Studia Linguistica Universitatis Iagellonicae Cracoviensis 137 (2020): 245–258
doi:10.4467/20834624SL.20.020.12982
www.ejournals.eu/Studia-Linguistica

DAMIAN HERDA
damian.t.herda@gmail.com
Jagiellonian University in Kraków
ORCID: 0000-0002-0925-1805

ARRANGEMENT CLASSIFIERS, COLLOCATIONS, AND NEAR-SYNONYMY: A CORPUS-BASED STUDY WITH REFERENCE TO POLISH

Keywords: arrangement classifiers, collocations, near-synonymy, lexical semantics, Polish

Abstract

The aim of this paper is to contrast the near-synonymous Polish classifiers *kupa* 'heap', *sterta* 'pile', and *stos* 'stack', all of which encode upward-oriented arrangements of objects or substances and thus prototypically combine with concrete inanimate nouns, by means of a collocational analysis conducted on naturally-occurring data derived from the National Corpus of Polish. The results of the empirical investigation point to a tendency for *kupa* 'heap' to combine predominantly with mass nouns denoting amorphous, frequently natural, stuff, whereas *sterta* 'pile' and *stos* 'stack' exhibit a pronounced predilection for count N2-collocates referring to artefacts. In a similar vein, while both *sterta* 'pile' and *stos* 'stack' typically stand for aggregates formed by a volitional human agent, it is not infrequent for *kupa* 'heap' to classify portions of substances whose shape is a result of the forces of nature or merely constitutes a by-effect of activities intended to achieve goals other than arranging stuff into units. What differentiates between *sterta* 'pile' and *stos* 'stack', however, is that constructional solidity appears a more salient feature of the latter item, hence its capability of applying to vertical collections of entities marked by an orderly internal structure.

1. Introduction[1]

Discussing the semantics of the Polish arrangement classifiers (henceforth also ACLs) *kupa* 'heap', *sterta* 'pile', and *stos* 'stack', all of which encode vertical arrangements

[1] The abbreviations employed in this paper are as follows: ACL – arrangement classifier; AUG – augmentative form; DIM – diminutive form; GEN – genitive case; N1 – the first nominal element in a binominal syntagm; N2 – the second nominal element in a binominal syntagm.

of objects or substances and hence prototypically combine with concrete inanimate nominals, Bednarek (1994: 116–117) tentatively proposes that the reference of a noun delimited by *kupa* 'heap' is conceptualized as a portion of undifferentiated mass, whereas both *sterta* 'pile' and *stos* 'stack' are more conceptually specific in that they typically classify sets composed of perceptually discrete entities. However, no quantitative data are offered to substantiate this hypothesis.

In the light of the above observations, the present paper sets out to confront the aforementioned intuition-based claim with empirical, corpus-derived evidence. Since more often than not, the countability status of a nominal predicate reflects the cognitively salient distinction between bounded entities and non-bounded substances, the conceptual specificity of a particular ACL will be operationalized as the frequency of its co-occurrence with (i) count and (ii) non-count N2-collocates, the assumption being that the more numerous the collocates belonging to the latter category, the less conceptually specific the ACL. Accordingly, it is expected that *kupa* 'heap' should exhibit a markedly higher percentage of uses involving mass N2-collocates than is the case with *sterta* 'pile' and *stos* 'stack'.

The structure of the paper is as follows. Section 2 offers basic information pertaining to classifiers and their role in the individuation of nominals. Section 3 outlines the research hypothesis, and section 4 describes the applied methodology. The results of an analysis of naturally-occurring Polish data are presented in section 5 and further discussed in section 6. Finally, section 7 summarizes the main conclusions arrived at in the empirical investigation as well as suggests prospects for future research on the topic.

2. Classifiers, individuation, and (un)countability

Classifiers, also known as measure nouns and partitives, will be defined here as items which serve to individuate the reference of the accompanying nominals, i.e. to impose "a quantitative limit on the extension of the predicates they apply to" (Willim 2006: 45),[2] by means of either unit-excerption, part/portion-excerption or unit-creation. In the former cases, classifiers co-occur with mass nouns and thus make it possible to achieve countability thereof (cf. Grochowski 1992: 70), as in (1a), whereas in the latter, they combine with plural nouns and as such further unitize discrete entities, as in (1b):

[2] Even though such individuating elements can be found cross-linguistically, not all languages are traditionally regarded as classifier languages (cf., among others, Allan 1977; Dik 1989; Bednarek 1994). The primary distinction between classifier languages and non-classifier ones is that in the former, nouns are, as a rule, obligatorily accompanied by classifiers in quantifying constructions, which is not the case in the latter. As for classifier languages, Dik (1989: 123) therefore posits the existence of ensemble nouns, i.e. "nominal predicates which can be used to refer to ensembles without forcing the question whether these are sets or masses". Non-classifier languages, on the other hand, are known to make a syntactic distinction between count and non-count, i.e. mass, predicates. Hence, it is primarily nouns representative of the latter category that occur with various classificatory items.

(1) a. *trzy* ***kupy*** *błota* 'three heaps of mud' (**trzy błota* 'three muds'), *pięć* ***kilogramów*** *piasku* 'five kilograms of sand' (**pięć piasków* 'five sands'), *siedem* ***wiader*** *żwiru* 'seven buckets of gravel' (**siedem żwirów* 'seven gravels')

b. *trzy* ***grupy*** *ludzi* 'three groups of people', *kolejny* ***zestaw*** *narzędzi* 'another set of tools', *kilka* ***kolekcji*** *obrazów* 'a few collections of pictures'

As can be seen, a classifier is typically placed between a quantifier and the quantified NP (cf. Topolińska 1984: 369), so that the classifier phrase is embedded within a larger quantifier construction. This applies specifically to mass nouns (cf. (1a)), which are not susceptible to enumeration without having been individuated by a classificatory item (cf. Lyons 1977: 462).[3]

Another important characteristic of classifiers is that for the most part, they indicate "some salient perceived or imputed characteristic of the entity to which an associated noun refers" (Allan 1977: 285). In other words, such items exhibit specific selectional requirements in that they individuate the concomitant nominals in terms of certain cognitively prominent features of their referents, such as consistency, shape, or configuration (cf., among others, Lee 1987; Craig 1992; Croft 1994; Aikhenvald 2006). As for Polish, Bednarek (1994) identifies over a hundred classificatory elements which display sensitivity towards the semantic properties of the classified N2s (cf. also Nowosad-Bakalarczyk 2013: 235–236). By way of illustration, the classifier *kropla* 'drop' is only compatible with nouns denoting liquids (cf. (2)), while *kostka* 'cube' co-occurs with N2s referring to substances of rather solid consistency (cf. (3)):

(2) a. *kropla wody* 'a drop of water'

b. **kropla cukru* 'a drop of sugar'

(3) a. *kostka cukru* 'a cube of sugar'

b. **kostka wody* 'a cube of water'

At the same time, both classifiers reveal additional information as to the shape of the portion of the relevant mass: while the former implies a roundish shape, the latter points to the referent of the classifier construction being saliently three-dimensional. As already mentioned, the items *kupa* 'heap', *sterta* 'pile', and *stos* 'stack', in turn, convey the arrangement of what the associated nominal predicate stands for, and may combine with both count and non-count (solid) concrete nouns whose referents are capable of forming a pile-like structure.

Building on Allan's (1977) comprehensive typological study of classificatory morphemes, Lehrer (1986: 111) proposes a semantic taxonomy of classifiers encompassing

[3] Mass nouns can nevertheless be reclassified so as to denote portions or kinds of the pertinent substance. As Lehrer (1986: 110) points out, non-count predicates can be "freely used as count nouns", their reference being "determined by the pragmatics of the situation". For instance, the phrase *three coffees* may stand for either three portions or kinds of coffee. It should be emphasized, however, that there exist certain intra- as well as cross-linguistic differences with respect to the acceptability of such instances of conversion (cf. Willim 2006: 59).

seven categories of such items. Although originally formulated in relation to English, the pertinent classification can easily be extended to Polish, as evidenced by the examples provided below:

Classifier type	Example
(i) unit counters	*kilka* ***sztuk*** *bydła* 'a few head of cattle'
(ii) fractional classifiers	*dwie* ***trzecie*** *ciasta* 'two-thirds of the cake'
(iii) number set classifiers	*pięć* ***tuzinów*** *kwiatów* 'five dozen flowers'
(iv) collective classifiers	*kolejny* ***zbiór*** *danych* 'another set of data'
(v) varietal classifiers	*różne* ***rodzaje*** *muzyki* 'various kinds of music'
(vi) measure classifiers	*trzy* **litry** *wina* 'three litres of wine'
(vii) arrangement classifiers	*wiele* ***stert*** *książek* 'many piles of books'

Table 1. Categories of classifiers

As stated before, of special interest here are elements belonging to the last category, i.e. ACLs, which themselves bear semantic resemblance to collective classifiers. Analyzing the two types of items, Xiao and McEnery (2010: 57) observe that, as far as count N2-collocates are concerned, both "provide a collective reference", yet only the former "focus on the constellation aspect", i.e. carry information as to how the classified entities are grouped together, and can therefore be looked at as "largely motivated by the cognitive basis of shape" (Xiao, McEnery 2010: 66). Similarly, Jackendoff (2012: 1141) remarks that of the classificatory nouns which he calls aggregates, some possess an inherent shape, e.g. *heap*, *stack*, and *row*, as against *swarm* or *flock*, the latter lacking this quality.

3. Research hypothesis

As noted by, among others, Xiao and McEnery (2006: 108), collocations, i.e. instances of regular co-occurrence of (types of) words, frequently bear witness to slight divergences in the meanings of lexical elements commonly thought of as denotationally identical or at least highly similar, hence the term *near-synonymy*, applicable to closely semantically related items which nonetheless differ in their combinatorial preferences. A case in point may be the three ACLs *kupa* 'heap', *sterta* 'pile', and *stos* 'stack': in the Universal Dictionary of Polish, for instance, *sterta* 'pile' and *stos* 'stack' are listed as synonyms of *kupa* 'heap' (cf. Dubisz 2003a: 565), and the former items are likewise present in each other's definitions (cf. Dubisz 2003b: 533 and 544, respectively).

In view of Bednarek's (1994: 116–117) tentative claim, according to which *kupa* 'heap' refers to portions of undifferentiated mass, while both *sterta* 'pile' and *stos*

'stack' stand for sets composed of perceptually discrete objects, and given the overwhelming tendency for count and non-count concrete nouns to denote, respectively, inherently bounded entities and amorphous stuff (cf., among others, Bogusławski 1973; Allan 1980; Drożdż 2017), it is expected that the former ACL will combine with mass nominals to a substantially higher degree than is the case with *sterta* 'pile' and *stos* 'stack', the latter items being instead hypothesized to collocate mostly with count N2s whose referents retain their conceptual granularity. In other words, despite the spatial boundedness and vertical arrangement of the N2-referent(s) of each of the scrutinized ACLs, what *kupa* 'heap' applies to is assumed to constitute an aggregate of objects or fragments of a substance of such a small size that they are not conceived of as separate entities, but are rather perceived as forming an internally homogeneous whole, whereas both *sterta* 'pile' and *stos* 'stack' classify collections of fairly sizeable and thus readily distinguishable units.

As suggested above, differences in the countability status of nouns are not, however, invariably reflective of the salient discrepancy between individuated objects and non-bounded substances (Willim 2006: 25), as evidenced by the occasional availability of two distinct nominal forms, one being syntactically count, and the other non-count, referring to the same piece of reality (cf., e.g. *liście* 'leaves' and *listowie* 'foliage'). Moreover, some syntactically mass predicates, such as English *jewelry* and its Polish counterpart *biżuteria* 'jewelry', are in fact referentially countable, i.e. rather than standing for amorphous substance, they apply to spatially delimited objects which can easily be counted. Conversely, when pluralized, count nouns such as *śmieć* 'a piece of rubbish' typically refer to what can be described as masses consisting of perceptually indiscriminable entities.[4] Apart from grouping the nominals accompanying the analyzed ACLs into count and non-count, it will therefore be necessary to closely examine the referential status of the detected collocates of a given ACL.

4. Methodology

The first stage of the investigation involved the extraction of random samples of occurrences of *kupa* 'heap', *sterta* 'pile', and *stos* 'stack' in the binominal N1 N2.GEN-construction, where N1 stands for a particular ACL, and N2 refers to the nominal being assessed in terms of quantity, from the National Corpus of Polish (henceforth NKJP) by means of the Poliqarp search engine.[5] The aim was to collect a set of 100 attestations of each of the above-listed items, i.e. 300 uses in total, in which it individuates the reference of a concrete inanimate N2-collocate. This entailed

[4] Even more challenging to traditional treatments of (un)countability is the individuation of abstract predicates. For an overview of the pertinent problems, see Grimm (2014).

[5] The following commands were used: [base="kupa"] [pos=subst & cas=gen], [base="sterta"] [pos=subst & cas=gen], and [base="stos"] [pos=subst & cas=gen]. If the N2-slot was occupied by two or more nominal elements, it was the countability status of the first, i.e. leftmost, one that was taken into consideration in the data classification procedure.

the removal of all irrelevant instances, chiefly purely quantificational uses (cf., e.g. *kupa czasu* 'a lot of time') as well as figurative extensions (cf., e.g. *kupa złomu* 'lit.: a heap of scrap iron' employed metaphorically in relation to an old, malfunctioning car). The obtained attestations were then classified into uses involving (i) count and (ii) non-count N2-collocates.[6] The final step consisted in formulating semantic generalizations over the types of nouns typically collocating with the ACLs under scrutiny.

5. Results

Presented below are the results of the quantitative analysis of the collocational patterns of the Polish ACLs *kupa* 'heap', *sterta* 'pile', and *stos* 'stack':

ACL	Count N2-collocates (%)	Non-count N2-collocates (%)
Kupa 'heap'	25 (25%)	75 (75%)
Sterta 'pile'	77 (77%)	23 (23%)
Stos 'stack'	82 (82%)	18 (18%)

Table 2. Percentages of count and non-count N2-collocates

In consonance with the advanced hypothesis, there is a significant disparity between *kupa* 'heap' on the one hand and *sterta* 'pile' and *stos* 'stack' on the other in that the former typically collocates with mass nouns, whereas the latter display a comparably conspicuous preference for count N2-collocates: χ^2 (2, N = 300) = 84.02, $p < .001$. In what follows, additional qualitative comments are offered as to the collocability of the scrutinized items. Each subsection begins with a table containing a list of the most frequent count and non-count collocates of a particular ACL.

5.1 *Kupa* 'heap'

As can be easily noted, a vast majority of the non-count N2-collocates of *kupa* 'heap', such as *gruz* 'rubble' and *gruzy* 'debris', *gnój* 'dung', *śnieg* 'snow', *chrust* 'brushwood', *muł* 'river mud', and *nawóz* 'muck' stand for amorphous substances rather than referentially countable objects. Also the count nouns co-occurring with the ACL in question generally apply to entities which are hardly distinguishable from one another due to their relatively small size, visual similarity or even identity, and typically numerous co-appearance, e.g. *śmieci* 'pieces of rubbish' and *odpadki* 'wastes', *kamienie* 'stones', as well as *liście* 'leaves'. What also points to a low degree of

[6] Pluralia tantum nouns standing for collections of entities, such as *akta* 'files; dossier' and *pierze* 'feathers', which have intrinsically cumulative reference, were subsumed under the non-count category.

Count N2-collocates (No. of attestations)	Non-count N2-collocates (No. of attestations)
1. *Śmieci* 'pieces of rubbish' (6)	1. *Gruz* 'rubble' (19)
2. *Kamienie* 'stones' (4)	2. *Gnój* 'muck' (9)
3. *Liście* 'leaves' (2)	3. *Gruzy* 'debris' (7)
4. *Odpadki* 'wastes' (2)	4. *Śnieg* 'snow' (4)
5. *Papiery* 'papers; documents' (2)	5. *Chrust* 'brushwood' (3)
6. *Szmaty* 'cloths' (2)	6. *Muł* 'river mud' (3)
7. *Banknoty* 'banknotes' (1)	7. *Nawóz* 'manure' (3)
8. *Cegły* 'bricks' (1)	8. *Piasek* 'sand' (3)
9. *Kartofle* 'potatoes' (1)	9. *Błoto* 'mud' (2)
10. *Kwity* 'receipts' (1)	10. *Obornik* 'manure' (2)
11. *Listy* 'letters' (1)	11. *Piach* 'sand.AUG' (2)
12. *Wióry* 'shavings' (1)	12. *Popiół* 'ash' (2)
13. *Węgle* 'embers' (1)	13. *Węgiel* 'coal' (2)
	14. *Żelastwo* 'iron stuff' (2)
	15. *Żwir* 'gravel' (2)
	16. *Drzewo* 'wood' (1)
	17. *Jedwab* 'silk' (1)
	18. *Kompost* 'manure' (1)
	19. *Pierze* 'feathers' (1)
	20. *Posłanie* 'bedding' (1)
	21. *Ruiny* 'ruins' (1)
	22. *Ziemia* 'earth' (1)
	23. *Złom* 'scrap iron' (1)
	24. *Złoto* 'gold' (1)
	25. *Żarcie* 'food' (1)

Table 3. The most frequent N2-collocates of the ACL *kupa* 'heap'

conceptual granularity of the extensions of these nouns is the fact that they can be replaced by mass counterparts, i.e. it is not infrequent for *śmieć* 'a piece of rubbish' and *kamień* 'stone' to generically refer to (pieces of) rubbish and stone(s), respectively, with the already mentioned item *listowie* 'foliage' being the non-count variant encoding the same fragment of reality as the plural form of the count noun *liść* 'leaf'.[7]

5.2 *Sterta* 'pile'

The above table demonstrates that he ACL *sterta* 'pile' exhibits a marked predilection for count collocates referring to objects made of paper. Among the pertinent examples are the N2s *gazety* 'newspapers', *dokumenty* 'documents', *książki* 'books', *czasopisma* 'magazines', *papiery* 'papers; documents', *pudełka* 'boxes', and *wydruki*

[7] Note also that in English, what the plural forms of the three Polish nouns under discussion refer to may likewise be expressed by mass predicates, i.e. *rubbish*/*trash*, *stone*, and *foliage*, respectively.

Count N2-collocates (No. of attestations)	Non-count N2-collocates (No. of attestations)
1. *Śmieci* 'pieces of rubbish' (8)	1. *Gruz* 'rubble' (4)
2. *Gazety* 'newspapers' (6)	2. *Gruzy* 'debris' (3)
3. *Dokumenty* 'documents' (5)	3. *Papier* 'paper' (3)
4. *Książki* 'books' (5)	4. *Żelastwo* 'iron stuff' (2)
5. *Cegły* 'bricks' (3)	5. *Akta* 'files; dossier' (1)
6. *Worki* 'bags' (3)	6. *Bagaż* 'luggage' (1)
7. *Czasopisma* 'magazines; periodicals' (2)	7. *Bielizna* 'underwear' (1)
8. *Gałęzie* 'branches' (2)	8. *Cegła* 'brick' (1)
9. *Kamienie* 'stones' (2)	9. *Drewno* 'wood' (1)
10. *Odpadki* 'wastes' (2)	10. *Obornik* 'manure' (1)
11. *Papiery* 'papers; documents' (2)	11. *Piach* 'sand.'AUG (1)
12. *Płyty* 'discs' (2)	12. *Popiół* 'ash' (1)
13. *Pudełka* 'boxes' (2)	13. *Siano* 'hay' (1)
14. *Ubrania* 'clothes' (2)	14. *Ziemia* 'earth' (1)
15. *Wydruki* 'printouts' (2)	15. *Śnieg* 'snow' (1)
16. *Buty* 'shoes' (1)	
17. *Deski* 'planks' (1)	
18. *Głazy* 'boulders' (1)	
19. *Kartony* 'cartons' (1)	
20. *Kasety* 'cassettes' (1)	
21. *Kości* 'bones' (1)	
22. *Księgi* 'books.AUG' (1)	
23. *Kwiaty* 'flowers' (1)	
24. *Materace* 'mattresses' (1)	
25. *Niedopałki* 'butts' (1)	
26. *Papierzyska* 'papers.AUG; documents' (1)	
27. *Plastry mięsa* 'slices of meat' (1)	
28. *Plecaki* 'backpacks' (1)	
29. *Płótna* 'canvas paintings' (1)	
30. *Podróbki* 'counterfeits' (1)	
31. *Prace* 'works' (1)	
32. *Rachunki* 'bills' (1)	
33. *Reklamy* 'advertisements' (1)	
34. *Raptularzyki* 'diaries.DIM' (1)	
35. *Roczniki* 'towels' (1)	
36. *Rupiecie* '(pieces of) junk' (1)	
37. *Szmaty* 'cloths' (1)	
38. *Talerze* 'plates' (1)	
39. *Tobołki* 'bundles' (1)	
40. *Towary* 'commodities' (1)	
41. *Ulotki* 'leaflets' (1)	
42. *Walizki* 'suitcases' (1)	
43. *Wory* 'bags.AUG' (1)	
44. *Ziemniaki* 'potatoes' (1)	

Table 4. The most frequent N2-collocates of the ACL *sterta* 'pile'

'printouts'. Yet, its most frequent N2-collocate in the analyzed data, namely *śmieci* 'pieces of rubbish', on a par with *kamienie* 'stones' and *odpadki* 'wastes', suggests that what forms an aggregate described as *sterta* 'pile' need in fact not be conceived of as a collection of readily differentiable entities. As for its non-count collocates, *sterta* 'pile' is again similar to *kupa* 'heap' in that both combine with the N2s *gruz* 'rubble' and *gruzy* 'debris', although obviously with radically distinct frequencies (cf. 26 occurrences of *kupa gruzu/gruzów* 'a heap of rubble/debris' vs. a mere 7 attestations of *sterta gruzu/gruzów* 'a pile of rubble/debris'). Curiously, aside from *gruz* 'rubble' and *gruzy* 'debris', the N2-collocates *obornik* 'manure', *piach* 'sand.AUG', *popiół* 'ash', *siano* 'hay', *ziemia* 'earth', and *śnieg* 'snow' likewise apply to amorphous stuff rather than sets of discrete objects.

5.3 *Stos* 'stack'

Count N2-collocates (No. of attestations)	**Non-count N2-collocates (No. of attestations)**
1. *Papiery* 'papers; documents' (6)	1. *Drewno* 'wood' (4)
2. *Książki* 'books' (6)	2. *Chrust* 'brushwood' (2)
3. *Śmieci* 'pieces of rubbish' (5)	3. *Gruzy* 'debris' (1)
4. *Gazety* 'newspapers' (4)	4. *Korespondencja* 'correspondence' (1)
5. *Kamienie* 'stones' (3)	5. *Obuwie* 'footwear' (1)
6. *Opony* 'tyres' (3)	6. *Opał* 'firewood' (1)
7. *Dokumenty* 'documents' (2)	7. *Pieniądze* 'money' (1)
8. *Krzesła* 'chairs' (2)	8. *Porcelana* 'china' (1)
9. *Materiały* 'materials' (2)	9. *Prasa* 'press' (1)
10. *Pisma* 'letters; documents' (2)	10. *Słonina* 'bacon' (1)
11. *Rupiecie* 'pieces of junk' (2)	11. *Złom* 'scrap iron' (1)
12. *Talerze* 'plates' (2)	12. *Śnieg* 'snow' (1)
13. *Bagaże* 'pieces of luggage' (1)	13. *Żelastwo* 'iron stuff' (1)
14. *Banknoty* 'banknotes' (1)	14. *Żywność* 'food' (1)
15. *Blaszanki* 'tins' (1)	
16. *Czasopisma* 'magazines; periodicals' (1)	
17. *Druczki* 'printouts.DIM' (1)	
18. *Gałęzie* 'branches' (1)	
19. *Igły* 'needles' (1)	
20. *Jabłka* 'apples' (1)	
21. *Karty* 'cards' (1)	
22. *Kartki* 'sheets of paper' (1)	
23. *Kartony* 'cartons' (1)	
24. *Kije* 'sticks' (1)	
25. *Kostki trotylu* 'cubes of trotyl' (1)	
26. *Kości* 'bones' (1)	
27. *Krawaty* 'ties' (1)	
28. *Kwiaty* 'flowers' (1)	
29. *Materace* 'mattresses' (1)	

Count N2-collocates (No. of attestations)	Non-count N2-collocates (No. of attestations)
30. *Niedopałki* 'butts' (1)	
31. *Notatki* 'notes' (1)	
32. *Notatniki* 'notebooks' (1)	
33. *Odpady* 'wastes' (1)	
34. *Palety* 'pallets' (1)	
35. *Papierki* 'papers.DIM; documents' (1)	
36. *Papierzyska* 'papers.AUG; documents' (1)	
37. *Płyty* CD 'CDs' (1)	
38. *Podręczniki* 'handbooks' (1)	
39. *Poduszki* 'pillows' (1)	
40. *Projekty* 'projects' (1)	
41. *Prospekty* 'prospects' (1)	
42. *Prześcieradła* 'sheets' (1)	
43. *Publikacje* 'publications' (1)	
44. *Rekwizyty* 'props' (1)	
45. *Rysunki* 'drawings' (1)	
46. *Rzeczy* 'objects' (1)	
47. *Szpargały* 'pieces of junk' (1)	
48. *Teczki* 'folders' (1)	
49. *Tortille* 'tortillas' (1)	
50. *Towary* 'commodities' (1)	
51. *Ubrania* 'garments' (1)	
52. *Wieńce* 'wreaths' (1)	
53. *Wycinki* 'excerpts' (1)	
54. *Zapiski* 'notes' (1)	
55. *Zeszyty* 'notebooks' (1)	

Table 5. The most frequent N2-collocates of the ACL *stos* 'stack'

Analogously to *sterta* 'pile', *stos* 'stack' reveals a preference for paper-denoting count N2s. Importantly, however, the latter ACL more frequently combines with nouns referring to fairly large-size, clearly discriminable entities, such as *opony* 'tyres', *krzesła* 'chairs', and *talerze* 'plates', which, apart from a slightly higher proportion of its uses involving count N2-collocates, indicates that what falls under the denotation of *stos* 'stack' is typically a collection of conceptually distinct objects which may be placed one on top of another. Likewise, most of the discussed item's non-count collocates actually stand for referentially countable entities: *opał* 'firewood', for instance, applies to perceptually delimited pieces of wood rather than, e.g. hardly differentiable wood-shavings, *korespondencja* 'correspondence' denotes a set of letters, *obuwie* 'footwear' and *prasa* 'press' collectively refer to shoes and newspapers, respectively, and *porcelana* 'china' applies to distinct artefacts made of the relevant substance, such as cups or pots. Admittedly, it is not impossible for the ACL under analysis to also combine with predicates encoding more or less undifferentiated masses, such as the count noun *śmieci* 'pieces of rubbish' and the

non-count noun *chrust* 'brushwood', yet the number of such uses in the scrutinized data is negligible (cf. Table 5), especially when compared with that of *kupa* 'heap' and, to a lesser degree, *sterta* 'pile'.

6. Discussion

In accordance with the research assumption specified in section 3, the ACL *kupa* 'heap' differs from *sterta* 'pile' and *stos* 'stack' in that it displays a pronounced predilection for non-count N2-collocates, whereas the latter ACLs predominantly combine with count nouns. On closer inspection, however, it turns out that contra Bednarek (1994), who equals *sterta* 'pile' and *stos* 'stack' with respect to their semantics, the former item is similar to *kupa* 'heap' in that it may also be employed in relation to collections composed of hardly distinguishable entities or even portions of amorphous stuff, even though such uses are considerably more frequently attested for the latter ACL. *Stos* 'stack', in turn, can be claimed to exhibit the highest level of perceptual granularity of the three ACLs under scrutiny, since it combines with N2-collocates standing for objects of a fairly large size, which are clearly perceptually delimited, to a greater extent than is the case with *sterta* 'pile' and, especially, *kupa* 'heap'.

Still, what links the ACLs *sterta* 'pile' and *stos* 'stack' is the fact that both typically refer to aggregates formed by a volitional human agent, whereas stuff individuated by means of *kupa* 'heap' need not be intentionally arranged into such a structure; rather, its shape can be an outcome of the forces of nature (cf. (4–5)), or may merely constitute a by-effect of volitional activities intended to achieve goals other than the formation of heaps of the associated substances, e.g. the event of destroying a building (cf. (6)), as a result of which mounds of rubble inescapably come into existence:

(4) Z gęstwiny zarośli, spod wykrotu czy **kupy chrustu** w wilgotniejszych typach lasów, usłyszymy często gwałtowne terkotanie. Wydaje je strzyżyk – krótkoskrzydłe, brązowe maleństwo z ciemniejszymi prążkami rozsianymi po całym ciele. (NKJP)

'In more humid forests, one is likely to hear an intense rattle coming from the thicket, from under a fallen tree, or from a heap of brushwood. The sound is made by a wren – a little short-tailed brown creature with dark stripes all over its body.'

(5) Zarośnięte pnączami mury gdzieniegdzie były jeszcze całkiem wysokie, lecz nieubłaganie kruszyły się od góry. Większość zawaliła się jednak zupełnie, tworząc **kupy gruzu**, zagarnięte niemal całkowicie przez roślinność. (NKJP)

'The creeper-covered walls were still fairly high here and there, yet kept crumbling from the top. Most of them had nonetheless totally collapsed, creating heaps of rubble, completely seized by vegetation.'

(6) Wczoraj o 2.10 w nocy kilka natowskich rakiet zamieniło siedzibę państwowej telewizji jugosłowiańskiej w **kupę gruzu**. (NKJP)

'Yesterday at 2.10 AM a few NATO missiles turned the building of Yugoslavian state television into a heap of rubble.'

The above-mentioned facts are likewise connected with the general tendency for *sterta* 'pile' and *stos* 'stack' to classify aggregates of artefacts (cf. Table 4 and Table 5, respectively). *Kupa* 'heap', on the other hand, is more likely to be employed with reference to portions of natural substances (cf. Table 3).

Another factor differentiating between the discussed ACLs is the varying extent to which each of them profiles the feature of layered verticality, as corroborated by the disparity in the frequency of their co-occurrence with the verb *piętrzyć się* 'to mount; to rise', clearly invoking a vertical dimension: in the investigated data, only one instance of this kind has been detected for *kupa* 'heap', three such uses have been identified for *sterta* 'pile', and as many as nine for *stos* 'stack', of which a few examples are cited below:

(7) Eleonora wskazała głową na **stos materiałów** *piętrzący się* na stoliku. (NKJP)
'Eleonora nodded towards the stack of materials mounting on the table.'

(8) W obu pokoikach *piętrzyły się* **stosy rekwizytów** i trudno się było poruszać. (NKJP)
'Mounting in both of the little rooms were stacks of props, and it was difficult to move around.'

(9) Wielkie, opasłe tomy, **stosy teczek**, dokumentów i papierów *piętrzą się* koło biurek. (NKJP)
'Huge, bulky tomes, and stacks of folders, documents, and papers are mounting next to the desks.'

That the semantics of *stos* 'stack' incorporates a saliently layered vertical structure further translates into its capability of standing for orderly, geometrically neat arrangements of objects. Both *kupa* 'heap' and *sterta* 'pile', by contrast, tend to classify aggregates marked by a more chaotic internal organization.

7. Conclusion

The results of the empirical examination pursued here corroborate Bednarek's (1994: 116–117) intuitive claim pertaining to the semantics of the Polish ACLs *kupa* 'heap', *sterta* 'pile', and *stos* 'stack' only in part. In accordance with the assumed hypothesis, it has been demonstrated that as against *kupa* 'heap', which is usually accompanied by stuff-denoting non-count nominals, both *sterta* 'pile' and *stos* 'stack' combine mainly with count predicates.

Nevertheless, there exist certain notable discrepancies between the latter ACLs, a finding which is unexpected under Bednarek's (1994) account. Firstly, *sterta* 'pile' is more likely to co-occur with N2-collocates denoting amorphous substances, whereas *stos* 'stack' exhibits a more pronounced tendency to classify sets comprising clearly spatially delimited objects of a fairly large size, with a considerable proportion of its syntactically mass N2-collocates actually referring to countable entities. Secondly, constructional solidity seems to be a more conceptually salient feature of *stos* 'stack',

which implies the ACL's ability to stand for collections characterized by an orderly geometric arrangement. On the other hand, *sterta* 'pile' and, in particular, *kupa* 'heap' typically refer to internally more chaotic arrangements. Still, in addition to their predilection for count N2-collocates, *sterta* 'pile' and *stos* 'stack' are similar in that they usually apply to sets of artefacts arranged by a volitional human agent, while *kupa* 'heap' frequently classifies portions of natural substances whose shape is a result of the forces of nature or constitutes a by-effect of activities intended to achieve goals other than arranging stuff into units.

In order to shed further light on the semantics of the three classifiers, the corpus-based analysis can be complemented with an experimental investigation whereby native speakers of Polish will first be confronted with a number of photographs presenting vertical configurations of a specific kind of stuff which slightly differ with respect to their exact dimensions (e.g. a pyramid-like vs. a tower-like structure), and then asked to decide whether the arrangement shown in a particular picture should be best labelled as *kupa* 'heap', *sterta* 'pile', or *stos* 'stack'. Considering that the classificatory nouns at issue share a few of their most frequent collocates (cf. Tables 3–5), such an additional study will enable a more precise ascertainment of how the conceptualization of one and the same N2-item changes when it combines with distinct near-synonymous ACLs.

References

Aikhenvald A.Y. 2006. Classifiers and noun classes: Semantics. – Brown K. (ed.). *Encyclopedia of language and linguistics*. Oxford: 463–471.

Allan K. 1977. Classifiers. – *Language* 53.2: 285–311.

Allan K. 1980. Nouns and countability. – *Language* 56.3: 541–567.

Bednarek A. 1994. *Leksykalne wykładniki parametryzacji świata: Studium semantyczne*. Toruń.

Bogusławski A. 1973. Nazwy pospolite przedmiotów konkretnych i niektóre właściwości ich form liczbowych i połączeń z liczebnikami w języku polskim. – Topolińska Z., Grochowski M. (eds.). *Liczba, ilość, miara*. Wrocław: 7–35.

Craig C.G. 1992. Classifiers in a functional perspective. – Harder P., Fortescue M., Kristoffersen L. (eds.). *Papers from the Functional Grammar Conference*. Amsterdam: 277–301.

Croft W. 1994. Semantic universals in classifier systems. – *Word* 45: 145–171.

Dik S.C. 1989. *The theory of functional grammar. Part 1: The structure of the clause*. Dordrecht.

Drożdż G. 2017. *The puzzle of (un)countability in English: A study in cognitive grammar*. Katowice.

Dubisz S. (ed.). 2003a. *Uniwersalny słownik języka polskiego: Tom 2: H–N*. Warszawa.

Dubisz S. (ed.). 2003b. *Uniwersalny słownik języka polskiego: Tom 4: R–V*. Warszawa.

Grimm S. 2014. Individuating the abstract. – Etxeberria U., Falaus A., Irurtzun A., Leferman B. (eds.). *Proceedings of Sinn und Bedeutung 18*. Vitoria: 182–200.

Grochowski M. 1992. O metodzie wyjaśniania struktury semantycznej nazw substancji naturalnych. – Siatkowski J., Galster B. (eds.). *Z polskich studiów slawistycznych: Seria VIII: Językoznawstwo*. Warszawa: 69–74.

Jackendoff R. 2012. Language as a source of evidence for theories of spatial representation. – *Perception* 41: 1128–1152.

Lee M. 1987. The cognitive basis of classifier systems. – Aske J., Beery N., Filip H., Michaelis L. (eds.). *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: 395–407.

Lehrer A. 1986. English classifier constructions. – *Lingua* 68: 109–148.

Lyons J. 1977. *Semantics*. Cambridge.

NKJP = Bańko M., Górski R.L., Lewandowska-Tomaszczyk B., Łaziński M., Pęzik P., Przepiórkowski A. 2008–2012. *Narodowy Korpus Języka Polskiego*. [http://nkjp.pl/poliqarp/; 14.05.2019].

Nowosad-Bakalarczyk M. 2013. Linguistic categories in onomasiological perspective. The category of quantity in contemporary Polish. – Głaz A., Danaher D.S., Łozowski P. (eds.). *The linguistic worldview: Ethnolinguistics, cognition, and culture*. London: 227–244.

Topolińska Z. 1984. Składnia grupy imiennej. – Topolińska Z. (ed.). *Gramatyka współczesnego języka polskiego: Składnia*. Warszawa: 301–386.

Willim E. 2006. *Event, individuation and countability: A study with special reference to English and Polish*. Kraków.

Xiao R., McEnery T. 2006. Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. – *Applied Linguistics* 27.1: 103–129.

Xiao R., McEnery T. 2010. *Corpus-based contrastive studies of English and Chinese*. New York.