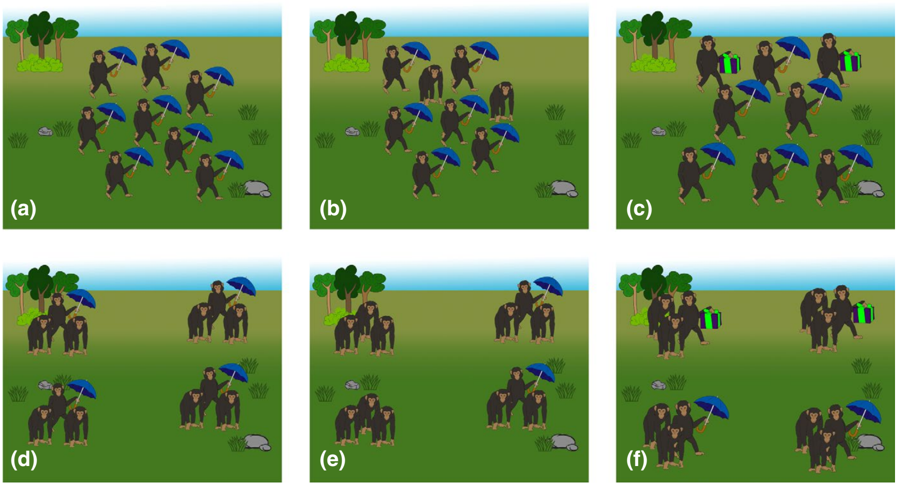

FIGURE 3 Six experimental conditions (2 × 3 study) for transitive sentences with *po*. (a) One–A, exhausted individuals. (b) One–B, nonexhausted individuals. (c) One–C, nonexhausted individuals with different objects. (d) Four–A, exhausted groups. (e) Four–B, nonexhausted groups. (f) Four–C, nonexhausted groups with different objects

individual monkeys, in the sample situation in figure 3. Under the three conditions labeled as Groups = four, on the other hand, the distributive key is *partitioned nonatomically*. The members of the set to be distributed over are thus plural individuals: triplets of monkeys, in the sample situation in figure 3. The visual features of each condition are as follows, taking figure 3 as the sample situation.

The A pictures are exhausted, because we either have every potential agent (monkey) holding an object (umbrella), in the conditions where Groups = one, or one monkey per group holding an umbrella, in the conditions where Groups = four. The B pictures are nonexhausted conditions because they contain either two monkeys (one–B) or two groups of monkeys (four–B) who are not holding an umbrella. Lastly, the C pictures are nonexhausted conditions because what is being held by two monkeys (one–C) or two groups of monkeys (four–C) is not an umbrella but a gift. Recapitulating, whether distribution is over singular or plural individuals, there are two ways of making the nonexhausted scenario true: either the monkeys are not holding anything (B conditions) or they are holding something else, namely gifts (C conditions).

Our predictions based on the conclusions for intransitive sentences are as follows. The test sentence should be judged true under both the one–A and four–A conditions since the distributive key (which by hypothesis can be either atomically or nonatomically partitioned) is exhausted. Conversely, under the B and C conditions, the test sentence should be judged false since the distributive key (whether it is partitioned atomically or not) is not exhausted.

To sum up, if the semantics of *po* involve universal quantification, our test sentences should be judged true only under the exhausted conditions, one–A and four–A. If, however, the speakers fail to reject the nonexhausted conditions as predicted, then there is no evidence for a distributive key, and we should explore an alternative analysis, such as event plurality, as originally proposed for *po* by Knežević 2015. Recall that this analysis predicts that all six conditions should be accepted because the minimum requirement of at least two events of monkeys holding umbrellas is met.

3.2 | Results of experiment 1

Results are in figure 4.[8] They show an unexpectedly low acceptance rate in the four–A condition; this rate was not predicted to be low by the universal-quantification analysis. We also expected the B conditions to be more strongly rejected than they were. Lastly, there was a strong rejection of the C conditions.

Looking closely at the results for the four–A condition, we see an hourglass shape for the distribution of per-participant means, suggesting there are relatively many participants who either reject or accept this condition. We thus checked these means and grouped participants according to their four–A acceptance and rejection rates. The criterion used for grouping was the following. Participants who accepted the four–A condition 4/5 or 5/5 times were in the yes-saying group; those who accepted it 0/5 or 1/5 times were in the no-saying group. Everyone else (10 participants in total) was in the "middle" group (2/5 and 3/5 acceptance), which was deemed to be at chance and was not statistically analyzed. By means of this grouping process, we identified two different populations of participants: out of the 70 participants tested, 36 overwhelmingly accepted items

[8]We used "pirateplots" in R (the Yarrr package) for our RDI graphs, using the inference method se (standard error).

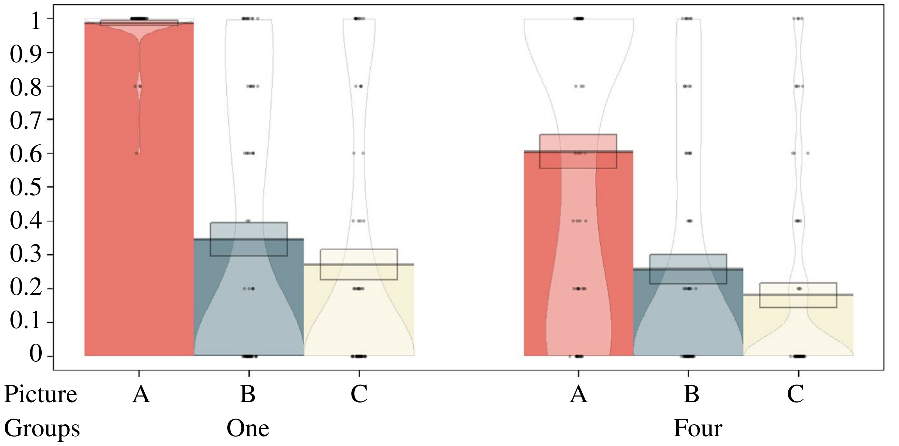

FIGURE 4 RDI plots of the results of experiment 1 with standard errors. The labels on the *x* axis refer to the two variables that define the six experimental conditions. Bars show aggregate acceptance rate (no answers being coded as 0 and yes answers as 1), lines show the distribution of answers, and each dot represents the mean of an individual participant's answers

under the four–A condition, while 24 rejected them. Separating the participants in this way we get very clear results, as shown in figure 5.

The no-saying population is very strict with their interpretations: they require individual monkeys to be exhausted in every situation, so they only accept the atomically exhausted one–A. This pattern of responses is thus similar to the response patterns for atomic universal distributive-key quantifiers (e.g., English *each*).

The yes-saying population, on the other hand, has a distinct pattern of interpretation: they appear to accept the nonexhausted B and C conditions at a much higher rate. However, further inspection of this pattern, once again examining the distribution of the per-participant means, revealed another population split. This time, we took the acceptance rate of the one–B condition, the highest-rated condition among the yes-saying population, as the basis for the split, using the same criterion as before. The result is in figure 6. One set of participants shows a pattern of accepting A conditions (exhausted) and rejecting B and C conditions (nonexhausted) regardless of the value of the Groups variable. This pattern is observed in 13 out of the 36 participants. The second set of 17 participants shows a pattern of systematically accepting (almost) all conditions. These two patterns are very distinct, as can be seen in the graphs in figure 6.[9]

For the statistical analyses we did a repeated-measures mixed-effects logistic regression using the glmer() function of the LME4 package (Bates et al. 2015) in R (R Core Team 2014). We performed the analysis on the whole population first and included Groups and Picture as predictors (fixed factors) and, as random factors, Subject and Item. We used a stepwise-variable-addition procedure and compared the models according to the Akaike Information Criterion. The model with the lowest value according to this criterion was considered the best fit, with a complex model being preferred over a simpler model only if its Akaike Information Criterion value was

[9]Six participants are in the middle when it comes to consistently accepting or rejecting the one–B condition, but they mostly reject other conditions, namely one–C, four–B, and four–C. This brings them closer to the yes-to-A population, but we are still leaving them out for the sake of bringing out more clear-cut, unambiguous response patterns.

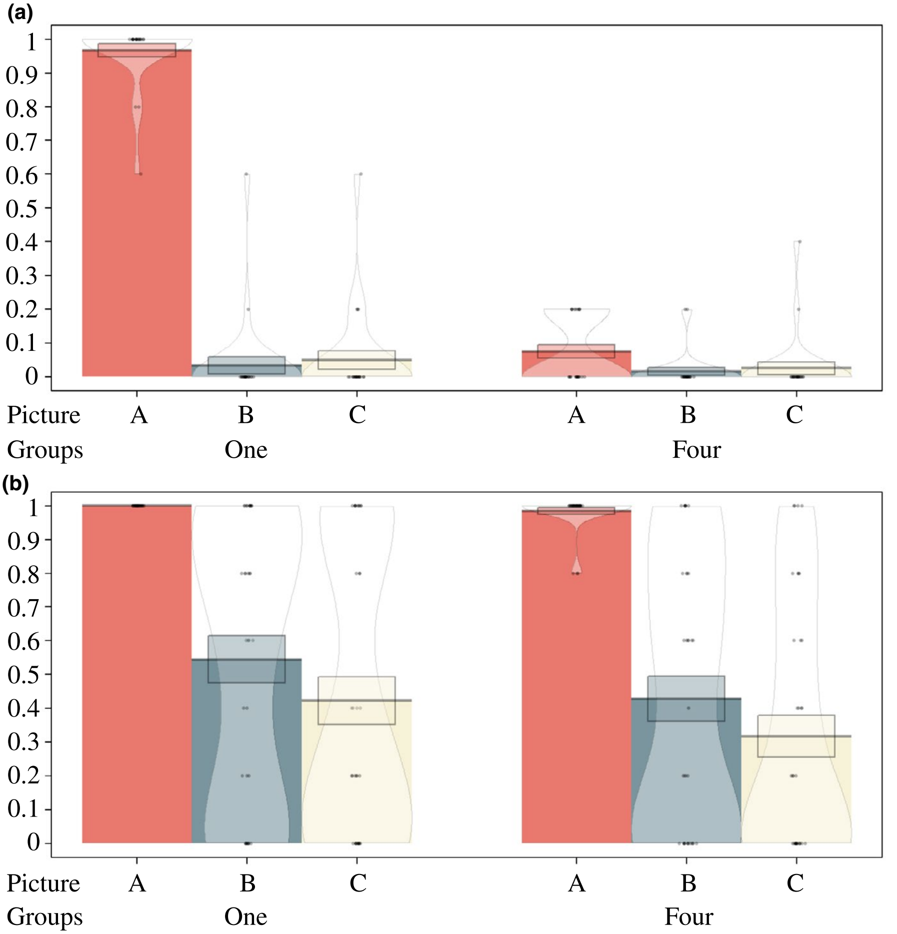

FIGURE 5 RDI plots for the two populations that were defined based on their responses to the four–A condition. (a) The no-saying population. (b) The yes-saying population. Bars show aggregate acceptance rate, lines show the distribution of answers, and each dot represents the mean of an individual participant's answers

two or more points lower (Akaike 1974). Models that included the random factor Item were no better than models without, so we were able to exclude Item from further analysis. The best-fitting model was a random slope of Picture for the random factor Subject with an interaction between Groups and Picture as fixed factors. See table A1 in appendix A.[10]

We took the four–A condition (exhausted groups) as a reference level (intercept) and compared other conditions to it. The estimate of the intercept is positive and significantly different from that of one–A: one–A is more likely to be judged true than four–A ($\beta = 6.855$, z value $= 8.613$, $p \leq 0.0001$). The nonexhausted conditions (one–B, one–C, four–B, and four–C) were significantly less likely to be accepted than the intercept. See the complete model output in table A1.

[10]We also compared models with additional factors, such as age, verb, gender, and education. Models including these factors were not significantly better than the best model reported above, so these factors did not significantly influence responses.

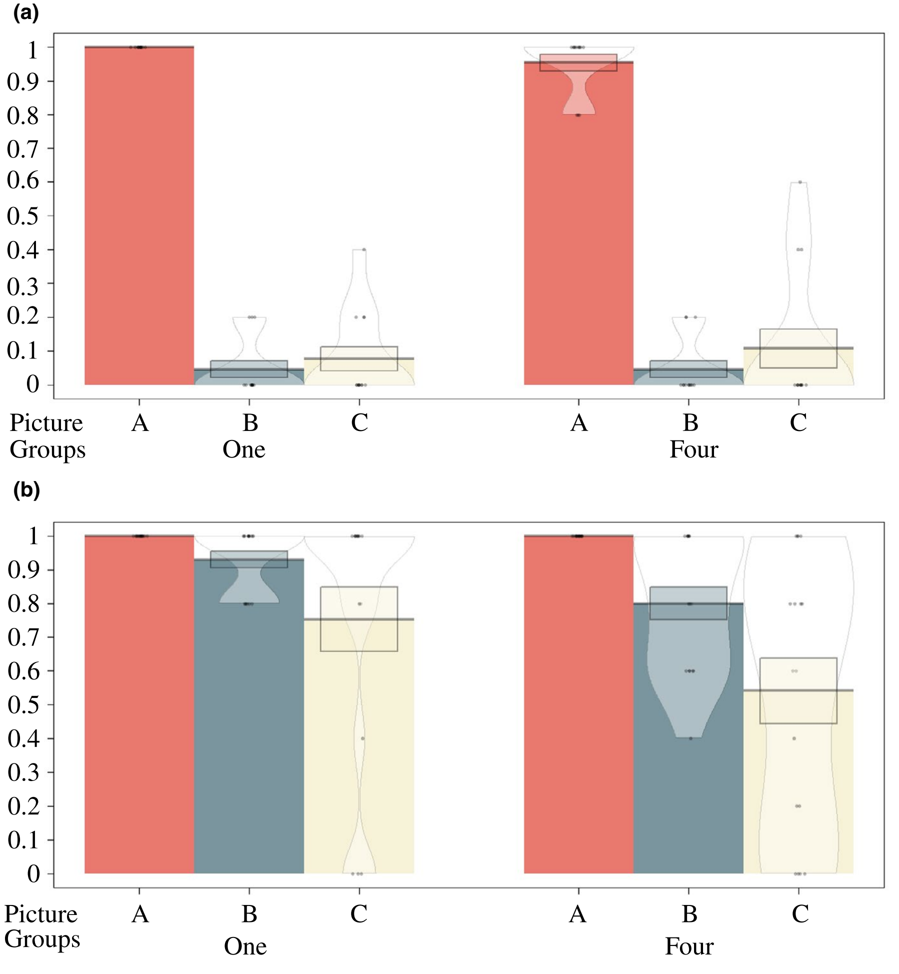

FIGURE 6 RDI plots for distinct patterns in the yes-saying population. (a) The yes-to-A population. These 13 people rejected B and C conditions. (b) The yes-to-all population. These 17 people accepted B and C conditions. Bars show aggregate acceptance rate, lines show the distribution of answers, and each dot represents the mean of an individual participant's answers

Additionally, we performed a post-hoc multiple-comparison Tukey test to check for significant differences between each level combination. Apart from the obvious differences between the combinations, we found significant crossconditional differences: one–B ($p < 0.0024$) and one–C ($p < 0.0006$) both had significantly higher acceptance than four–C, and one–B was also significantly higher than four–B ($p < 0.0014$). See table A2 for details. More importantly, nonexhausted conditions (B and C) are not different within their Groups level (one or four).

Since we also have clearly identified populations, each with its own response pattern, we performed a mixed-effects logistic-regression analysis on the no-saying population as a whole and on the yes-saying population as a whole. For both models, we used a stepwise-variable-addition procedure with random slopes, setting the four–A condition as the reference condition. Our fixed factors were Groups and Picture, while Item and Subject were the random factors. Item was excluded from further analysis because including it as a factor did not improve the models.

When it comes to the yes-saying population,[11] the model showed only a significant main effect of Groups and Picture as predictors. The best-fitting model was the additive model with the random slope of Picture for Subject. The one–A condition is still more likely to be judged true ($\beta = 1.038$, z value $= 4.451$, $p \leq 0.001$) than four–A is ($\beta = 5.678$, z value $= 3.991$, $p \leq 0.001$), while B and C conditions are more likely to be rejected. See the model output in table A3.

The analysis for the no-saying population yielded something different. The best model found a significant interaction between the fixed factors Groups and Picture and also included a random slope of Picture for Subject. As can be seen from figure 5a as well, one–A is the only condition that is significantly more likely to be accepted ($\beta = 6.446$, z value $= 3.431$, $p \leq 0.001$) than the other conditions. All other conditions were likely to be rejected, with some variance (e.g., four–B was not significantly different than four–A: $\beta = -4.623$, z value $= -1.747$, $p = 0.081$). The full model output is in table A4.

3.3 | Discussion of experiment 1

In our results, we identified three distinct populations: a no-saying population (24/70 people who said no to four–A, the exhausted-groups condition, and to all other conditions except one–A: figure 5a); a yes-to-A population (13/70 people who said yes to A conditions only: figure 6a); and a yes-to-all population (17/70 people who said yes to all conditions: figure 6b).

The no sayers show precisely the pattern of behavior expected under the proposal that *po* is a distributive universal quantifier that distributes down to atomic individuals, similar to English *each*. On the other hand, the yes-to-A pattern is exactly the pattern of behavior expected under the proposal that *po* is a distributive universal quantifier that can also distribute over "bigger," nonatomic entities (here groups/plural individuals): the test sentences are judged true only under the one–A and four–A conditions.

The yes-to-all pattern is the most interesting. It is unexpected on a universal-quantification analysis of *po*. It does, however, match the hypothesis that *po* marks event plurality without universal quantification, as Knežević 2015 and Knežević & Demirdache 2017, 2018 contend, since the only requirement to be met on this analysis is that there be more than one event of the relevant type, which is the case under all of the conditions. The issue then is whether the yes-to-all response pattern is evidence for the event-plurality hypothesis or whether there is an alternative interpretation of this pattern.

Essentially, we draw two major conclusions from the three patterns of responses uncovered:[12]

First, we can take the no-saying and yes-to-A patterns to both reflect a universal-quantification interpretation of *po*, on the assumption that *po* can distribute over nonatomic entities as well as

[11]In order to use a more powerful statistical analysis (mixed-effects models) we needed to change one positive answer into a negative one under the one–A condition because that condition originally had no variance (it was accepted 100% of the time). Otherwise, we would not be able to do mixed-effects modeling for these results, and it is impossible to exclude only one condition from the analysis.

[12]We take the full range of response patterns elicited, in particular the substantial number of participants who accepted the nonexhausted scenarios and the exhausted condition involving distribution over nonatomic entities/groups (four–A), to show that distributive-share markers such as *po* behave very differently from binominal *each*, whether or not the latter is also analyzed as a distributive-share marker (see footnote 2).

atomic entities. We thus subsume these two patterns under a single pattern, which we henceforth refer to as the universal-quantification pattern. Since the granularity parameter for *po* is left free, it can be set to atoms (singular individuals), yielding the no-saying pattern, or to bigger-size entities (plural individuals/groups), yielding the yes-to-A pattern. What determines how it gets set is the supporting (linguistic and visual) context, what the subject judges to be salient in the input provided. So, in principle, it could just as well be set to bigger-size entities such as sums or amounts of time; it is just that, in our experiment, this setting is not made available by the supporting context, the experimental stimuli provided.

Second, what about the yes-to-all pattern? Does it instantiate an event-plurality pattern of interpretation for *po*, or could it be explained away by appealing to auxiliary assumptions while maintaining the universal-quantification line of analysis? One thing to keep in mind is that Serbian bare plurals are ambiguous between definite and indefinite interpretations. This opens the possibility that the yes-to-all response pattern could be imputed to an indefinite interpretation of the subject argument serving as the distributive key for *po*. On this alternative analysis, *po* would be a universal quantifier distributing over an indefinite subject, and the resulting interpretation for (7) could be paraphrased as 'Some monkeys are each holding an umbrella', which will come out true under all the conditions since they all involve at least two monkeys each holding an umbrella.[13] In other words, it could well be that *po* is indeed a universal quantifier but that the exhaustivity requirement of *po* is hidden by an indefinite interpretation of the bare subject DP *majmuni* 'monkeys'.[14]

Summarizing, there are two alternative explanations for the yes-to-all pattern. It could instantiate either an event-plurality pattern of interpretation of (7) or a universal-quantification pattern of interpretation where the bare plural subject in (7) is construed as indefinite. To test whether indeed the yes-to-all response pattern can be explained away as an indefinite interpretation of the subject DP (the distributive key), we designed a follow-up experiment to force the subject to be interpreted as definite.

4 | EXPERIMENT 2: FORCED DEFINITE INTERPRETATION OF THE SUBJECT

The results from experiment 1 revealed two populations of speakers with distinct response patterns. On the one hand, we have a population of speakers who appear to interpret *po* as a universal quantifier: they either say no to all conditions except one–A (these no sayers set the granularity of *po* to atoms, on the basis of what they take to be salient in the supporting context) or say no to all conditions but one–A and four–A (these yes-to-A sayers set the granularity of *po* to groups). On the other hand, we also have a population of speakers who say yes to all the conditions. This response pattern can be taken to reflect either an event-plurality pattern of interpretation or a universal-quantification pattern of interpretation in which the subject DP (providing the distributive key for *po*) is construed indefinitely.

[13]We are indebted to an anonymous reviewer for highlighting this interpretation of the yes-to-all pattern.

[14]To be thorough, we conducted a small experiment with 29 Serbian speakers to confirm that when the subject is indefinite, *po* test sentences are accepted under all the experimental conditions. We made the subject DP indefinite by embedding it in an existential construction ('There are monkeys that are holding *po* one umbrella'). All experimental conditions were indeed accepted at ceiling.

A follow-up experiment was designed to force the subject to be interpreted as a definite plural in *po* sentences. We can linguistically control the interpretation of the bare plural subject of our test sentences (e.g., *majmuni* 'monkeys') rather straightforwardly via the context: if the subject is introduced in a sentence immediately preceding the target test sentence, then it will be interpreted as a definite in the target test sentence. The prediction is that, if *po* is a universal quantifier, the yes-to-all pattern of responses in the follow-up experiment should disappear, since the exhaustivity requirement can no longer be overridden by an indefinite interpretation of the subject DP (that is, of the distributive key). To clearly see the effect of definiteness, we moreover tested both sentences *with* and *without po* in these forced-definite-subject contexts.

4.1 | Method and procedure

4.1.1 | Participants

A total of 61 native speakers of Serbian (40 female and 21 male; mean age 26.43, minimum 16, maximum 53) completed an online questionnaire. No participants had taken part in experiment 1.

4.1.2 | Design and procedure

We used a picture-verification task, with a 2 × 2 (Groups × Picture) factorial design and five observations per condition. Additionally, we included 20 unique scenarios with sentences without *po*, which were identical for all participants. The comparison of sentences with and without *po* was added to check for the effects of definiteness in distributively unmarked sentences. Finally, we only used the A and B conditions from experiment 1 because these showed the most extreme response patterns and were expected to be sufficient to identify an effect of definiteness on potential exhaustivity requirements. Participants saw 20 sentences with *po*, 20 without *po*, five control items, and 10 fillers: 55 items in total.

4.1.3 | Items

We used a context sentence preceding the test sentence to manipulate the definiteness of the subject DP: by introducing the subject DP in an existential construction, we forced it to have a definite interpretation in the following test sentence. The test sentences with *po* were the same as in experiment 1, for a direct comparison. The test sentences without *po* were novel sentences with unambiguously distributive verbs. This ensured that a given sentence would not be mistakenly rejected because of a bias for the collective interpretation (this being the default interpretation of distributively unmarked sentences, according to Knežević 2012). Thus, the new scenarios that were created used distributive VPs that are hard to interpret collectively, such as 'brush one's teeth', 'play the guitar', 'wear boots', 'wear a cap', and 'eat a lollipop'.

The format of the test sentences (with context sentence preposed) was as in (8) for those with *po* and as in (9) for those without *po*. (The number in the context sentence, 'eight' or '12', depended on the scenario, specifically the value of Groups: the one-group pictures depicted a single group of eight individuals while the four-group pictures depicted four groups of three individuals.)

(8) 'There are eight/12 [subject]s in the photo. [Subject]s are [verb]ing **po** one [object].'

(9) 'There are eight/12 [subject]s in the photo. [Subject]s are [verb]ing [object].'

To illustrate the *po*-less items, figure 7, paired with the test sentence in (10), is an example of one of the nonexhausted conditions, one–B, while figure 8, paired with the test sentence in (11), is an example of the four–A condition, which exhausts groups but not individuals.

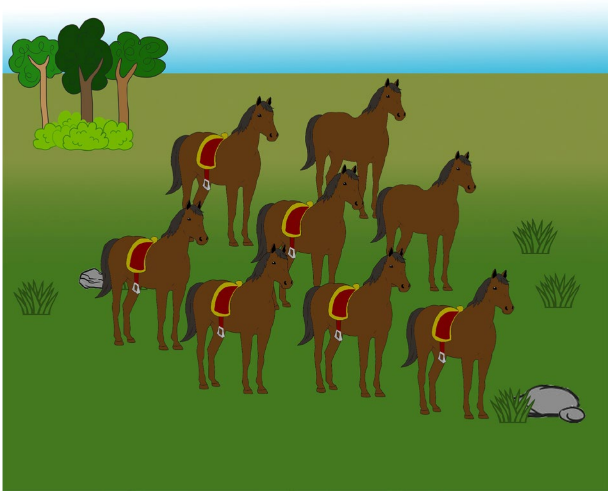

FIGURE 7 A nonexhausted scenario: horses in a single group, some with and some without saddles. Paired with the *po*-less sentence in (10)

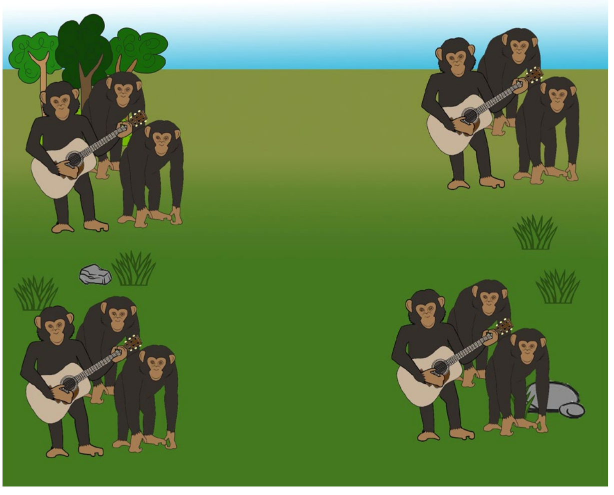

FIGURE 8 An exhausted scenario: monkeys in four groups, each group containing a monkey playing the guitar. Paired with the *po*-less sentence in (11)

(10) Na slic-i ima-∅ osam konj-a.
on photo-SG.LOC have-PRS.3SG eight horse-PL.GEN
Konj-i ima-ju sedl-o.
horse-PL.NOM have-PRS.3PL saddle-ACC
'In the photo, there are eight horses. The horses have a saddle.'

(11) Na slic-i ima dvanaest majmun-a.
on photo-SG.LOC have-PRS.3SG twelve monkey-PL.GEN
Majmun-i svira-ju gitar-u.
monkey-PL.NOM play-PRS.3PL guitar-ACC
'In the photo, there are twelve monkeys. The monkeys are playing the guitar.'

4.2 | Results of experiment 2

The results are shown in figure 9a for the *po*-marked sentences and in figure 9b for the unmarked sentences.

The one–A condition has a ceiling acceptance with test sentences with *po* (98%), as predicted if *po* is a universal quantifier, since the bare plurals here are interpreted as definite DPs. We also found a near-ceiling acceptance (89%) of one–A with *po*-less sentences. Interestingly, all other conditions, including four–A, were significantly more likely to be rejected for both sentence types. Recall that in experiment 1 we originally found evidence of three distinct patterns of responses: no sayers, yes-to-A sayers, and yes-to-all sayers. In experiment 2, we no longer find evidence of these distinct populations being represented by a substantial number of people. There is a subset of 18/61 participants that showed a ceiling acceptance of both one–A and four–A, but they were not uniform in their answers to the B conditions. Out of these 18 participants, four had a universal-quantification pattern (yes to one–A and four–A and no to one–B and four–B) and 14 had (by hypothesis, since the universal-quantification-plus-indefinite-subject analysis is no longer available) an event-plurality pattern (yes to one–A, one–B, and four–A;

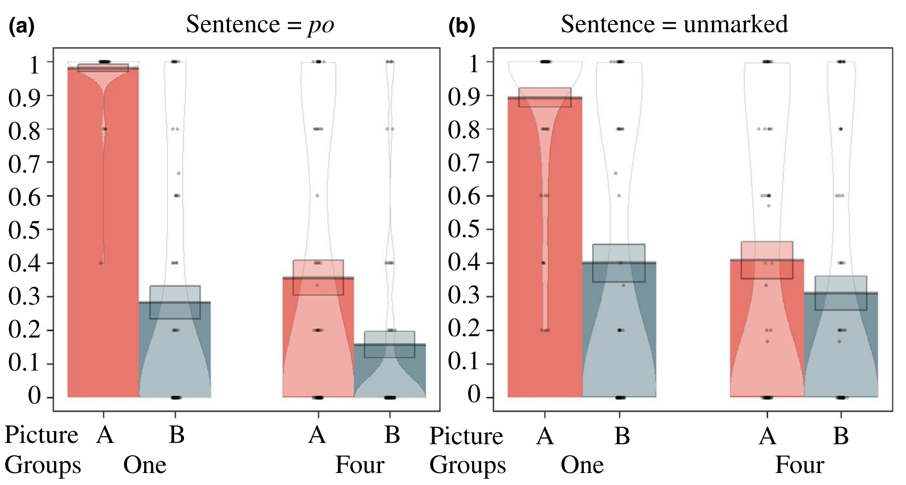

FIGURE 9 RDI plots with standard errors for the definiteness follow-up experiment. (a) Results for the sentences with *po*. (b) Results for the sentences without *po*. The labels on the *x* axis refer to the two variables that define the four experimental conditions. Bars show aggregate acceptance rate, lines show the distribution of answers, and each dot represents the mean of an individual participant's answers

four–B was at chance); they treated the sentences without *po* the same way. Meanwhile, 35/61 participants systematically rejected all conditions except one–A, showing the no-saying pattern from experiment 1, and 27 of these 35 did the same for the sentences without *po*. Lastly, 5/61 participants showed a peculiar pattern: they accepted all the conditions with the sentences without *po* but rejected the nonexhausted ones (together with four–A) with the *po* sentences.

We again used mixed-effects logistic regression (the glmer() function in R) to analyze our data and compare sentences with and without *po*. The reference condition was again set to be the crucial condition four–A with *po* sentences, being the one where acceptance rates drastically lowered (35.8%) in comparison to experiment 1 (60.5%). Using a stepwise-variable-addition procedure, we found that the best model had a triple interaction of Picture, Sentence Type (*po*-marked versus unmarked), and Groups and a random slope of these three predictors for Subject. There was no significant effect of Item in the model. See the complete model output of the fixed factors in appendix B. People are less likely to judge one–B, four–A, and four–B as true and are likely to accept one–A regardless of the presence of *po*. The difference between four–A without *po* and four–A with *po* is not statistically significant ($\beta = 0.197$, z value $= 0.448$, $p = 0.654$). However, one–B and four–B pictures with and without *po* are more likely to be rejected than four–A with and without *po*; see appendix B for details.

4.3 | Discussion of experiment 2

Recall the initial question we sought to address: can the yes-to-all response pattern uncovered in experiment 1 be explained as a consequence of an indefinite interpretation of the subject DP serving as the distributive key? Experiment 2 was designed to test this by forcing the plural subject of a *po* sentence to be interpreted as definite. We predicted that if this construal was forced, we should no longer find the yes-to-all response pattern. With a definite plural, *po* will enforce its exhaustivity requirement on the distributive key; nonexhausted scenarios should therefore be rejected. We predicted, then, that the responses in experiment 2 would support an analysis of *po* as a universal quantifier. This in turn would suggest that the yes-to-all responses in experiment 1 could plausibly be subsumed under a universal-quantification (+indefinite-subject) pattern of interpretation.

The results of experiment 2 reveal a more intricate story, however. There was an increase in the population showing a universal-quantification pattern of responses, as predicted. However, 14 people still accepted (almost) all conditions, contrary to our expectation that they should now reject the nonexhausted scenarios.

An important further observation was that 27 participants had the very same no-saying pattern (saying yes to one–A and no to all other conditions) in sentences with and without *po*. Given these results, there is now a question as to whether the observed exhaustivity requirements are actually originating from *po*, since they also appear to arise in the absence of *po*. One possibility is that the no-saying response pattern could be ascribed to a well-known characteristic of definite plural DPs, namely that they typically show maximality effects. The contexts in experiment 2 were set up to force a definite interpretation of the bare plural subject and avoid an indefinite interpretation of it, but doing this also introduces potential maximality effects.

We are now faced with a conundrum: how can we tell whether the exhaustivity effects in *po* sentences originate from *po* or are in fact maximality effects arising via a definite interpretation of the bare plural subject? As we shall now argue, there is an elegant way to experimentally distinguish maximality from exhaustivity and thus resolve this issue: so-called homogeneity effects in negative sentences containing either plural definite descriptions or universally quantified descriptions will serve as a diagnostic to tease maximality and exhaustivity apart.

5 | EXPERIMENT 3: MAXIMALITY VERSUS EXHAUSTIVITY—HOMOGENEITY EFFECTS WITH *PO*

What is the difference between *maximality* and exhaustivity? As has been extensively discussed in the literature (Brisson 1998, Schwarz 2013, Križ 2015, and references therein), statements with definite plurals tend to receive a maximal/universal interpretation: for the sentence in (12a) to be judged true, for example, *all* the entities satisfying the description *monkey* in the context have to satisfy the property of holding an umbrella.

(12) a. The monkeys are holding an umbrella.
True iff *all* the monkeys are holding an umbrella
b. Every monkey is holding an umbrella.
True iff *all* the monkeys are holding an umbrella

Importantly, however, sentences with plain definite plurals such as (12a) and sentences with universally quantified phrases such as (12b) do not have quite the same truth conditions. First, definite plurals do allow nonmaximal readings; that is, they tolerate exceptions. As Lasersohn 1999 observes, the sentence *The townspeople are asleep* is commonly judged true even if not all the townspeople are asleep. Nonmaximal interpretations arise via pragmatic weakening, given the right context.[15] Pragmatic weakening can be achieved by domain restriction: for instance, the sentence in (12a) would be acceptable in a situation where some monkeys are not holding an umbrella and are doing something else, if the hearer generously restricts the domain under consideration to only those monkeys who are holding an umbrella (see Brisson 1998 for more on domain restriction). Another possibility is to make a contrast with a different set: so if we add pandas to our picture, then even if there are nonparticipating monkeys, the sentence in (12a) would be fine because the set of monkeys contrasts with pandas. Lasersohn 1999 puts forward the notion of "pragmatic slack," essentially arguing that speakers can use sentences imprecisely as long as the exception in the context does not matter and is not relevant for the current purpose of the conversation. While definite plurals allow maximality violations under the appropriate pragmatic conditions, universal quantifiers like in (12b) do not allow exceptions. In Lasersohn's terms, universal quantifiers are *maximizers* (enforcing maximality) or *(pragmatic-) slack regulators* (enforcing precise, strict interpretations).

The difference between the truth conditions of sentences with definite plurals and sentences with universally quantified phrases comes out in contexts where the property ascribed to the plurality denoted by the subject DP does not hold uniformly, *homogeneously*:

(13) a. The monkeys are holding an umbrella.
Undefined if *some but not all* monkeys are holding an umbrella
→ Truth-value gap
b. Every monkey is holding an umbrella.
False if *some but not all* monkeys are holding an umbrella
→ No truth-value gap

[15]The opposite view has also been defended (e.g., Malamud 2012), namely that definite plurals have existential (nonmaximal) semantics and that maximality effects arise via pragmatic strengthening. Experimental evidence from Schwarz 2013, however, indicates that the maximal interpretation is the preferred one.

In such contexts sentences with definite plurals are trivalent in their truth conditions: they are neither completely true nor completely false, yielding a truth-value gap (13a). In accord with those who have extensively worked on the topic both experimentally and theoretically (see Križ 2015, Križ & Chemla 2015, Križ 2017, Tieu et al. 2019), we henceforth refer to such contexts as *gap scenarios* or nonhomogeneous scenarios. In contrast, the presence of a universal quantifier excludes truth-value gaps (13b).

The truth-conditional difference between definite plurals and universally quantified phrases comes out even more clearly in negative statements:

(14) a. The monkeys are **not** holding an umbrella.
True if *none* of the monkeys are holding an umbrella
Undefined if *some but not all* monkeys are holding an umbrella
→ Truth-value gap

b. Every monkey is **not** holding an umbrella.
True if *some but not all* monkeys are holding an umbrella (preferred reading)
(Less favored reading: **true** if *none* were holding an umbrella)

We see that the apparent equivalence illustrated in (12) between positive statements with definite plurals and positive statements with a universal disappears with negation: (14b), with the universal, is judged true in a context where some of the monkeys but not all of them are holding an umbrella, while (14a), with a definite, yields a truth-value gap (is neither true or false) in this context and comes out as true only when none of the monkeys are holding an umbrella.

The general idea found in the literature is that a predicate must hold homogeneously of the plurality denoted by a plural definite: either the monkeys are all holding an umbrella (12a) or none of the monkeys are holding an umbrella (14a). Since homogeneity cannot be satisfied in a gap context, definite plurals give rise to a truth-value gap. Plural definites are thus said to give rise to *homogeneity effects*—unlike universally quantified descriptions.

Turning to Serbian, the interpretation of bare nouns as either definite or indefinite is typically context dependent. Importantly, however, sentences containing a bare noun in Serbian carry an *implicature* of uniqueness, for singular bare nouns, or of maximality, for plural bare nouns: the hearer expects the referent to be unique/maximal in any given context, unless the context rules uniqueness/maximality out.[16] As shown in (15a), when a plural referent is introduced into the discourse, a subsequent DP with the same descriptive content refers back to the entire group and not to a subset of the group. However, this uniqueness/maximality effect is not a presupposition but an implicature, since it can be explicitly canceled, as shown in (15b), by adding a subsequent assertion (denying maximality).[17]

[16]Arsenijević 2018 points out that anaphoric/unique interpretations require a bare noun: as (i) illustrates, they are unavailable with a demonstrative *taj* 'that'.

(i) Lingvist-a i advokat-∅ su doš-li na žurk-u.
linguist-NOM and lawyer-NOM AUX come-PST.3PL on party-LOC
(*Taj) lingvist-a je done-o kolač-e.
that linguist-NOM AUX bring-PST.3SG cake-PL.ACC
'A linguist and a lawyer came to the party. The/*this linguist brought cookies.'

[17]This argument is adapted from Gillon 2014, which makes the very same point for Sḵwx̱wú7mesh Salish, another language with no determiners coding a definite–indefinite contrast.

(15) a.

Osam	žen-a	je	uš-lo	u	sob-u.
eight	woman-PL.NOM	AUX	enter-PST.3SG	in	room-LOC

Popriča-la	sam	sa	žen-ama.
talk-PST.1SG	AUX	with	woman-PL.INS

‘Eight women entered the room. I talked to (all of) the women/*some of the women.’

b.

Kupi-la	sam	deset	jaj-a.
buy-PST.1SG	AUX	ten	egg-PL.ACC

Jaj-a	su	se	slomi-la	na	put-u	do
egg-PL.NOM	AUX	REFL	break-PST.3PL	on	way-SG.LOC	until

kuć-e.	U stvari,	ni-su	sva.
home-SG.GEN	actually	not.be-PRS.3PL	all

‘I bought ten eggs. (The) eggs cracked on the way home. Actually, not all of them.’

Homogeneity effects have been discussed for other distributive-share markers crosslinguistically. Križ 2017 argues that *po* in Russian is compatible with the homogeneity-based trivalence of plural predication: that is, plural bare nouns interpreted as definites display trivalent truth values even when used with Russian *po*, but the Russian universal quantifier *každyj* does not lead to truth-value-gap responses. This suggests that *po* distributive numerals, at least in Russian, do not pattern like universally quantified expressions since they lack the hallmark property of universal quantifiers with respect to homogeneity: they are not slack regulators that can remove homogeneity effects. This contrasts with reduplicated numerals/dependent indefinites in languages like Hungarian, also considered to be a form of distributive-share marking, which do seem to function as slack regulators in that they do not allow truth-value-gap responses (Križ 2017, Kuhn 2017).

Returning to our Serbian results, the issue that our findings for experiment 2 raised was how to tell whether the exhaustivity effects found in *po* sentences originate from a universal-quantifier interpretation or are, in fact, maximality effects arising via a definite interpretation of the bare plural subject. We can use the presence (or lack) of homogeneity effects in negative contexts as a diagnostic to tease maximality and exhaustivity apart.

Experiment 3 was designed to achieve this by testing and comparing responses to three types of negative sentences: with *po* marking the object, with the universal quantifier *svaki* in subject position, and with only a definite plural in subject position. Our aim was to establish baseline responses for both *svaki* and definite plurals and then to compare the responses elicited for *po* with these baseline responses. If *po* has universal quantificational force, then (i) it should remove homogeneity effects in negative sentences with a definite subject, and (ii) just like negative *svaki* sentences, negative *po* sentences should not display truth-value-gap answers and should be judged true in nonhomogeneous/gap scenarios. If instead *po* is an event-plurality marker with no universal quantificational force of its own, then (i) it should not remove homogeneity effects in negative sentences with a definite subject, and (ii) responses for negative *po* sentences will pattern differently than for negative *svaki* sentences, displaying truth-value-gap answers in nonhomogeneous/gap scenarios.

5.1 | Method and procedure

5.1.1 | Participants

A total of 45 native speakers of Serbian (33 female and 12 male; mean age 33.53, minimum 20, maximum 48) completed a short online experiment. They were recruited via social media and were not compensated for their time.

5.1.2 | Design and procedure

We again used a picture-verification task, with a 3 × 1 (Sentence Type × Picture) factorial design distributed in three lists with five observations per condition. We manipulated three types of negative sentences (*svaki*, *po*, and definite plural) and one type of picture (one–B nonexhausted scenarios like figure 3b).[18] Additionally, we created 15 sentences with *svaki*, *po*, and definite plural subjects in positive contexts, for comparison. Participants saw 15 negative sentences (target items), the 15 positive sentences (identical for all participants), and five control items: 35 items in total.

5.1.3 | Items

In accord with Križ & Chemla's 2015's and Tieu et al. 2019's methods for testing homogeneity effects in English and French, we tested only gap situations, such as figure 10, with three types of sentences, such as those in (16). Just like in experiment 2, we controlled for the definite interpretation with a preceding context: 'There are eight elephants in the photo'.

(16)	a.	Slon-ovi	ne	nos-e	šešir-∅.	
		elephant-PL.NOM	NEG	wear-PRS.3PL	hat-ACC	
		'The elephants are not wearing a hat.'				
	b.	Slon-ovi	ne	nos-e	**po**	šešir-∅.
		elephant-PL.NOM	NEG	wear-PRS.3PL	DISTR	hat-ACC
		'The elephants are not wearing DISTR a hat.'				
	c.	Ne	nos-i	**svaki**	slon-∅	šešir-∅.[19]
		NEG	wear-PRS.3SG	every	elephant-SG.NOM	hat-ACC
		'Not every elephant is wearing a hat.'				

[18]For the question at hand (and given the simplicity of the design), it was not necessary to look at nonatomic/group cases such as four–B.

[19] It has been claimed that Serbian does not allow inverse-scope readings (see Progovac 1994, Bošković 2012), so we changed the word order to obtain a surface-scope reading in which negation scopes over *svaki*. Note also that a surface-wide-scope reading of *svaki* over negation ('every' > 'not') is very unnatural. Because this adjustment creates an imbalance in the test items, we also checked for the effect of word order with a follow-up experiment using sentential (wide-scope) negation for all three types of sentences (e.g., 'It is not correct that every elephant is wearing a hat'/'It is not correct that elephants are wearing (*po*) a hat'). For reasons of space, we do not go into the details of this experiment here, but we can attest that the results of the follow-up were not significantly different from the experiment we describe here ($p = 0.178$). See Bosnić & Demirdache 2020 for details.

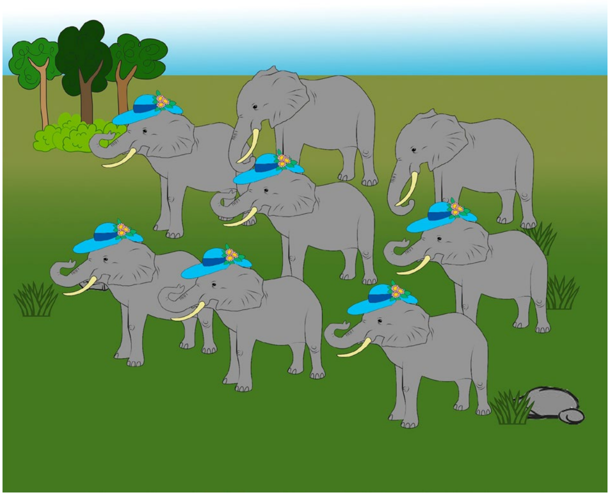

FIGURE 10 A nonexhausted (nonhomogeneous) scenario. Used with the three sentences shown in (16)

Crucially, we provided participants with three possible responses: yes, no, and a third response, ‘not really yes and not really no’, which we simply call a *gap answer* here.[19]

What would be the predictions for each of these conditions? The one–B nonexhausted scenario (figure 10) is a gap scenario: the group of elephants is not homogeneous with respect to the property of wearing a hat. This condition should thus allow us to verify our hypotheses, repeated here:

If *po* has universal quantificational force, then it should remove homogeneity effects in negated sentences with a definite subject. This means that negated sentences with *po*, such as (16b), will pattern like negated sentences with the universal quantifier *svaki*, such as (16c), both leading participants to overwhelmingly accept the one–B nonexhausted scenarios.

If, however, *po* is an event-plurality marker with no universal quantificational force of its own, then it should not remove homogeneity effects in negated sentences with a definite subject. This means that negated *po* sentences, such as (16b), should pattern like negated sentences with definite subjects only, such as (16a), yielding a gap answer (or a no answer) for the one–B nonexhausted scenarios.

5.2 | Results of experiment 3

Results are in figure 11. We only show the results for negated sentences, broken down by sentence type. The sentences with the universal quantifier *svaki* indeed showed an overwhelming acceptance for nonexhausted scenarios, as predicted. The sentences with definite plurals only

[19] We also had control items with clear yes and no answers, to check if participants were paying attention and understood that the gap answer could not be given in those cases. Further independent evidence to show that participants knew how to use the gap answer was that the *svaki* condition did not have a significant percentage of gap answers, which was as expected since universal quantifiers, unlike definite plurals, do not allow truth-value gaps.

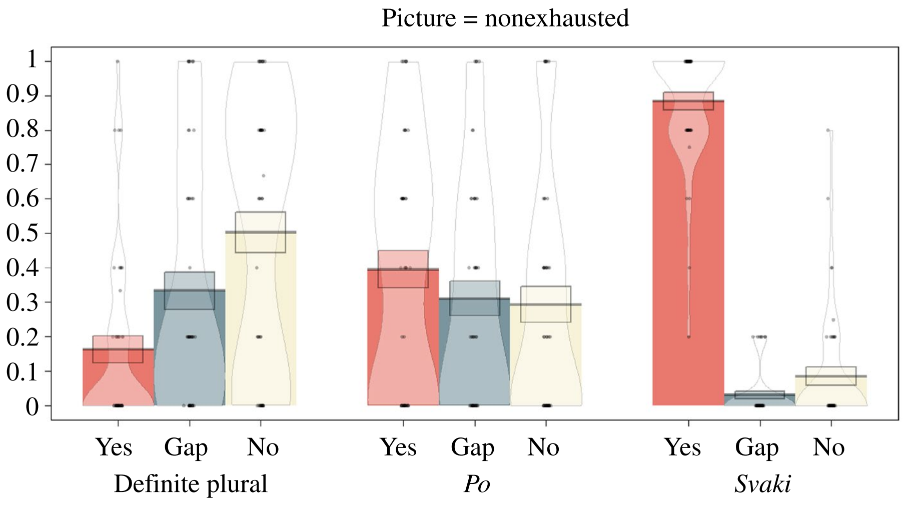

FIGURE 11 RDI plots of the results of experiment 3, with standard errors. The bars show the overall mean frequency of each of the three possible answers, yes, gap, and no, for the three types of sentences (from left to right): negative sentences with a definite plural only, negative sentences with *po*, and negative sentences with *svaki*. The lines show the distribution of answers, and each dot represents the mean of an individual participant's answers

were mostly rejected, which is predicted under the homogeneity hypothesis; in addition, they yielded gap answers, which is not the case for sentences with *svaki*. The results with *po* seem at first to be again inconclusive: the pattern is somewhere between that found with *svaki* and that found with the definite plural only.

Looking at the results for *po*, we can identify an hourglass shape of the yes answers of the individual participants, which, once again, suggests we should inspect this condition further. The population split we identified is shown in figure 12. This time the criterion we adopted for identifying the populations was a bit less strict: we grouped together participants who said yes 3/5, 4/5, or 5/5 times and participants who said yes 0/5, 1/5, or 2/5 times. (The less extreme graphs reflect the less strict criterion.)[20] In addition to applying this criterion to *po* sentences, it was also important to look at the definite-plural condition, because we had to eliminate people who accepted this condition systematically (only four participants). We ended up with two populations with distinct patterns of responses: (i) There is a group of participants (17/45) who treat *po* as having its own universal quantificational force. Their responses to *po* sentences show a ceiling acceptance rate and few gap answers, just like responses to *svaki*. This universal-quantification pattern is shown in figure 12a. (ii) There is a group of participants (24/45) who treat *po* as a mere event-plurality marker, with no universal quantificational force of its own. For them, *po* does not remove homogeneity effects, resulting in either a gap answer or a no answer, in very comparable proportions to the sentences with definite subjects, not the *svaki* sentences. This homogeneity (-effect) pattern is shown in figure 12b.

[20]The reason for the less strict criterion for the split is practical: we had to look at two conditions at the same time with a smaller sample of participants. If we were strict on both conditions, we would not be able to distinguish two populations. Future research may include a replication with a higher number of speakers and/or more observations per condition. The "middle" group is also an interesting point of discussion, as raised by an anonymous reviewer: could it be that the participants are oscillating between two meanings of *po*? Based on the current sample, we cannot discuss this further, but more research is called for to determine whether this group of participants is significantly represented.

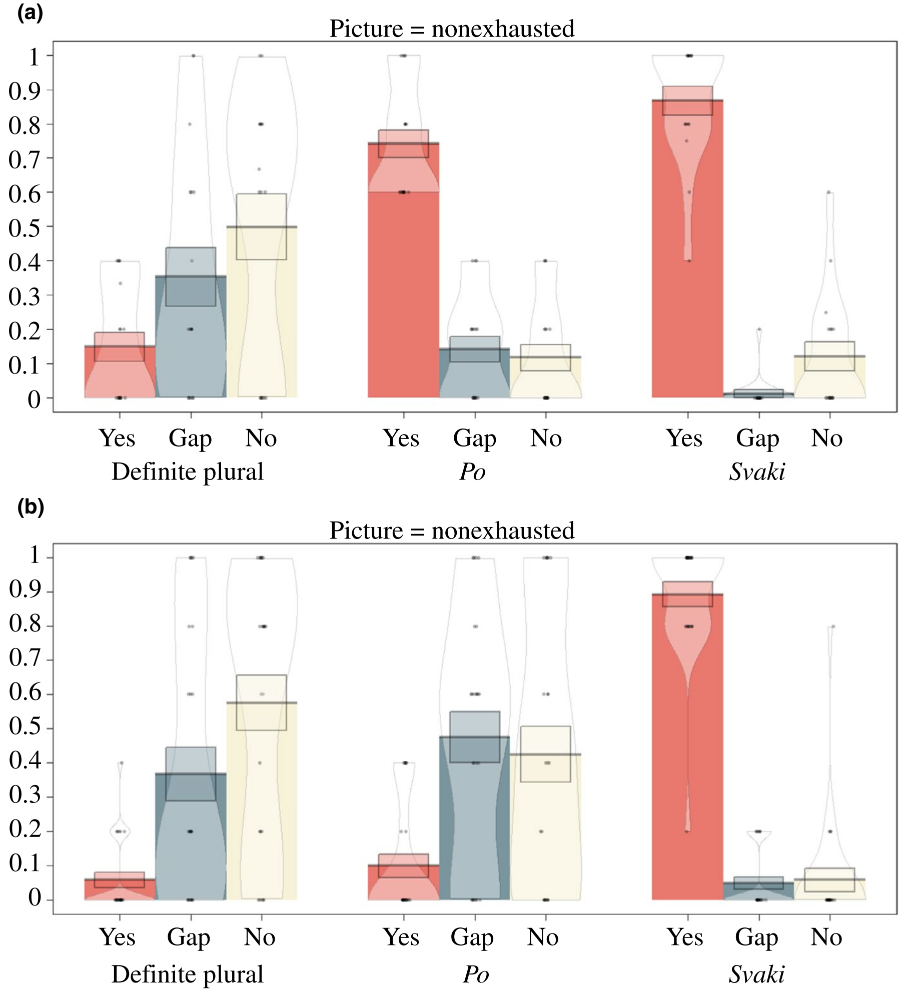

FIGURE 12 RDI plots of the two populations identified in experiment 3 based on responses to negative sentences with *po*. (a) Universal-quantification pattern of responses. (b) Homogeneity pattern of responses. Bars show the overall mean frequency of the three possible answers, lines show the distribution of answers, and each dot represents the mean of an individual participant's answers

For the statistical analysis, we again performed a repeated-measures mixed-effects logistic regression on the whole population using the glmer() function of the LME4 package (Bates et al. 2015) in R (R Core Team 2014), and we evaluated our models by comparing the Akaike Information Criterion values. We ran the models separately on different types of answers: yes and gap. The only independent variable with multiple levels was Sentence Type (negative *po*, negative definite, negative *svaki*), and the random factor Item was excluded in all the models because it did not result in a better fit. The best model for yes answers in the overall results was a random slope of Sentence Type for Subject. The reference level was set as negative *po*. People are less likely to judge negative definite sentences true than negative *po* sentences, and they are in general, not likely to accept this condition ($\beta = -1.890$, z value $= -3.335$, $p = 0.001$). As predicted, people are more likely to accept negative *svaki* sentences ($\beta = 3.281$, z value $= 6.114$, $p < 0.000$) both in general and compared to negative *po* sentences. The best model for gap answers was a random-intercept model.[21] Overall, people are more likely to choose gap answers for negative definite sentences than for negative *po* sentences, but this difference was not statistically significant ($\beta = 0.158$, z value $= 0.633$, $p = 0.526$). Of course, gap answers are significantly less likely to be chosen for negative *svaki* sentences ($\beta = -3.535$, z value $= -7.089$, $p < 0.000$). See tables C1 and C2 in appendix C for complete model outputs for both types of answers.

When looking at the best models for the two populations separately—universal quantification and homogeneity—there is nothing surprising that cannot be inferred from the graphs alone. However, it is interesting to point out that, in the homogeneity pattern, gap answers are more likely to be given (even if they are at chance) for negative *po* sentences than for negative definite sentences ($\beta = -0.739$, z value $= -2.150$, $p < 0.03$). The complete outputs for these models are given in tables C3–C6.

5.3 | Discussion of experiment 3

Our findings for experiment 2 led us to ask whether the exhaustivity effects found in *po* sentences were not in fact maximality effects arising via a definite interpretation of the bare plural subject. We proposed to use the distribution of homogeneity effects in negative contexts as a diagnostic to tease maximality and exhaustivity apart.

Note first that our findings provide a nice experimental confirmation from Serbian that the differences between the truth conditions of statements with definite plurals and statements with universally quantified expressions come out very clearly in negative statements. The Serbian sentence in (16c), 'Not every elephant is wearing a hat', was overwhelmingly judged true in the one–B scenario where some but not all elephants were wearing a hat; in contrast, the sentence in (16a), 'The elephants are not wearing a hat', was either judged false in this context (one expected response, since the sentence is true iff none of the elephants are wearing a hat) or neither false nor true (also an expected response, since definites but not universals yield truth-value-judgment gaps).

Experiment 3, moreover, revealed fascinating results for *po* sentences, validating *both* of the hypotheses we set out to confirm, thus providing experimental evidence for two populations with different interpretations of *po*. On the one hand, 17/45 participants show the pattern of response that is expected if *po* is a distributor with its own universal quantificational force (the universal-quantification pattern: no gap answers and ceiling acceptance of *po* sentences). On the other hand, 24/45 participants show the pattern of response that is expected if *po* is an event-plurality marker, with no universal quantificational force of its own (the homogeneity pattern: some gap answers and some no answers). In the concluding section of this article, we speculate that the coexistence of these two patterns suggests an ongoing diachronic process of semantic weakening from a distributor with universal force to a mere marker of event plurality.

How do our findings for experiment 3 bear on our previous findings? They confirm the coexistence of the two patterns of interpretation (universal quantification and event plurality) uncovered by experiment 1. They provide criteria for teasing exhaustivity and maximality apart: If a speaker says no to all nonexhausted conditions and has the universal-quantification pattern, rejection can be plausibly imputed to an exhaustivity requirement that *po* enforces. But if the speaker shows a homogeneity pattern, then rejection can no longer be imputed to the assumed

[21]The random-slope model had an Akaike Information Criterion value two points lower than the simpler random-intercept model, so the simpler model was chosen.

universal force of *po*. In that case it can plausibly be attributed to a maximal interpretation of the bare plural subject, because the distribution of homogeneity effects is statistically comparable across negative definite sentences and negative *po* sentences for this population.

6 | CONCLUSIONS AND OPEN QUESTIONS

We have presented here three experiments designed to investigate the interpretations of transitive sentences with *po* and thus empirically test competing theories of distributive-share markers: universal quantification over events versus event plurality.

With experiment 1, we sought to determine whether the findings for intransitive sentences in Bosnić, Spenader & Demirdache 2020 would carry over to transitive sentences: would the interpretation of transitive *po* sentences provide evidence that distribution must be exhaustive and can take place over nonatomic entities, be they entities bigger than atoms/singular individuals (i.e., pluralities of atomic individuals) or entities from noncount domains (i.e., time)? The results of experiment 1 provide evidence for two distinct populations: (i) speakers whose behavioral patterns of response reflect a universal-quantification interpretation of *po* where *po* distributes down to atoms (no sayers) or over bigger, nonatomic entities as well (yes-to-A sayers); (ii) speakers who say yes to all conditions. This response pattern can be taken to reflect either an event-plurality pattern of interpretation or a universal-quantification pattern of interpretation in which the subject DP (providing the distributive key for *po*) is construed indefinitely.

With experiment 2, we sought to test whether the yes-to-all pattern could indeed be explained away as a consequence of an indefinite interpretation of the bare plural subject, by forcing the subject to be interpreted as definite. The results do not provide the evidence that would allow us to subsume the yes-to-all pattern under the universal-quantification pattern. Rather, they suggest a different take on the data altogether, namely that the exhaustivity effects found in *po* sentences are in fact maximality effects arising via a definite interpretation of the bare plural subject.

With experiment 3, we sought to tease maximality and exhaustivity apart by testing homogeneity effects in negative sentences with plural definite descriptions and with universally quantified descriptions. The results confirm the findings of experiment 1. That is, we do indeed have two distinct populations: (i) speakers who have a universal-quantification pattern (ceiling acceptance of negative *po* sentences in nonhomogeneous contexts; no gap answers), who interpret *po* as a universal quantifier; and (ii) speakers who have a homogeneity pattern (giving a gap answer or a no answer for negative *po* sentences in nonhomogeneous contexts), who do not interpret *po* as a universal quantifier but as an event-plurality marker.

In sum, we found evidence for two populations with different interpretations of *po*: as a universal quantifier versus as an event-plurality marker. But how should we understand these results? We start by addressing three questions relating to this issue. We then address the difference in availability between atomic and nonatomic event-quantificational readings.

First, is there independent evidence that in other domains *po* has a pluractional meaning? There are uses and readings that are common to *po* across many Slavic languages and that are, interestingly, also characteristic of (types of) pluractional markers attested in a wide variety of other languages (Knežević 2015; for a typological overview of pluractional markers, see Cabredo Hofherr 2010, Cabredo Hofherr & Laca 2012): (i) *po* is a perfective verbal prefix indicating that the verb denotes multiple events (e.g., *po-gasiti svetla* 'turn off multiple lights in different locations, usually one by one' versus *u-gasiti svetla* 'turn off the lights, not necessarily in different locations, usually at once'); (ii) *po* occurs with reduplicated numerals, yielding temporal distribution (e.g., *dva po dva* 'two at a time'); (iii) *po* is also a locative distributive preposition (e.g., *po podu* 'scattered all over the floor' versus *na podu* 'on the floor'); and (iv) it can attach to adverbs (e.g., *po-malo* 'little by little') and adjectives (e.g., *po-velik* 'big-ish'). So, there are linguistic reasons to believe that *po* can also function in some contexts as a purely pluractional marker. It is therefore not unreasonable to conjecture that some speakers interpret *po* as having only distributive and plural content.

Second, given that these two populations with these two interpretations exist, what is the relationship between the two? A common explanation when two populations are discovered is that it might be evidence of ongoing language change. If that is the case, it seems most likely that for some speakers, *po* has lost its universal quantificational force and is simply interpreted as a weaker, pluractional marker. Supporting (though anecdotal) evidence comes from comments made by many of our Serbian informants. Frequently, when presented with sentences containing a *po*-marked object, participants would repeat back the sentences in a slightly different form, with the universal quantificational determiner *svaki* modifying the subject, that is, using the common *svaki–po* construction (see footnote 5). Because *svaki* is a universal quantifier that enforces an exhaustive interpretation of its distributive key, speakers may begin to reanalyze *po* as making a different, weaker contribution, treating it simply as a pluractional marker. For these speakers, if no *svaki* is present, there is no exhaustive import. This explanation would mean that our two populations are the consequence of an ongoing diachronic process of semantic weakening. Whether or not this explanation is correct could be experimentally investigated by testing interpretation preferences with respect to exhaustivity in a range of ages and across regions. We leave this for future research.

Third, how do we reconcile our results here, which conclusively show that we have two coexisting populations, with the results of Bosnić, Spenader & Demirdache 2020, which provide evidence that *po* (as well as Korean *-ssik*) is a universal quantifier across all participants when used in intransitive sentences? We found a significant population of event-plurality participants in the experiments with transitive sentences, so we expect that there must be a similar population split underlying our previous results. But why then did we not find evidence of event-plurality interpretations?

One possible explanation is that event-plurality responders actually have a universal-quantification interpretation of *po* and that their event-plurality responses in the transitive experiments are simply pragmatically driven, weakened interpretations of universal-quantifier *po* undergoing a diachronic change.

Our conjecture is that such participants (with a "surface" event-plurality pattern but an underlying universal-quantification pattern) only show evidence of a pluractional reading in contexts where they have no evidence that the speaker intended exhaustivity. That is to say, by considering potential alternative expressions that could have been said, these individuals make conclusions about what was actually said, a type of conversational implicature. In particular, *po* in transitive sentences appears to often co-occur with the unequivocally exhaustive quantifier *svaki* (see footnote 5). The co-occurrence of *svaki* with *po* (in a given sentence) serves to unambiguously signal to the hearer that distribution must be construed as exhaustive. But intransitive sentences have no linguistically expressed distributive-key argument that could co-occur with the (sole) *po* argument in the sentence to signal an exhaustive distributive interpretation. This could be why pragmatic weakening is not possible in intransitive sentences and why *po* thus retains its exhaustive import. Nevertheless, more research is needed to see if this explanation holds up.

Finally, returning to the other findings of the present article, consider just the population that interprets *po* as a universal quantifier. In our analysis of the experiment 1 results, we subsumed two separate patterns under universal quantification: that of participants who only accepted atomic distributive readings and that of participants who allowed event-distributive readings over nonatomic entities/groups. In experiment 2 and to a lesser degree in experiment 1 as well, event-distributive readings over groups appeared to be less available than distributive readings over (singular) individuals. Only four participants showed the yes-to-A pattern in experiment 2. Why is this?

Recall that since the granularity parameter for *po* is free, it can be set to atoms or to bigger-size entities (groups, sums/amounts of time, chunks of space). What determines how it gets set is the supporting context: what the subject judges to be salient in the (linguistic and visual) input provided. One reason for the poor representation of the yes-to-A pattern in experiment 2 might then be that, since the lead-in sentence introduced the characters by explicitly mentioning the cardinality of the superset (e.g., 'There are 12 monkeys ...'), it was harder to interpret the subsets as forming the relevant distributive key over which to distribute. According to other work on universal event quantifiers (Zimmermann 2002, Balusu & Jayaseelan 2013, Champollion 2016b), event-distributive readings are only felicitous with a supportive relevant context, and in fact an appropriate context may be essential for these readings. Our frame story did not have a context that encouraged an event-distributive reading (nor a context that discouraged it, for that matter). The event-distributive reading could be encouraged linguistically by using a different target sentence. Bosnić, Velich & Spenader 2020 reports on an experiment with the German distributive-share marker *jeweils*, testing transitive sentences whose object was marked with it. Like our participants, the German-speaking participants also preferred the distributive key to be set to atoms, finding atomic distributive readings more acceptable. In a second experiment, Bosnić, Velich & Spenader only tested sentences with singular indefinite subjects, in a Groups = four condition. This manipulation makes only event-distributive readings possible, since the only plurality that can be chosen for the distributive key is some implicit spatial argument, such as the multiple groups made salient in the visual context. Bosnić, Velich & Spenader found that about half the participants accepted this reading in a test sentence presented with no context. Thus, the German results, like the Serbian results, show that even though participants can clearly get event-distributive readings over nonatomic entities/groups, without a supportive context these readings are less available than individual distributive readings. The role of contextual support in nonatomic readings has hardly been addressed but is an intriguing topic for future work.

The focus of this article was the distributive marker *po* in Serbian. Only future research will tell us to what extent our experimental and theoretical findings carry over to other types of distributive-share markers reported crosslinguistically (in Korean, Japanese, Telugu, and other Slavic languages besides Serbian, among others). Preliminary investigations carried out with Korean and Telugu informants suggest that our experimental findings extend to a certain degree to such languages. We hope, however, to have shown here the strong theoretical relevance of experimental research on distributive-share markers such as *po*, complementing the work on distributive-key markers such as *each/every* and thus enhancing our understanding of distributivity across languages.

ACKNOWLEDGMENTS

Hamida Demirdache and Jennifer Spenader should be considered joint senior authors of this article. We thank Nataša Knežević, Boban Arsenijević, Seth Cable, Orin Percus, Bart Hollebrandse, Angeliek van Hout, Anamaria Fălăuş, Jakub Dotlačil, Anna de Koster, and Oana Lungu for invaluable discussions, advice, comments, and feedback at different stages of this study. We also thank our research groups and labs for useful meetings and discussions: the Nantes Linguistics Lab (Joint Research Unit 6310, French National Center for Scientific Research/University of Nantes) and the Acquisition Lab and the Semantics and Cognition group (University of Groningen). Conducting this research would not have been possible without numerous informants and participants in our experiments, whose input and judgments we greatly appreciate. This research was partly sponsored by Ammodo (Groningen–Nantes Van Gogh research grant).

DATA-AVAILABILITY STATEMENT

The original data discussed in this article are available from the corresponding author upon reasonable request.

ORCID

Ana Bosnić https://orcid.org/0000-0002-9117-4692

REFERENCES

Akaike, Hirotugu. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19.6.716–723.

Arsenijević, Boban. 2018. Atypical demonstratives in an articleless language. **In:** Marco Coniglio, Andrew Murphy, Eva Schlachter, and Tonjes Veenstra, editors. *Atypical demonstratives: Syntax, semantics and pragmatics*. Berlin, Germany: De Gruyter. 161–196.

Balusu, Rahul. 2006. Distributive reduplication in Telugu. **In:** Christopher Davis, Amy Rose Deal, and Youri Zabbal, editors. *NELS 36: Proceedings of the thirty-sixth annual meeting of the North East Linguistic Society*. Amherst, MA: Graduate Linguistics Student Association, University of Massachusetts Amherst. 39–52.

Balusu, Rahul, and K. A. Jayaseelan. 2013. Distributive quantification by reduplication in Dravidian. **In:** Kook-Hee Gil, Stephen Harlow, and George Tsoulas, editors. *Strategies of quantification*. Oxford, United Kingdom: Oxford University Press. 60–86.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.1.1–48.

Bošković, Željko. 2012. On NPs and clauses. **In:** Günther Grewendorf and Thomas Ede Zimmermann, editors. *Discourse and grammar: From sentence types to lexical categories*. Berlin, Germany: De Gruyter Mouton. 179–242.

Bosnić, Ana, and Hamida Demirdache. 2020. Mind the gap: Homogeneity effects with bare plurals and distributive markers in Serbian. Unpublished paper. **Presented at:** 15th annual meeting of the Slavic Linguistics Society, Indiana University, September 5.

Bosnić, Ana, Jennifer Spenader, and Hamida Demirdache. 2020. Dancing monkeys in Serbian and Korean—exhaustivity requirements on distributive share markers. *Glossa* 5.1.78.

Bosnić, Ana, Maximilian Velich, and Jennifer Spenader. 2020. Event-distributivity and exhaustivity: Insights into distributive share markers from experiments with German *jeweils*. **In:** Elena Tribushinina and Mark Dingemanse, editors. *Linguistics in the Netherlands* 37. Amsterdam, the Netherlands: John Benjamins Publishing Company. 2–22.

Brisson, Christine M. 1998. *Distributivity, maximality, and floating quantifiers*. Doctoral thesis. New Brunswick, NJ: Rutgers University.

Cable, Seth. 2014. Distributive numerals and distance distributivity in Tlingit (and beyond). *Language* 90.3.562–606.

Cabredo Hofherr, Patricia. 2010. Verbal plurality and event plurality. Course materials. DGfS-CNRS Summer School on Linguistic Typology, August 15–September 3. https://www.eva.mpg.de/lingua/conference/2010_summerschool/files/courses.html#cabredo.

Cabredo Hofherr, Patricia, and Brenda Laca, editors. 2012. *Verbal plurality and distributivity*. Berlin, Germany: Walter de Gruyter.

Cabredo Hofherr, Patricia, Jérémy Pasquereau, and Carolyn O'Meara. 2018. Event plurality in Seri. **In:** Kimberly Johnson and Alexander Göbel, editors. *SULA 10: Proceedings of the tenth conference on the Semantics of Under-represented Languages in the Americas*. Amherst, MA: Graduate Linguistics Student Association, University of Massachusetts Amherst. 1–15.

Champollion, Lucas. 2012. *Each* vs. *jeweils*: A cover-based view on distance distributivity. **In:** Maria Aloni, Vadim Kimmelman, Floris Roelofsen, Galit W. Sassoon, Katrin Schulz, and Matthijs Westera, editors. *Logic, language and meaning*. Berlin, Germany: Springer Berlin Heidelberg. 251–260.

Champollion, Lucas. 2016a. Covert distributivity in algebraic event semantics. *Semantics and Pragmatics* 9.15.

Champollion, Lucas. 2016b. Overt distributivity in algebraic event semantics. *Semantics and Pragmatics* 9.16.

Choe, Jae-Woong. 1987. *Anti-quantifiers and a theory of distributivity*. Doctoral thesis. Amherst, MA: University of Massachusetts Amherst.

É. Kiss, Katalin, Mátyás Gerőcs, and Tamás Zétényi. 2013. Preschoolers' interpretation of doubly quantified sentences. *Acta Linguistica Hungarica* 60.2.143–171.

É. Kiss, Katalin, and Tamás Zétényi. 2018. Linguistic pathway to multiplication. *Acta Linguistica Academica* 65.2–3.201–225.

Faller, Martina. 2001. The problem of Quechua *-nka*: Distributivity vs. group forming. **In:** Ji-Yung Kim and Adam Werle, editors. *The proceedings of SULA: The Semantics of Under-represented Languages in the Americas*. Amherst, MA: Graduate Linguistics Student Association, University of Massachusetts Amherst. 38–46.

Gil, David. 1982. *Distributive numerals*. Doctoral thesis. Los Angeles, CA: University of California, Los Angeles.

Gil, David. 1990. Markers of distributivity in Japanese and Korean. **In:** Hajime Hoji, editor. *Japanese/Korean Linguistics* 1. Stanford, CA: Center for the Study of Language and Information. 385–393.

Gil, David. 1995. Universal quantifiers and distributivity. **In:** Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee, editors. *Quantification in natural languages*. Dordrecht, the Netherlands: Springer Netherlands. 321–362.

Gillon, Carrie. 2014. *The semantics of determiners: Domain restriction in Sḵwx̱wú7mesh*. Newcastle upon Tyne, United Kingdom: Cambridge Scholars Publishing.

Knežević, Nataša. 2012. Serbian distributive children. **In:** Joanna Błaszczak, Bożena Rozwadowska, and Wojciech Witkowski, editors. *Current issues in generative linguistics: Syntax, semantics, and phonology*. Wrocław, Poland: Center for General and Comparative Linguistics. 26–41.

Knežević, Nataša, 2015. *Numerals and distributivity in Serbian: At the syntax-semantics-acquisition interface*. Doctoral thesis. Nantes, France: University of Nantes.

Knežević, Nataša, and Hamida Demirdache. 2017. On the acquisition of distributive numerals in Serbian. **In:** Jiyoung Choi, Hamida Demirdache, Oana Lungu, and Laurence Voeltzel, editors. *Language acquisition at the interfaces: Proceedings of GALA 2015*. Newcastle upon Tyne, United Kingdom: Cambridge Scholars Publishing. 83–107.

Knežević, Nataša, and Hamida Demirdache. 2018. Universal quantification and distributivity in Serbian child language. **In:** Katalin É. Kiss and Tamás Zétényi, editors. *Linguistic and cognitive aspects of quantification*. Dordrecht, the Netherlands: Springer International Publishing. 115–140.

Križ, Manuel. 2015. Homogeneity, non-maximality, and *all*. *Journal of Semantics* 33.3.493–539.

Križ, Manuel. 2017. In Soviet Russia, alcohol is dependent on you. **In:** Clemens Mayr and Edwin Williams, editors. *Festschrift für Martin Prinzhorn.* (*Wiener Linguistische Gazette* 82.) Vienna, Austria: Department of Linguistics, University of Vienna. 173–180.

Križ, Manuel, and Emmanuel Chemla. 2015. Two methods to find truth-value gaps and their application to the projection problem of homogeneity. *Natural Language Semantics* 23.3.205–248.

Kuhn, Jeremy. 2017. Dependent indefinites: The view from sign language. *Journal of Semantics* 34.3.407–446.

Lasersohn, Peter. 1995. *Pluralities, conjunction and events*. Dordrecht, the Netherlands: Springer Netherlands.

Lasersohn, Peter. 1998. Generalized distributivity operators. *Linguistics and Philosophy* 21.1.83–93.

Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75.3.522–551.

Link, Godehard. 1987. Generalized quantifiers and plurals. **In:** Peter Gärdenfors, editor. *Generalized quantifiers*. Dordrecht, the Netherlands: Springer Netherlands. 151–180.

Malamud, Sophia A. 2012. The meaning of plural definites: A decision-theoretic approach. *Semantics and Pragmatics* 5.3.

Matthewson, Lisa. 2000. On distributivity and pluractionality. **In:** Brendan Jackson and Tanya Matthews, editors. *Proceedings from Semantics and Linguistic Theory X*. Ithaca, NY: Cornell University. 98–114.

McKercher, David, and Yookyung Kim. 2000. What does *ssik* in Korean really mean? **In:** Mineharu Nakayama and Charles J. Quinn Jr., editors. *Japanese/Korean Linguistics* 9. Stanford, CA: CSLI Publications. 239–252.

Müller, Ana, and Esmeralda Negrão. 2012. On distributivity in Karitiana. **In:** Patricia Cabredo Hofherr and Brenda Laca, editors. *Verbal plurality and distributivity*. Berlin, Germany: Walter de Gruyter. 159–184.

Oh, Sei-Rang. 2001. Distributivity in an event semantics. **In:** Rachel Hastings, Brendan Jackson, and Zsofia Zvolenszky, editors. *Proceedings from Semantics and Linguistic Theory XI*. Ithaca, NY: Cornell University. 326–345.

Oh, Sei-Rang. 2006. *Plurality markers across languages*. Doctoral thesis. Storrs, CT: University of Connecticut.

Progovac, Ljiljana. 1994. *Negative and positive polarity: A binding approach*. Cambridge, United Kingdom: Cambridge University Press.

Przepiórkowski, Adam. 2015. A weakly compositional analysis of distance distributivity in Polish. **In:** Małgorzata Szajbel-Keck, Roslyn Burns, and Darya Kavitskaya. *Annual workshop on Formal Approaches to Slavic Linguistics: The first Berkeley meeting*. Ann Arbor, MI: Michigan Slavic Publications. 262–281.

R Core Team. 2014. *R: A language and environment for statistical computing. Software*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.

Safir, Ken, and Tim Stowell. 1988. Binominal *each*. **In:** James Blevins and Julie Carter, editors. *Proceedings of NELS 18*. Amherst, MA: Graduate Linguistics Student Association, University of Massachusetts Amherst. 426–450.

Schwarz, Florian. 2013. Maximality and definite plurals—experimental evidence. **In:** Emmanuel Chemla, Vincent Homer, and Grégoire Winterstein. editors. *Sinn und Bedeutung 17: Proceedings*. Paris, France: École Normale Supérieure. 509–526.

Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Tieu, Lyn, Manuel Križ, and Emmanuel Chemla. 2019. Children's acquisition of homogeneity in plural definite descriptions. *Frontiers in Psychology* 10.2329.

Zimmermann, Malte. 2002. *Boys buying two sausages* each: *On the syntax and semantics of distance-distributivity*. Utrecht, the Netherlands: Netherlands National Graduate School of Linguistics.

How to cite this article: Bosnić, Ana, Hamida Demirdache, and Jennifer Spenader. 2022. Exhaustivity and homogeneity effects with distributive-share markers: Experimental evidence from Serbian *po*. *Syntax* 25.1–38. https://doi.org/10.1111/synt.12223

APPENDIX A

Statistical analysis of experiment 1 results

TABLE A1 Output of best model, whole population. Formula: Answer ~ Picture ∗ Groups + (1 + Picture | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Four–A (intercept)	1.413	0.581	2.432	.015 *
Four–C	−4.915	0.594	−8.279	.000 ***
Four–B	−4.232	0.555	−7.632	.000 ***
One–A	6.855	0.796	8.613	.000 ***
One–C	−5.671	0.847	−6.693	.000 ***
One–B	−5.772	0.844	−6.841	.000 ***

TABLE A2 Tukey contrasts. Results are given on the logit scale (not the response scale). Confidence level used: 0.95; *p*-value adjustment: Tukey method for comparing a family of six estimates

Contrast	Estimate	Standard error	*z* value	*p* value
Four–A : four–C	4.9151709	0.5937260	8.279	<.0001
Four–A : four–B	4.2324117	0.5545881	7.632	<.0001
Four–A : one–A	−6.8551681	0.7959551	−8.613	<.0001
Four–A : one–C	3.7314045	0.5542638	6.732	<.0001
Four–A : one–B	3.1496332	0.5252151	5.997	<.0001
Four–C : four–B	−0.6827592	0.5024893	−1.359	.7517
Four–C : one–A	−11.7703391	1.0108700	−11.644	<.0001
Four–C : one–C	−1.1837664	0.2891111	−4.095	.0006
Four–C : one–B	−1.7655377	0.4697285	−3.759	.0024
Four–B : one–A	−11.0875799	0.9921515	−11.175	<.0001
Four–B : one–C	−0.5010072	0.4539321	−1.104	.8801
Four–B : one–B	−1.0827785	0.2779503	−3.896	.0014
One–A : one–C	10.5865727	0.9886908	10.708	<.0001
One–A : one–B	10.0048013	0.9766699	10.244	<.0001
One–C : one–B	−0.5817713	0.4173874	−1.394	.7308

TABLE A3 Output of best model, yes sayers. Formula: Answer ~ Picture + Groups + (1 + Picture | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Four–A (intercept)	5.678	1.423	3.991	.000 ***
Four–C	−7.250	1.479	−4.901	.000 ***
Four–B	−6.326	1.426	−4.436	.000 ***
One–A	1.038	0.233	4.451	.000 ***

TABLE A4 Output of best model, no sayers. Formula: Answer ~ Picture ∗ Groups + (1 + Picture | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Four–A (intercept)	−2.561	0.372	−6.879	.000 ***
Four–C	−5.272	2.658	−1.983	.047 *
Four–B	−4.623	2.646	−1.747	.081
One–A	5.979	0.658	9.086	.000 ***
One–C	−5.011	1.064	−4.707	.000 ***
One–B	−5.159	1.150	−4.487	.000 ***

APPENDIX B

Statistical analysis of experiment 2 results (definiteness follow-up with and without *po*)

TABLE B1 Output of best model. Formula: Answer ~ Picture ∗ Groups ∗ Sentence Type + (1 + Picture + Groups + Sentence Type | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Four–A with *po* (intercept)	−1.387	0.520	−2.667	.008 **
Four–B with *po*	−3.733	0.653	−5.712	.000 ***
One–A with *po*	8.343	0.906	9.208	.000 ***
Four–A without *po*	0.197	0.440	0.448	.654
One–B with *po*	−5.383	0.983	−5.475	.000 ***
Four–B without *po*	1.487	0.499	2.981	.003 **
One–A without *po*	−2.923	0.894	−3.269	.001 **
One–B without *po*	2.070	1.000	2.070	.038 *

APPENDIX C

Statistical analysis of experiment 3 results (homogeneity effects with *po*)

TABLE C1 Output of best model, yes answers, whole population. Formula: Answer_yes ~ Sentence Type + (1 + Sentence Type | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Negative *po* (intercept)	−0.883	0.427	−2.068	.039 *
Negative definite	−1.890	0.567	−3.335	.001 ***
Negative *svaki*	3.281	0.537	6.114	.000 ***

TABLE C2 Output of best model, gap answers, whole population. Formula: Answer_gap ~ Sentence Type + (1 | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Negative *po* (intercept)	−1.294	0.356	−3.629	.000 ***
Negative definite	0.159	0.250	0.633	.527
Negative *svaki*	−3.536	0.499	−7.089	.000 ***

TABLE C3 Output of best model, yes answers, universal-quantification pattern. Formula: Answer_yes ~ Sentence Type + (1 | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Negative *po* (intercept)	1.036	0.248	4.175	.000 ***
Negative definite	−2.814	0.399	−7.057	.000 ***
Negative *svaki*	0.856	0.408	2.101	.036 *

TABLE C4 Output of best model, gap answers, universal-quantification pattern. Formula: Answer_gap ~ Sentence Type + (1 | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Negative *po* (intercept)	−2.265	0.501	−4.517	.000 ***
Negative definite	1.472	0.437	3.371	.001 ***
Negative *svaki*	−2.780	1.068	−2.602	.009 **

TABLE C5 Output of best model, yes answers, homogeneity pattern. Formula: Answer_yes ~ Sentence Type + (1 | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Negative *po* (intercept)	−2.739	0.469	−5.841	.000 ***
Negative definite	−0.650	0.521	−1.246	.213
Negative *svaki*	5.384	0.638	8.435	.000 ***

TABLE C6 Output of best model, gap answers, homogeneity pattern. Formula: Answer_gap ~ Sentence Type + (1 | Subject)

	Estimate	Standard error	*z* value	Pr(>|*z*|)
Negative *po* (intercept)	−0.269	0.506	−0.532	.595
Negative definite	−0.739	0.344	−2.150	.032 *
Negative *svaki*	−4.146	0.612	−6.769	.000 ***