Slavic derived collective nouns as spatial and social clusters

Marcin Wągiel
Masaryk University in Brno

1 Introduction

A puzzling property of collective nouns is that they simultaneously evoke a sense of plurality and singularity. Though it is commonly assumed that they are specific to the domain of individuals, e.g., *committee of women,* in fact the category is much more general and covers also abstract objects such as eventualities, e.g., *series of murders,* and numbers, e.g., *sequence of integers.* In this paper, I argue that Slavic derivational morphology reflects two modes of collectivity. In particular, I examine two types of Slavic derived collectives exemplified by the Polish nouns in (1) and (2), which I will call SPATIAL and SOCIAL COLLECTIVES, respectively.[1]

(1) *kwiat* ⇒ *kwieci-e*
　　flower.SG　flower-COLL
　　'flower'　　'clump of flowers'

(2) *duchowny* ⇒ *duchowień-stwo*
　　priest.SG　priest-COLL
　　'priest'　　'collective of priests, clergy'

The main claim is that both types of Slavic derived collective nouns denote properties of clusters, i.e., structured configurations of entities. I will argue that spatial collectives (1) are predicates true of spatial clusters, i.e., topological arrangements of objects in physical space, whereas social collectives (2) are predicates true of social clusters, i.e., abstract configurations of roles that individuals can bear in social space.

2 Two types of collectives

For a long time, it was standardly assumed that collective nouns constitute a uniform category (Landman, 1989; Barker, 1992; Schwarzschild, 1996). However, recent findings suggest that there are two different kinds of such expressions: social collectives designating organizations based on membership, e.g., *committee (of women),* and spatial collectives referring to topologicals configurations, e.g., *pile (of dishes)* (Pearson, 2011; de Vries, 2015; Henderson, 2017; Zwarts, 2020).

A number of diagnostics to distinguish the two types have been proposed in the literature: (i) plural agreement in British and Canadian English (3), (ii) ability to antecede plural pronouns (4), (iii) embedding in partitive constructions (5), (iv) quantificational domain of *half* (6): (6a) quantifies over individual committee members, whereas (6b) quantifies over any part of the bouquet (not only individual flowers), (v) reference to larger cardinalities (7), (vi) truth conditions of negated existential statements (8), (vii) compatibility with spatial modifiers (9) and (viii) compatibility with certain expressions like the Dutch noun *lid* 'member' (10).

(3) a.　The committee are old.　　　　　　　　　　　　(Barker, 1992)
　　b.　* The bunch of flowers are tall.　　　　　　　　(Pearson, 2011)

(4) a.　The committee is in the backyard. They are by the river.
　　b.　The bouquet is in the backyard. #They are by the river.　(Henderson, 2017)

---

[1]The orthographic differences between the singular and the collective forms in (1)–(2) represent standard morphophonological alternations in Polish.

(5) a. Three of the committee came to the meeting.

    b. * Three of the bunch of flowers had died.   (Pearson, 2011)

(6) a. Half of the committee had been painted yellow.

    b. Half of the bunch of flowers had been painted yellow.   (Pearson, 2011)

(7) a. Bill needs to learn to cook for a family of two.

    b. # John planted a grove of two redbud trees.   (Henderson, 2017)

(8) a. Each member of the committee travels to a different state to visit family.
       ⊭ The committee no longer exists.

    b. Someone takes each flower from the bouquet and places it in a different room of the house. ⊭ The bouquet no longer exists.   (Henderson, 2017)

(9) a. ? *midden in een comité*
       middle in a   committee
       Intended: 'in the middle of a committee'

    b. *midden in een menigte*
       middle in a   crowd
       'in the middle of a crowd'   Dutch (Zwarts, 2020)

(10) a. *Anna is een lid   van het comité.*
       Anna is a   member of   the committee
       'Anna is a member of the committee.'

    b. ? *Anna is een lid   van de menigte.*
       Anna is a   member of   the crowd
       'Anna is a member of the crowd.'   Dutch (Zwarts, 2020)

It has been observed, however, that (in)animacy can play a role in the behavior of collectives (de Vries, 2015) and it turns out that many of the tests above distinguish between animate and inanimate collections rather than between spatial and social collections. For instance, *crowd* designates a spatial configuration and yet in British English it does allow for plural agreement (11) and licenses plural anaphora (12), unlike *bouquet*. Furthermore, as observed by Zwarts (2020), (13) patterns with (6a) in that it can only quantify over persons (and not arbitrary parts thereof). Finally, collectives such as *fleet (of trucks)* and *network (of computers)* designate higher-order configurations of objects, which are based on function rather than spatial proximity, and thus are examples of inanimate social collectives. Interestingly, unlike (7b), they do not require large pluralities (14) and, unlike (8b), do not give rise to spatial existential inferences.

(11)   The crowd are cheerful.

(12)   The crowd is in the backyard. They are by the river.

(13)   Half of the crowd had been painted yellow. ⊭ Half of the people.   (Zwarts, 2020)

(14)   The company owns a fleet of two trucks for unexpected deliveries.

(15)   Each truck from the fleet travels to a different state to deliver goods.
       ⊭ The fleet no longer exists.

Therefore, I conclude that out of the tests in (i)–(viii) only (v) reference to larger cardinalities, (vi) truth conditions of negated existential statements, (vii) compatibility with spatial modifiers and (viii) compatibility with expressions like Dutch *lid* 'member' are reliable diagnostics for distinguishing social and spatial collectives. Moreover, the data in (11–15) show that (in)animacy is orthogonal to the social/spatial distinction and in fact there are two dimensions of collectivity illustrated in Tab. 1 (cf. Zwarts 2020 for a similar classification though without specifying social inanimate collections).

Table 1: Dimensions of collectivity

	SPATIAL collections	SOCIAL collections
ANIMATE collections	crowd, swarm	club, committee
INANIMATE collections	bouquet, pile	fleet, network

3 Slavic derived collectives

Additional evidence in favor of the relevance of the distinction between spatial and social collections comes from Slavic derivational morphology. Slavic languages have a rich inventory of collectivizing affixes. I argue that different classes of such morphemes correspond semantically to the spatial/social distinction. This will be demonstrated on two types of derived collectives attested in every branch of Slavic Tab. 2 and 3.

What all of the collective forms in Tab. 2 have in common is that they (i) are derived from inanimate nouns, (ii) are morphosyntactically singular, (iii) denote large collections of objects conceptualized as aggregates, i.e., topological configurations of entities that either touch each other or remain in close proximity. For instance, Czech *rákosí* does not denote arbitrary reeds but rather a reed bed, Russian *listva* means 'foliage', Macedonian *snopje* 'bundle of sheaves'.

Table 2: Slavic derived spatial collectives

	GLOSS	SINGULAR	PLURAL	COLLECTIVE
Polish	'flower'	kwiat	kwiaty	kwiecie
Czech	'reed'	rákos	rákosy	rákosí
Slovak	'rock'	kameň	kamene	kamenie
Russian	'leaf'	list	list'ja	listva
BCMS	'flower'	cvet	cvetovi	cveće
Macedonian	'sheaf'	snop	snopovi	snopje
Slovenian	'bush'	grm	grmi	grmovje

The collectives in Tab. 3 (i) are typically derived from human nouns denoting social roles and capacities associated with profession, social class and status, (ii) are morphosyntactically singular, (iii) usually designate pluralities of individuals but (iv) can also refer to institutions or organizations independent from their members. For instance, Czech *učitelstvo* and Slovak *študentstvo* refer to a body of teachers and students, respectively.

Table 3: Slavic derived social collectives

	GLOSS	SINGULAR	PLURAL	COLLECTIVE
Polish	'priest'	duchowny	duchowni	duchowieństwo
Czech	'teacher'	učitel	učitelé	učitelstvo
Slovak	'student'	študent	študenti	študentstvo
Russian	'soldier'	voin	voiny	voinstvo
BCMS	'worker'	radnik	radnici	radništvo
Macedonian	'citizen'	graǵanin	graǵani	graǵanstvo
Slovenian	'leader'	vodja	vodji	vodstvo

Based on Polish data, I propose four diagnostics for distinguishing semantically derived spatial and social collectives: (i) compatibility with VPs headed by the verb *należeć* 'belong' (3),

(ii) existence independent from the constituent members (17), (iii) behavior with PPs headed by *wśród* 'among, amid' (18): (18a) can mean that one of the spotted priests is intriguing (rather than an intriguing non-priest was spotted surrounded by priests), whereas (18b) only means that the spotted thing is not a flower; and finally (iv) compatibility with kind predicates (19).

(16) a. *Ten mężczyzna należy   do duchowieństwa.*
          this man   belongs to priest.COLL.GEN
          'This man belongs to the clergy.'

    b. # *Ta niezapominajka należy   do kwiecia.*
          this forget.me.not   belongs to flower.COLL.GEN
          Intended: 'This forget-me-not belongs to the clump of flowers.'   Polish

(17) a. *Obecnie nikt   nie należy do duchowieństwa.*
          currently no.one NEG belongs to priest.COLL.GEN
          'Currently, no one belongs to the clergy.'

    b. # *Obecnie nic   nie jest częścią kwiecia.*
          currently nothing NEG is   part   flower.COLL.GEN
          Intended: 'Currently, nothing is part of the clump of flowers.'   Polish

(18) a. *Ania zauważyła kogoś   intrygującego wśród duchowieństwa.*
          Ania spotted   someone intriguing   among priest.COLL.GEN
          'Ania spotted someone intriguing among the clergy.'

    b. *Ania zauważyła coś   intrygującego wśród kwiecia.*
          Ania spotted   something intriguing   among flower.COLL.GEN
          'Ania spotted something intriguing amid the clump of flowers.'   Polish

(19) a. *Duchowieństwo było powszechne w XX   wieku.*
          priest.COLL   was widespread in 20th century
          'Clergy was widespread in the 20th century.'

    b. # *Kwiecie   było powszechne w trzeciorzędzie.*
          flower.COLL was widespread in Tertiary
          Intended: 'Flowers were widespread in the Tertiary Period.'   Polish

Based on the examined data, I posit the following generalization. Slavic collectivizing suffixes form structured collections of entities denoted by the noun root. The suffixes that attach to inanimate entities (Tab. 2) demonstrably make reference to clusters, i.e., topological configurations of objects in close proximity. The suffixes that attach to human nouns (Tab. 3) refer to structured social institutions, i.e., organizations whose members perform similar roles.

4   Mereotopology

In order to account for pluralities arranged in a structured manner, I follow Grimm (2012) and adopt MEREOTOPOLOGY, a theory of wholes extending standard mereology with topological notions (Casati & Varzi, 1999; Varzi, 2007). In mereotopology, CONNECTEDNESS (C) allows for capturing a configuration in which two entities share a boundary. Given C, it is possible to define more complex mereotopological notions to capture subtle distinctions between different spatial configurations. Inspired by Grimm (2012), I propose a revised formulation of the property TRANSITIVELY CONNECTED (TC) (20a), which determines whether two objects are connected through a series of mediating entities. TC allows for defining the concept of CLUSTER (CLSTR) (20b). An entity x is a cluster relative to a connection relation C and a property P iff x is a sum of entities falling under the same property which are all transitively connected relative to some set Y under the same property and connection relation.

(20)   a.   For a finite sequence $Z = \langle z_1, \ldots, z_n \rangle$, $\text{TC}(x, y, P, C, Z)$ holds iff
       $z_1 = x, z_n = y$, $\text{C}(z_i, z_{i+1})$ holds for $1 \leq i < n$ and $P(z_i)$ holds for $1 \leq i \leq n$.

   b. $\text{CLSTR}_C(P)(x) \overset{\text{def}}{=} \exists Z[x = \bigsqcup Z \land \forall z \forall z' \in Z \exists Y \subseteq Z[\text{TC}(z, z', P, C, Y)]]$

5   Roles

I follow Zobel (2017) in extending ontology with social roles (type r), which are modeled as independent ontological objects. While class nouns such as *cat* denote properties of individuals (type $\langle e, t \rangle$), role nouns denote properties of roles (type $\langle r, t \rangle$). In addition, an individual can be associated with a role via the operation PLAY (5), which takes a set of roles P and yields a set of (potentially plural) individuals x for which there are a role r and an eventuality e such that r is a P-role and $\langle r, e \rangle$ is part of the specific role structure $\mathscr{R}_x$ of x.

(21) $[\![\text{PLAY}]\!] = \lambda P_{\langle r, t \rangle} \lambda x_e \exists r_r \exists e_v [P(r) \land \langle r, e \rangle \in \mathscr{R}_x]$ (Zobel, 2017)

Furthermore, I observe that roles, just like ordinary individuals, enter part-whole relations and can form pluralities. The evidence comes from the behavior of conjunction within *as*-phrases, e.g., (22a) can get either the distributive (Paul earns 8,000 euros in total) or the non-distributive reading (Paul earns a total of 4,000 euros in total), similarly to (22b).

(22)   a.   Paul earns 4,000 euros as a judge and a lecturer.
       b.   Paul gave 4,000 euros to Tom and Amy.

6   Collectives as properties of clusters

Since the derivational processes yielding derived spatial collectives belong to a larger class that should receive a unified semantics, I postulate that all derivational suffixes for collective nouns involve the notion of a cluster in some way. Together with the independently motivated idea that some collective nouns denote predicates of social role pluralities, this entails that some collective nouns involve clusters in social space. For instance, the Polish suffix *-e* (23a) takes an inanimate class noun (23b) and yields a predicate true of spatial clusters (23c). On the other hand, the suffix *-stwo* (24a) selects for social nouns (24b) and returns a property of social clusters (24c), i.e., conglomerates of closely related roles in social space. If needed, (24c) can easily be associated with a plurality of individuals via the PLAY operator (5).

(23)   a. $[\![\text{-e}]\!] = \lambda P_{\langle e, t \rangle} \lambda x_e [\text{CLSTR}_{\text{SP}}(P)(x)]$
       b. $[\![\text{kwiat}]\!] = \lambda x_e [\text{FLOWER}(x)]$
       c. $[\![\text{kwiecie}]\!] = \lambda x_e [\text{CLSTR}_{\text{SP}}(\text{FLOWER})(x)]$
(24)   a. $[\![\text{-stwo}]\!] = \lambda P_{\langle r, t \rangle} \lambda r_r [\text{CLSTR}_{\text{SC}}(P)(r)]$
       b. $[\![\text{duchowny}]\!] = \lambda r_r [\text{PRIEST}(r)]$
       c. $[\![\text{duchowieństwo}]\!] = \lambda r_r [\text{CLSTR}_{\text{SC}}(\text{PRIEST})(r)]$

The proposed analysis has a number of advantages. It provides a unified treatment for both types of collectives since both types of the discussed collectivizing suffixes denote predicate modifiers yielding a property of clusters. At the same time, it explains the differences between spatial and social collectives. While the former refer to clusters of entities, the latter denote clusters of roles, which as abstract objects have different properties than individuals performing them, e.g., an institution does not necessarily cease to exist if it temporarily has no members. Finally, the proposal shows how mereotopology can be extended to abstract domains.

References

Barker, Chris. 1992. Group terms in English: Representing groups as atoms. *Journal of Semantics* 9(1). 69–93. doi:10.1093/jos/9.1.69.

Casati, Roberto & Achille C. Varzi. 1999. *Parts and places: The structures of spatial representation*. Cambridge, MA: MIT. `https://mitpress.mit.edu/books/parts-and-places`.

Grimm, Scott. 2012. *Number and individuation*: Stanford University dissertation. `http://www.sas.rochester.edu/lin/sgrimm/publications.html`.

Henderson, Robert. 2017. Swarms: Spatiotemporal grouping across domains. *Natural Language & Linguistic Theory* 35(1). 161–203. doi:10.1007/s11049-016-9334-z.

Landman, Fred. 1989. Groups, I. *Linguistics and Philosophy* 12(5). 559–605. doi:10.1007/BF00627774.

Pearson, Hazel. 2011. A new semantics for group nouns. In Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer & Barbara Tomaszewicz (eds.), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, 160–168. Somerville, MA: Cascadilla. `http://www.lingref.com/cpp/wccfl/28/index.html`.

Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer. doi:10.1007/978-94-017-2704-4.

Varzi, Achille C. 2007. Spatial reasoning and ontology: Parts, wholes, and locations. In Marco Aiello, Ian E. Pratt-Hartmann & Johan van Benthem (eds.), *Handbook of spatial logics*, 945–1038. Berlin: Springer. doi:10.1007/978-1-4020-5587-4_15.

de Vries, Hanna. 2015. *Shifting sets, hidden atoms: The semantics of distributivity, plurality and animacy*: Utrecht University dissertation. `http://dspace.library.uu.nl/handle/1874/312186`.

Zobel, Sarah. 2017. The sensitivity of natural language to the distinction between class nouns and role nouns. In Dan Burgdorf, Jacob Collard, Sireemas Maspong & Brynhildur Stefáns-dóttir (eds.), *Proceedings of Semantics and Linguistic Theory 27*, 438–458. Ithaca, NY: CLC. doi:10.3765/salt.v27i0.4182.

Zwarts, Joost. 2020. Contiguity and membership and the typology of collective nouns. In Michael Franke, Nikola Kompa, Mingya Liu, Jutta L. Mueller & Juliane Schwab (eds.), *Proceedings of Sinn und Bedeutung 24*, 539–554. Osnabrück: Osnabrück University. `https://semanticsarchive.net/Archive/mZhNDA4Y/`.