Counting minimal groups: towards a unified semantics for the Polish measure word \textit{para}

Marcin Wągiel
\textit{Palacký University, Olomouc; Masaryk University, Brno}

Abstract. In this squib I discuss a puzzle concerning the Polish measure word \textit{para} ('pair of') and its supposedly heterogeneous behavior in phrases in which it combines with regular count nouns and, on the other hand, with pluralia tantum. The puzzle concerns the different cardinality of denoted objects and the relationship between atomicity and collectivity. Building on the theory of Krifka (1995) which I extend with the group-forming operation developed by Landman (1989) I propose a unified semantic interpretation of \textit{para}. I posit that the denotation of regular count nouns includes only sums of individuals whereas pluralia tantum denote complete semi­lattices. In each case, \textit{para} selects a minimal element of the denotation of a noun, i.e., either a sum consisting of 2 individuals or an atom, and turns it into a group which can be further counted.

Keywords. measure words, pluralia tantum, collectivity, atomicity, plurality

1. Introduction

In some languages, measure words allowing for counting entities denoted by pluralia tantum and group nouns referring to collections of two individuals are homophonous, conf. English \textit{pair}, Russian \textit{para}, Spanish \textit{par} etc. In this squib I propose a unified semantic analysis for the Polish measure word \textit{para} ('pair of') which can shed some more light on the relationship between atomicity, counting, and collectivity.

The squib is outlined as follows. In Section 2 I discuss Polish data concerning the behavior of \textit{para} in phrases with pluralia tantum and regular count nouns, specifically the cardinality of denoted entities and issues related to collectivity and distributivity. In Section 3 I introduce the framework of Krifka (1995) which I extend with the group-forming operation postulated by Landman (1989). In Section 4 I sketch a proposal for a unified semantics of Polish \textit{para} and discuss some of its advantages. Section 5 concludes the squib.

\footnote{I would like to thank the audience of the 7th Conference on Syntax, Phonology, and Language Analysis (SinFonIJA VII) for their helpful questions and comments. I am especially grateful to Sarah Zobel for some technical hints.}
2. Puzzle

It is widely known that pluralia tantum are systematically ambiguous between singular and plural readings. In this section I will examine the case of pluralia tantum nouns in Polish.

(1) a. Nożyczk_i_ leźały na stole. → 1 object

scissors leźały on table-LOC → more than 1 object

'Scissors lied on the table.'

b. Książki leźały na stole. → *1 object

books lied lied table-LOC → more than 1 object

'Books lied on the table.'

Though the only accessible reading of (1b) is that there was more than one book on the table, the sentence in (1a) is true both in a scenario where there are several pairs of scissors on the table and in a scenario in which there is only one pair of scissors on the table. This fact may suggest that there is a difference between domains within which regular count nouns and pluralia tantum refer.

Another well-known fact is that similarly to many other languages, Polish pluralia tantum can never combine directly with basic cardinal numerals. A specialized numeral, see (2b), or the measure word para ('pair of'), see (2c), is required.

(2) a. *dwa

nożyczk_i_ two scissors

b. dwoje

nożyczek two-OJE scissors-GEN
c. dwie

pary nożyeczek two pars scissors-GEN

However, the measure word para can also combine with regular count nouns denoting animate beings or objects which typically occur in pairs, e.g., shoes, socks, gloves. Interestingly, phrases in which para combines with pluralia tantum do not show ambiguity with respect to the singular/plural interpretation, i.e., they refer to singular objects, see (3a). On the other hand, phrases where para combines with regular count nouns denote collections of 2 objects, see (3b)-(3c).3

One should notice that in comparison to the basic numeral dwa ('two') in (2a), the specialized numeral dwoje ('two') in (2b) is morphologically more complex. It is, thus, legitimate to assume that basic cardinal numerals are semantically simpler, i.e., lack a semantic property allowing for counting entities denoted by pluralia tantum. On the other hand, a specialized suffix attached to the root in numerals such as dwoje behaves as para in (2c), i.e., introduces an operator which allows for such quantification (cf. Wagiel 2014).

3 In some Slavic languages phrases such as (3b), e.g., Czech par studentů, Serbian par studenata, have yet another reading where they mean 'several students'. In Polish such a reading is unavailable since a grammaticalized Accusative form of para, i.e., para/paru ('several'), is used in this context, e.g., paru studentów ('several students'). It has different syntactic and semantic properties than para and it seems that Czech par and Serbian par in the phrases presented above are also grammaticalized Accusative forms. Though I believe that the
(3) a. *para nożyczek* — 1 object
   pair scissors-GEN
   'pair of scissors'

   b. *para studentów* — 2 objects
   pair students-GEN
   'pair of students'

   c. *para butów* — 2 objects
   pair shoes-GEN
   'pair of shoes'

Moreover, although sentences with basic cardinal numerals are ambiguous between
distributive and collective interpretations, see (4a), sentences with NPs such as (3b)—
(3c) do not show such ambiguity and have only collective readings, see (4b).*

(4) a. Dwóch studentów napisało esej.
   Two students-GEN wrote-3.SG.N essay
   'Two students wrote an essay.'

   b. Para studentów napisała esej.
   pair-F students-GEN wrote-3.SG.F essay
   'A pair of students wrote an essay.'

(4a) can be interpreted either as if both students wrote an essay together, i.e., a total
of one essay was written, or as if each student wrote an essay on their own, i.e., a
total of two essays were written. (4b) lacks the second reading and can only be
interpreted in a collective manner.

The data presented above suggest that for a proper treatment of *para*, one
should assume that in Polish there are in fact two homonymous lexical items the first
one being a measure word and the second one a group noun similar to nouns such as
committee. However, such an explanation of the supposedly heterogeneous behavior
of *para* does not seem to be plausible since similar homonymy would have to be
assumed cross-linguistically. Therefore, a unified semantic analysis is desirable. Such
an analysis should account for three issues: i) the fact that *para* allows for counting
of entities denoted by pluralia tantum though basic numerals do not, ii) different
cardinality of objects denoted by phrases such as (3a) and (3b)–(3c), and
iii) obligatory collective interpretations of sentences such as (4b). In this squib I
provide a neat semantic approach that accounts for these issues. Before we propose a
draft of a possible solution, let us introduce some theoretical background.

*Note that in (4a) the verb is set to default whereas in (4b) it agrees with *para*.
3. Theoretical Framework

3.1. Common Nouns

I will adopt here an extensionalized version of the theory of Krifka (1995). In this framework bare nouns are basically expressions referring to kinds or, more generally, to concepts. One can derive count nouns in languages such as English or Polish by interpreting them as predicates denoting objects realizing a kind via the realization relation $R$ (Carlson 1977). Assuming that $x$ is an object-level individual and $k$ is a kind, $R(x, k)$ simply states that $x$ is a specimen of $k$. Such specimens can be counted by the OU function, i.e., object unit operation, which selects a kind and gives back a measure function that measures the number of its specimens, e.g., if $x$ consists of two objects realizing a kind $k$, then $OU(k)(x) = 2$.

Krifka posits that there is no semantic difference between representations of singular and plural expressions. Nevertheless, in this squib I will follow the line of Link (1983) and many others who assume that there is such a distinction. In the theory of Link singular terms refer within the domain of atoms whereas plural terms are assumed to refer within the domain of atoms and sums. However, following Hoeksema (1983) and Chierchia (1998a, b) and unlike Link (1983) I will assume that atoms are not included in the denotation of plural terms, i.e., plural nouns denote sums only. Let us assume the following interpretations of singular and plural count nouns in Polish.

\[
\begin{align*}
\text{a. } & \text{[student]} = \lambda x[R(x, \text{Student}) \land \text{OU(Student)}(x) = 1] \rightarrow \text{atoms} \\
\text{b. } & \text{[studenci]} = \lambda x[R(x, \text{Student}) \land \text{OU(Student)}(x) \geq 2] \rightarrow \text{sums}
\end{align*}
\]

Fig. 1: Denotations of singular and plural count nouns

As one can see in Fig. 1, singular terms denote atoms only, while plural terms denote sums only. Thus, in the model where there are only three atomic individuals having the property of being a student, e.g., a, b, and c, the denotation of a singular predicate student ('student') would be \{a, b, c\}. On the other hand, a plural noun such as

---

5 For sake of simplicity, I will ignore the distinction between kinds and concepts here and I will use the term 'kind' as referring to both kinds and concepts.
students ('students') would denote \{aUb, aUc, bUc, aUbUc\}, where \( \cup \) is a typical join (sum) operator. Sums, of course, consist of atomic subparts, such subparts themselves, however, are not in the denotation of a plural term.

Mass nouns, however, are semantically less complex. They can be interpreted as having no \( O \) function incorporated in their semantics; hence to allow for counting entities denoted by mass nouns measure terms introducing appropriate measure functions are required. Moreover, let us follow the mereological treatment of mass nouns assuming that they denote complete semi-lattices (Chierchia 1998a). I adopt a definition of a complete semi-lattice proposed by Bale & Khanjian (2008).

\[ (7) \text{ A denotation } X \text{ is a complete semi-lattice iff for all members } y \text{ and } z \text{ of } X, \]
\[ y \cup z \text{ is a member of } X \text{ and, if } y \cap z \text{ is not the empty sum (0), then } y \cup z \text{ is a member of } X. \]

The symbol \( \cup \) in the definition above corresponds to the typical join (sum) operator whereas \( \cap \) is the typical meet operator. A mass noun \textit{piach} ('sand') would then translate as follows.

\[ (6) \text{ [piach]} = \lambda x[R(x, Sand)] \rightarrow \text{ complete semi-lattice no slot for counting} \]

The mass noun \textit{piach} ('sand') denotes both atomic units of sand and all possible sums formed from these atoms, i.e., portions of sand. If in a given context grains of sand were a, b, and c, then it would denote \{a, b, c, aUb, aUc, bUc, aUbUc\}. In such a case the join of any two members of the denotation is itself a member of the denotation. The meet of any two members is either the empty set (in the case of a\cap b, a\cap c, and b\cap c) or a member of the denotation as well (in all other cases).

### 3.2. Groups

The theory of Krifka (1995) can be further extended with the group-forming operation \( \dagger \) (Landman 1989, 2000) which can account for collective readings of non-quantificational plural NPs (Landman 1997). A group-forming operation maps a sum onto a group, i.e., a complex atomic individual in its own right. The definition of \( \dagger \) is given in (8) and Fig. 2 illustrates its mechanics.

\[ (8) \text{ \dagger is a one-one function from SUM into ATOM such that:} \]
1. \( \forall d \in \text{SUM-IND}; \dagger(d) \in \text{GROUP} \)
2. \( \forall d \in \text{IND}; \dagger(d) = d \)

---

\(^6\) Bale & Khanjian (2008) use \( \vee \) and \( \wedge \) symbols respectively.
The group-forming operator \( \triangleright \) takes a sum of individuals as its argument and returns a group, i.e., a complex atomic entity whose inner structure is inaccessible. In Landman’s system singular predication is collective predication, hence the fact that a group behaves as if it was an atom accounts for collectivity.

Having these tools in place let us consider a possible solution to the puzzle described in Section 2.

4. Proposal

First of all, following one of the possible interpretations of English numerals proposed by Krifka (1995) I assume that basic cardinal numerals in Polish simply denote numbers, i.e., they have no integrated measure function. For example, a numeral \( \text{dwa} \) ('two') merely denotes 2.

Second, I posit that pluralia tantum are semantically less complex than regular plural terms and have semantics similar to mass nouns. That is to say that in both cases the denotation of a noun is a complete semi-lattice and that it has no built-in measure function which would allow for counting.

\[
(9) \quad [\text{nóżyczki}] = \lambda x[R(x, \text{Scissors})] \quad \rightarrow \quad \text{complete semi-lattice, no slot for counting}
\]

The proposed semantics accounts for the fact that Polish pluralia tantum enable both singular and plural interpretations in any context and explains why they cannot combine with basic cardinal numerals. However, it should be emphasized that I do not claim that entities denoted by pluralia tantum and mass nouns form an ontologically uniform class. In fact they are very different. Although it might be unknown what an atom of an entity denoted by a mass noun is, an atom of an entity denoted by a plurale tantum noun is well defined and easily identifiable.\(^7\) Nevertheless, since a plurale tantum noun denotes both sums and atoms, an

\(^{7}\text{Chierchia (1998a, 2010) suggests that unlike count nouns mass expressions have unstable semantic building blocks, i.e., it is vague whether elements of the denotation represent atoms or sums, and that this accounts for the fact they cannot be counted. With respect to this aspect, pluralia tantum seem to be similar to the so-called fake mass nouns such as furniture.}\)
additional means is required to distinguish between atomic and non-atomic individuals.

Third, let us introduce a minimal element operator Min which is true for an individual being a minimal element of the denotation of a noun. A minimal element can be defined in the following way:

\[(10) \text{An individual } y \text{ is a minimal element of a denotation } X \text{ iff for all members } y \text{ and } z \text{ of } X, \text{ if } y \geq z, \text{ then } y = z\]

The motivation behind introducing the Min operator concerns the fact that entities denoted by pluralia tantum cannot be counted with cardinals. Given that they are semantically similar to mass nouns, a means to define what part of the denotation can be used as a counting unit is required. At first sight, the Min operator might seem to be similar to standard atomic element functions, e.g., AT in the system of Chierchia (2010). Nevertheless, it is crucial that Min is concerned barely with minimal entities which are still members of the denotation of a noun and not with ultimately minimal entities, i.e., atoms. For example, let us assume a plural individual \(x\) such that it is a member of the denotation \(X\) and consists of two singular, i.e., atomic, individuals \(y\) and \(z\), but neither \(y\) nor \(z\) belongs to \(X\). Though both \(y\) and \(z\) are subparts of \(x\), it is not the case that either \(y\) or \(z\) is a minimal element of \(X\) just because they are not members of \(X\).

Let us now propose a unified semantics for the Polish measure word \textit{para} ('pair of').

\[(11) [\text{para}] = \lambda n \lambda P . \lambda x [\text{Min}(x) \land \uparrow(x) \land |\uparrow(x)| = n \land P] \]

The measure word \textit{para} selects a minimal element from the denotation of a noun and forms a group-atom which can be further counted via the measure function \(|...|\) which returns the cardinality of entities it is applied to. Since atoms are not included in the denotation of plural regular count nouns, when \textit{para} combines with such a noun, it takes the smallest sum, i.e., a sum consisting of 2 atomic elements, e.g., \(aUb\), \(aUc\), \(bUc\), and forms a group, e.g., \(\uparrow(aUb)\), \(\uparrow(aUc)\), \(\uparrow(bUc)\). This accounts for the fact that phrases such as (3b)-(3c) denote pairs of objects and have obligatorily collective readings. Such minimal groups can further be counted by the \(|...|\) function. On the other hand, when \textit{para} combines with a plurale tantum noun, i.e., a noun whose denotation is a complete semi-lattice, it has to select an atom, e.g., \(a\), \(b\), \(c\), and not until then it forms \(\uparrow a\), \(\uparrow b\), \(\uparrow c\) which is by definition equivalent to \(a\), \(b\), \(c\), see (8). Furthermore, the cardinality function allows for counting objects denoted by pluralia tantum. This explains why phrases such as (3a) denote singular objects and why phrases such as (2c) are well-formed whereas phrases such as (2a) are not.

5. Conclusion

In this squib I have proposed a draft of a unified semantics for the Polish measure word \textit{para} ('pair of'). The analysis is based on the framework of Krifka (1995) which I have extended to include a group-forming operation \(\uparrow\) from the theory of Landman (1989, 2000). I have proposed that \textit{para} selects only minimal elements of the
denotation of a noun with which it combines, technically via a minimal element operation Min. Furthermore, I have posited that para has a built in cardinality function $\mid \ldots \mid$ and that it introduces the $\uparrow$ operator. Since plural regular count nouns denote sums only while the denotation of pluralia tantum is a complete semi-lattice, in the first case para takes a minimal sum and transforms it into a group whereas in the latter case it has to select an atom. This explains the allegedly heterogeneous behavior of phrases containing para and regular count nouns on the one hand, and phrases with para and pluralia tantum on the other.

Although a parallel between Polish and at least some of the languages in which there is homonymy between measure terms allowing for counting pluralia tantum and group nouns referring to collections of two entities can be drawn, due to limited space I restricted my focus to Polish data exclusively. Thus, possible extensions of the idea developed here have to be left for further research.

References


Counting minimal groups: towards a unified semantics for the Polish measure (eds.), *Nominal Structures: All in Complex DPs*. Olomouc: Palacký University, pp. 69-84.