

Three distributive elements *po* in Polish (without missing lexical generalisations)

Adam Przepiórkowski

The aim of this paper is to discuss an ugly detail of Polish grammar, whose unseemliness has not been fully appreciated in Polish linguistics so far, and to sketch an HPSG analysis of this quirk.

1 Introduction: two functional elements *po*

The distributive *po*, exemplified below, has been known to be troublesome at least since Łojasiewicz 1979; it is traditionally analysed as a preposition, apparently combining either with a locative NP (noun phrase, as in (1)) or with an accusative NumP (numeral phrase, as in (2)).

- (1) Dałem im po jabłku. (2) Dałem im po dwa jabłka.
gave-I them.DAT *po* apple.LOC gave-I them.DAT *po* two.ACC apples.ACC
'I gave them an apple each.' 'I gave them two apples each.'

Przepiórkowski 2010 argues that there are in fact *two* different functional lexemes *po* in Polish: a preposition, exemplified by (1), and an adnumeral operator (in the sense of the classification of functional parts of speech in Polish by Grochowski 1997), exemplified by (2). The main difference between them is that the distributive preposition *po* assigns the locative case, which is a strictly prepositional case in Polish, and the adnumeral operator *po* is transparent to any case assignment mechanisms, i.e., the accusative case in (2) is assigned to the direct object of *dałem* 'gave'.

The main reason for distinguishing such an adnumeral operator *po* is that it may occur with numeral phrases in cases other than the previously noticed accusative, as shown by the marginal but attested dative example (3) below:

- (3) Broń została przekazana po dwóm osobom z każdego ugrupowania.
weapon AUX transferred.PASS *po* two.DAT person.DAT.PL from each group
'The weapon was handed in to two people from each group.'

More robustly, adnumeral *po*-phrases¹ may also occur in the subject position, as already noted by Łojasiewicz 1979, where they agree with the verb, which shows that they bear the nominative case (in Polish, as in other European languages, the finite verb only agrees with nominative subjects):

- (4) W pokojach będą po dwa fotele.
in rooms be.FUT.PL *po* two.NOM/ACC armchair.NOM/ACC.PL
'There will be two armchairs in each room.'
- (5) Na ławkach leżały po trzy arkusze papieru.
on benches lie.PAST.PL *po* three.NOM/ACC sheet.NOM/ACC.PL paper.GEN.SG
'There lay three sheets of paper on each bench.'

The relevant numeral phrases are marked as *NOM/ACC* to indicate that they are morphosyntactically syncretic and could in principle be analysed as accusative, although this would clash with the fact that they agree with the finite verb. Moreover, attested examples can be readily found which involve unambiguously nominative (non-syncretic) *po*-phrases in the subject position, e.g.:

- (6) Do Senatu wybierani są po dwaj senatorzy z każdego stanu.
to senate selected be.PRES.PL *po* two.NOM senators.NOM.PL from each state
'Two senators from each state are selected for the Senate.'

In summary, while *po* in (1) is clearly a preposition assigning the locative case, the synonymous and at the same time homonymous *po* in (2)–(6) does not assign the locative, so it must be analysed as another functional lexeme, e.g., as an adnumeral operator transparent to case assignment.

¹We use the term *po-phrases* pretheoretically, without any prejudication as to their internal syntax.

2 More data: three functional elements *po*

Unfortunately, facts are even more ugly than the previous section would indicate. Consider the following attested examples (from the National Corpus of Polish, <http://nkjp.pl/>).

- (7) ...w każdym meczu nie wykorzystuje *po* kilka dobrych okazji...
in each match not take advantage *po* a few.ACC good opportunities
'... (he) doesn't take advantage of a few good opportunities in each match...'
- (8) Zaobserwowano również przypadki rozdawania *po* dwa egzemplarze kart...
observed.IMPS also cases distributing *po* two.ACC copies.ACC cards.GEN
'Also cases, where two cards were handed, were observed...'

In both sentences *po*-phrases occupy genitive positions. The following examples, constructed – by removing the element *po* – as minimally different from (7)–(8), show that these positions, while corresponding to the accusative position of a non-negated non-gerundial verb, are strictly genitive when the verb is negated or nominalised:

- (9) ...w każdym meczu nie wykorzystuje kilku/*kilka dobrych okazji...
in each match not take advantage a few.GEN/*ACC good opportunities
'... (he) doesn't take advantage of a few good opportunities in each match...'
- (10) Zaobserwowano również przypadki rozdawania dwu egzemplarzy / *dwa egzemplarze
observed.IMPS also cases distributing two.GEN copies.GEN two.ACC copies.ACC
kart...
cards.GEN
'Also cases, where two cards were handed, were observed...'

And yet, despite these genitive positions, the forms occurring after *po* in (7)–(8) are not genitive, but must rather be analysed as accusative.²

It should be noted that judgements concerning the distribution of *po* are often hazy (see § 4 below), but the facts above seem relatively clear. For example, the corpus query [pos=ger] *po* [pos=num], asking for a gerund followed by *po* followed by a numeral, gave 98 results when run on the 300-million-segment balanced subcorpus of the National Corpus of Polish. Among these, there were 16 occurrences of the distributive *po* (there are other homographic lexemes in Polish, without the distributive meaning), and all of them were followed by accusative rather than genitive forms. Also native speakers, while sometimes not happy with (7)–(8), uniformly judge them as palpably more acceptable than versions with *po* followed by the genitive (not shown here).

This behaviour is different from the behaviour of the two elements *po* postulated above: neither is it a preposition assigning the locative case, nor is it some kind of a modifier invisible to case assignment. Rather, *po* in (7)–(8) behaves like another preposition *po*, assigning the accusative case.

On one hand, this is not very surprising, as in Polish there are exactly two non-distributive prepositions *po*, combining with the locative and the accusative case, respectively. So not just one of them is recycled with the distributive meaning (the one exemplified in (1), with the locative), but the other one as well (with the accusative, as in (7)–(8)). On the other hand, this conclusion is unwelcome, as now we are forced to say that in Polish there are not just two – as Przepiórkowski 2010 would have it – but in fact *three* functional lexemes with the same distributive meaning and the same form *po*. Clearly, postulating 3 unrelated lexical entries would have the bitter taste of a *missed generalisation*.

3 Solutions to the *missed generalisation* problem

3.1 Previous solution

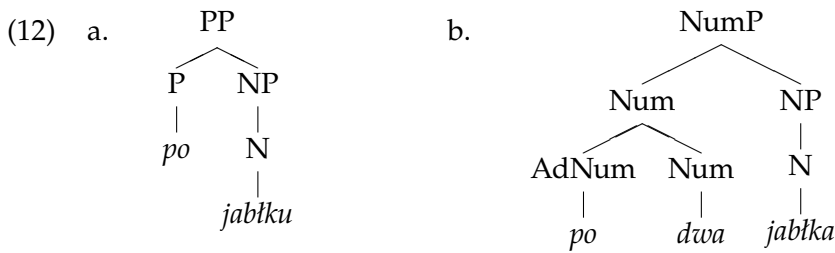
Przepiórkowski 2010, distinguishing between two distributive *po* elements, proposes an analysis couched in Sign-Based Construction Grammar (SBCG; Sag 2012) and seeks to maintain the generalisation by postulating a construction (*construct*, in SBCG parlance), (11), which combines distributive

²To be precise, the forms *kilka* and *dwa egzemplarze* are syncretic between accusative and nominative, but analogous examples may be constructed which make it clear that the accusative case is observed here.

semantics (not explicated in that paper) with the form *po*. The two observed distributive constructions, prepositional and adnumeral, are subtypes of this construction and make no further reference to the form *po* or to its semantic import.

$$(11) \text{ distr-cxt} \Rightarrow \left[\text{MTR} \left[\begin{array}{l} \text{FORM } \langle po, \dots \rangle \\ \text{SEM } (\text{distributive semantics}) \end{array} \right] \right]$$

What seems to be crucial in this proposal is that the form *po* is mentioned simply as the first word in the construction (the first element of FORM), i.e., in terms of linearisation, and not in terms of constituent structure (e.g., as the leftmost daughter). This is because Przepiórkowski 2010 assumes rather different syntactic structures of the two constructions:



The reason given for analysing the adnumeral *po* as modifying the numeral directly in (12b) instead of modifying the numeral phrase is the following contrast between a preposition in (13a–b),³ and the adnumeral operator in (14a–b) – only in the former can the adjective be placed in front of the numeral, within the numeral phrase, while in the latter the adnumeral operator combines with the numeral directly, thus forbidding such fronting:

- (13) a. Posłał go po 2 smaczne jabłka. (14) a. Dał każdemu po 2 smaczne jabłka.
 sent him for 2 tasty apples gave each one po 2 tasty apples
 ‘He sent him to fetch 2 tasty apples.’ ‘He gave each of them 2 tasty apples.’
- b. Posłał go po smaczne 2 jabłka. b. *Dał każdemu po smaczne 2 jabłka.
 sent him for tasty 2 apples gave each one po tasty 2 apples
- c. *Dał każdemu smaczne po 2 jabłka.
 gave each one tasty po 2 apples

This analysis seems to predict that the adjective may be placed before the numeral complex containing the adnumeral (as in (14c)), but this possibility is explicitly blocked by a stipulation that adnumeral operators only combine with non-gapped numerals, i.e., numerals with no extracted dependents.

3.2 Current solution

In the current analysis we dispute the importance of the data in (13)–(14). First, the acceptability contrasts are not as strong as the annotations in (14) would imply: while (14a) is clearly acceptable and (14c) is clearly out, (14b) has an intermediate status. In fact, Derwojedowa 2011, p. 145, judges an example similar to (14b) as grammatical (her (220b)).

Second, while we agree with Przepiórkowski 2010 that there is some acceptability contrast between (14a–b) that is lacking in (13a–b), a similar contrast may be observed in (15a–b), which involves the distributive preposition *po* assigning the accusative case, rather than an adnumeral operator.

- (15) a. ?Nie dał każdemu po 2 smaczne jabłka. b. ?*Nie dał każdemu po smaczne 2 jabłka.
 not gave each one po 2 tasty apples not gave each one po tasty 2 apples
 ‘He didn’t give each of them 2 tasty apples.’

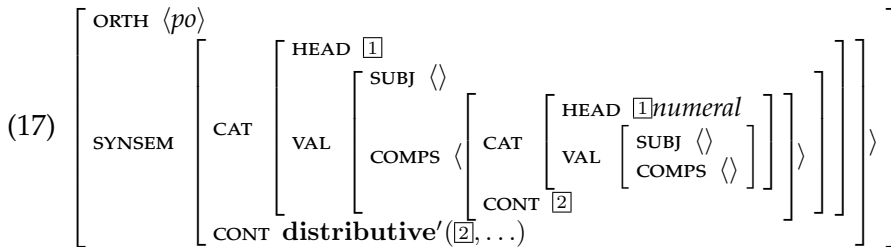
In conclusion, to the extent that (14b) and (15b) really are ungrammatical, we assume that a linearisation constraint common to all distributive elements *po* will have to take care of this effect.

³Note that this is *not* the distributive preposition *po*, but rather a homonymous preposition mentioned earlier.

Given these considerations, an HPSG analysis presents itself which is conceptually simpler than the SBCG analysis sketched above. First of all, even assuming the simplest approach to the HPSG lexicon, i.e., a Word Principle⁴ as in (16a) (where LE_i are lexical entries), the common part of the three distributive elements PO , i.e., minimally, its form and its semantics, could be factored out, as in (16b), where $\text{PO}_{\text{common}}^d$ stands for the description that is shared by the three distributive elements, and PO_{LOC}^d , etc., stand for descriptions that are idiosyncratic to them. Hence, generalisations are not missed at the description level.

- (16) a. $\text{word} \rightarrow LE_1 \vee LE_2 \vee \dots \vee LE_n$
 b. $\text{word} \rightarrow LE_1 \vee LE_2 \vee \dots \vee (\text{PO}_{\text{common}}^d \wedge (\text{PO}_{\text{LOC}}^d \vee \text{PO}_{\text{ACC}}^d \vee \text{PO}_{\text{ADNUM}}^d)) \vee \dots \vee LE_n$

Second, we assume that all three distributive elements are heads of respective PO -phrases. This is clear in case of the two prepositions, but needs to be explicated for the third PO , surfacing at least in (3)–(6). Here, we assume a version of HPSG without markers, as advocated by Tseng 2002. Hence, the adnumeral modifier PO is a weak(ish) head which takes over all relevant morphosyntactic features from its object, the numeral phrase; this is schematically shown in (17), which should be understood as a subset of the description “ $\text{PO}_{\text{common}}^d \wedge \text{PO}_{\text{ADNUM}}^d$ ” in the sense of (16b) (with semantics only indicated here).



Thanks to this analysis of the adnumeral PO , and in contradistinction to treating it as a modifier, the phrase following PO is analysed as its complement in all three cases, which simplifies the formalisation of its semantic contribution. Moreover, the structure-sharing of HEAD values has the effect of sharing case, number and gender, which accounts for the agreement observed in (4)–(6). There is, however, an important difference between (17) and typical weak heads of Tseng 2002, p. 273, or Abeillé *et al.* 2003, § 3.2: PO does not share the CONT value with its complement, but rather acts as a semantic functor. This means that the adnumeral PO is only syntactically weak. Hence, PO is analysed here in a way similar to the analysis of coordinating conjunctions in Abeillé 2006, § 5.2 (and earlier in Abeillé 2003, § 3).^{5,6}

The semantic impact of the distributive PO is not explicated in (17), as the semantics of distributivity is complex and a matter of ongoing work (cf., e.g., Zimmermann 2002 and Dotlačil 2012). The key problem, which has led to some non-compositional treatments of the semantically analogous binominal *each* in English (as in *I gave them an apple each*, with *each* arguably attaching to the preceding NP; Safir and Stowell 1988), is that – apart from the nominal or numeral phrase to which such a distributive element attaches (so-called *distributed share*; Choe 1987), e.g., *jabłku* ‘an apple’ in (1) – it also takes another semantic argument, which occurs elsewhere in the sentence (called *sorting key* in Choe 1987), e.g., *im* ‘them’ in (1). Moreover, contrary to what might be suggested by the simple constructed example (1), the *sorting key* may be both linearly and configurationally distant from PO (see, e.g., (7), where the *sorting key* is embedded within an adjunct PP), may be implicit ((8) could perhaps be analysed this way, with the missing dative complement of *rozdawania* ‘distributing, handing’ acting as the *sorting key*) and may even be contained within the *distributed share* itself (as in (3) and (6)).

While we do not have a detailed analysis of the semantics of distributive PO to offer at present, we envisage that the apparently non-compositional effects could be formalised in HPSG in terms of the

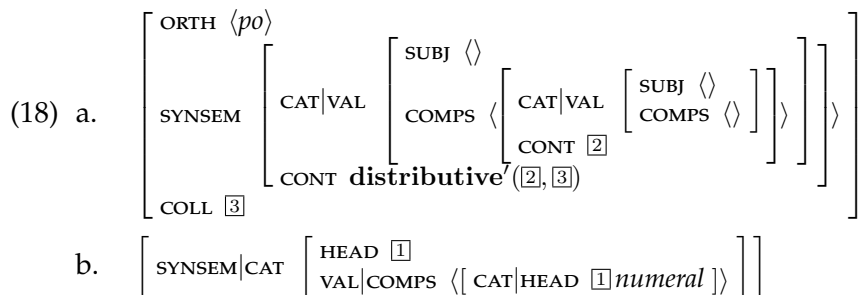
⁴See, e.g., Höhle 1999 and Meurers 1999 for discussion.

⁵Given the existence of such elements, it seems that *weak heads* in the sense of Tseng 2002 is a derived notion and can be defined as the intersection of the set of *syntactically weak heads* and the set of *semantically weak heads*, i.e., known in HPSG as *semantically vacuous elements* (cf., e.g., Pollard and Yoo 1998 and Przepiórkowski 1998).

⁶An alternative unified analysis could be attempted by treating all PO lexemes as *minor adpositions* of Van Eynde 2004, i.e., as functors. Such an analysis would be more controversial, though, as it would result in the existence of bare locative NPs (selected and marked by PO), even though elsewhere locative is assigned only by prepositions in Polish. Nevertheless, this possibility is worth exploring further, as the distributive elements discussed here seem to have all the features and secondary characteristics of minor adpositions.

COLL feature (Richter and Sailer 1999), as explicated in Sailer 2003, §8.2, possibly with restrictions argued for in Soehn 2004. The common part of the lexical entries for the distributive PO, the “PO_{common}^d” of (16b), will make a reference to the value of COLL, i.e., to the whole utterance (Sailer 2003) or its appropriate constituent (Soehn 2004), ensuring that it contains a suitable *sorting key* and incorporating it into its semantics.

In summary, PO_{common}^d can be schematically represented as in (18a), with the idiosyncratic contribution of PO_{ADNUM}^d indicated in (18b) (and similarly for PO_{LOC}^d and PO_{ACC}^d, omitted here for lack of space).



4 A note on the distribution of po

In this paper we have concentrated on the internal syntax and morphosyntax of PO-phrases. Their external syntax, i.e., their syntactic distribution, is equally baffling. First of all, it is not clear how prepositional PO-phrases (i.e., those headed by words described by PO_{LOC}^d and PO_{ACC}^d in (16b)) are licensed in case position. For the time being we adopt the *not my problem* approach, as in Polish there are also other kinds of PPs – and even more complex prepositional constructions – regularly occurring in case positions, e.g., *Na koncert przyjdzie około dwustu osób* ‘Around 200 people will come to the concert’, lit. ‘on concert will-come around 200.GEN people.GEN’, where the genitive on the numeral (instead of the expected accusative) shows that *około* ‘around’ is a preposition here.

Second, the distribution of prepositional PO-phrases is very constrained. Przepiórkowski 1999, §5.1.4.3, claims that they may occur exactly in structural case positions (in Polish these would be nominative, accusative and these genitive positions which alternate with the accusative), and (7)–(8) seem to confirm this, but if so, the differences between their distribution and that of other “structural” phrases considered there must be accounted for. Also, it is not clear whether PO_{ACC}^d may occur in non-structural (i.e., inherent) genitive positions – acceptability judgements are particularly murky here. Finally, we currently see no way of deciding whether accusative PO-phrases in accusative positions, as in (2) and (14), involve the preposition governing the accusative case or the adnumeral operator transparent to case assignment.

Third, also the distribution of the adnumeral PO is very limited; for example, the unacceptability of examples similar to (7)–(8) but with genitive forms following PO suggests that the adnumeral PO cannot occur in structural genitive positions, while (4)–(6) show that they easily occur in the nominative. On the other hand, as illustrated in (3), they may occur in the dative, but – as both native acceptability judgements and failed corpus searches suggest – for some reason not in the instrumental.

As acceptability judgements regarding the syntactic distribution of PO-phrases are often not clear, it seems that wide-scale psycholinguistic experiments are needed to conclusively establish empirical facts here.

5 Conclusion

Some aspects of the work reported here are finished, other are in progress. We establish the fact that there are three morphosyntactically different distributive elements PO in Polish – this is a novel empirical contribution of this paper, however narrow its focus. While the general treatment of this potential case of *missed generalisation*, given in (16b), is trivial, arguing that all three elements may be treated as heads and that the HPSG approach to phraseology based on COLL may be put to work here is perhaps interesting. Other aspects, including the formalisation of the distributive semantics and the empirical investigation into the syntactic distribution of the three elements, are a matter of ongoing work, with some results hopefully available by the time of HPSG 2013.

References

- Abeillé, A. (2003). A lexicon- and construction-based approach to coordinations. In S. Müller, ed., *Proceedings of the HPSG 2003 Conference*, pp. 5–25, Stanford, CA. CSLI Publications.
- Abeillé, A. (2006). In defense of lexical coordination. *Empirical Issues in Syntax and Semantics*, 6, 7–36.
- Abeillé, A., Bonami, O., Godard, D., and Tseng, J. (2003). The syntax of French *à* and *de*: an HPSG analysis. In *Syntax and semantics of prepositions*, pp. 147–162, Dordrecht. Springer-Verlag.
- Borsley, R. D. and Przepiórkowski, A., eds. (1999). *Slavic in Head-Driven Phrase Structure Grammar*. CSLI Publications, Stanford, CA.
- Choe, J.-W. (1987). *Anti-quantifiers and a theory of distributivity*. Ph. D. dissertation, University of Massachusetts, Amherst.
- Derwojedowa, M. (2011). *Grupy liczebnikowe we współczesnym języku polskim. Zarys opisu zależnościowego*. Uniwersytet Warszawski, Wydział Polonistyki, Warsaw.
- Dotlačil, J. (2012). Binominal *each* as an anaphoric determiner: Compositional analysis. In A. A. Guevara, A. Chernilovskaya, and R. Nouwen, eds., *Proceedings of Sinn und Bedeutung 16: Volume 1*, vol. 16 of *MIT Working Papers in Linguistics*, pp. 211–224.
- Grochowski, M. (1997). *Wyrażenia funkcyjne. Studium leksykograficzne*. Wydawnictwo IJP PAN, Cracow.
- Höhle, T. N. (1999). An architecture for phonology. In Borsley and Przepiórkowski (1999), pp. 61–90.
- Łojasiewicz, A. (1979). O budowie wyrażen z przyimkiem *po* dystrybutywnym. *Polonica*, V, 153–160.
- Meurers, W. D. (1999). *Lexical Generalizations in the Syntax of German Non-Finite Constructions*. Ph. D. dissertation, Universität Tübingen.
- Pollard, C. and Yoo, E. J. (1998). A unified theory of scope for quantifiers and *wh*-phrases. *Journal of Linguistics*, 34, 415–445.
- Przepiórkowski, A. (1998). ‘A Unified Theory of Scope’ revisited: Quantifier retrieval without spurious ambiguities. In G. Bouma, G.-J. M. Kruijff, and R. T. Oehrle, eds., *Proceedings of the Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar, and Categorical Grammar*, pp. 185–195, Saarbrücken.
- Przepiórkowski, A. (1999). *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Ph. D. dissertation, Universität Tübingen, Tübingen.
- Przepiórkowski, A. (2010). Towards a Construction Grammar account of the distributive *po* in Polish. *Études Cognitives*, 10, 163–176.
- Richter, F. and Sailer, M. (1999). LF conditions on expressions of Ty2: An HPSG analysis of Negative Concord in Polish. In Borsley and Przepiórkowski (1999), pp. 247–282.
- Safir, K. and Stowell, T. (1988). Binominal *each*. In *Proceedings of the North East Linguistic Society (NELS)*, Vol. 18, pp. 426–450.
- Sag, I. A. (2012). Sign-based construction grammar: An informal synopsis. In H. Boas and I. A. Sag, eds., *Sign-Based Construction Grammar*, pp. 69–202. CSLI Publications, Stanford, CA.
- Sailer, M. (2003). *Combinatorial Semantics and Idiomatic Expressions in Head-driven Phrase Structure Grammar*. Ph. D. dissertation, Universität Tübingen.
- Soehn, J.-P. (2004). License to *COLL*: How to bind bound words and readings to their contexts. In S. Müller, ed., *Proceedings of the HPSG 2004 Conference*, pp. 261–273, Stanford, CA. CSLI Publications.
- Tseng, J. (2002). Remarks on marking. In F. van Eynde, L. Hellan, and D. Beermann, eds., *Proceedings of the HPSG 2001 Conference*, pp. 267–283, Stanford, CA. CSLI Publications.
- Van Eynde, F. (2004). Minor adpositions in Dutch. *Journal of Comparative Germanic Linguistics*, 7, 1–58.
- Zimmermann, M. (2002). ‘Boys Buying Two Sausages Each’ – On the Syntax and Semantics of Distance Distributivity. Ph. D. dissertation, Universiteit van Amsterdam.