Preposition-Pronoun Contraction in Polish

Beata Trawiński

SFB 441
Universität Tübingen
trawinski@sfs.uni-tuebingen.de

The Second ACL-SIGSEM Workshop on
The Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications
University of Essex, Colchester, United Kingdom
April 19th-21st, 2005
The Phenomenon

Polish prepositions exhibit the ability to coalesce with the third person pronouns they select for.

(1)  
   a. na niego ‘on him’ $\rightarrow$ nań ‘on_him’  
   b. w niego ‘in him’ $\rightarrow$ weń ‘in_him’

The treatment of PPC relates to the following issues:

$\rightarrow$ Does PPC pertain exclusively to the phonological or, more precisely, prosodic domain, or is it also subject to any syntactic processes?

$\rightarrow$ Should PPC be analyzed in terms of affixation or cliticization?
Overview

- Empirical Data
  - Case, Gender and Number Restrictions
  - Categorial, Syntactic, Semantic and Prosodic Properties
  - Lexical Restrictions
- The HPSG Analysis
- Summary and Outlook
There are seven morphological cases in contemporary Polish:
- nominative,
- genitive,
- dative,
- accusative,
- instrumental,
- locative and
- vocative.

Cases assigned by prepositions:
- nominative,
- genitive,
- dative,
- accusative,
- instrumental,
- locative.
Genitive and Accusative

According to traditional approaches to PPC, only genitive and accusative singular masculine third person pronouns can contract with prepositions to form PPCs (cf. e.g., (Saloni 1981)).

(2) a. Maria pracuje dla Piotra / dla niego / dlań od roku.
   Maria works for Piotr gen / for him gen / for_him for year
   ‘Maria has been working for Piotr / for him for one year.’

b. Jan zapłacił za obiad / za niego / zań gotówką.
   Jan payed for dinner acc / for it acc / for_it cash
   ‘Jan payed for the dinner / for it in cash.’

(Doroszewski and Wieczorkiewicz 1972) even claim that PPCs are not allowed in the case of Ps requiring any other cases than accusative. To that effect, PPC in (2a) should be considered ungrammatical. However, according to the latest dictionaries of the Polish language such as (Dubisz 2003) or (Bańko 2000), as well as on the basis of corpus evidence, we consider prepositions contracted with genitive third person pronouns fully grammatical.
The following examples seem to support the assumption that only genitive and accusative singular masculine third person pronouns can contract with prepositions: Neither nominative nor dative or instrumental pronouns are possible within PPCs.

(3) a. Piotr tańczy jak Michael Jackson / jak on / *jakoń.
Piotr dances like Michael Jackson$_{\text{nom}}$ / like he$_{\text{nom}}$ / like he
‘Piotr dances like Michael Jackson / like he.’

b. Maria pochyliła się ku Piotrowi / ku niemu / ??kuń.
Maria leaned RM towards Piotr$_{\text{dat}}$ / towards him$_{\text{dat}}$ / towards him
‘Maria leaned towards Piotr / towards him.’

c. Anna szła przed Piotrem / przed nim / *przedeń obrażona.
Anna walked before Piotr$_{\text{instr}}$ / before him$_{\text{instr}}$ / before him offended
‘Anna walked offended before Piotr / before him.’
However, contrary to commonly accepted claims, PPCs seem to be possible with locative pronouns as well:

(4) Polskie przedsiebiorstwa przemysłowe najlepiej prosperują w Mazowieckiem i Śląskiem. Pod koniec sierpnia br., uzyskało weź employment altogether above 1 million persons [...] ‘Polish industrial companies best prosper in Mazowieckie and Śląskie. By late August of this year, a total of over 1 million people gained employment there.’
(5) Taką wiedzę trzeba odpowiednio zdobyć, a to nie jest tak łatwe jak przeczytanie książki o potworach i zdobycie oń wiedzy.

‘One must gain such a knowledge in an appropriate way and it is not as easy as to gain knowledge about monsters by reading a book about them.’

Although the common assumed restrictions on the case, gender and number of pronouns contracting with Ps seem to be too strong, we will temporarily adopt the common claim that only

- genitive and accusative
- masculine (human, animal, inanimate)
- singular

third person pronouns can contract with Ps.
Syntactic Status

PPCs cannot combine with any further NPs:

(6) a. Maria pracuje dlań (*Piotra) od roku.
    Maria works for_him Piotr_{gen} for year
    ‘Maria has been working for him for one year.’

   b. Anna robi zań (*Jana) wszystko.
    Anna does instead of_him Jan_{acc} everything
    ‘Anna does everything instead of him.’

This fact indicates that PPCs are syntactically saturated expressions corresponding to ordinary PPs.
Categorial Status

While examples where PPCs act as adjuncts, may suggest the possibility that PPCs are adverbs, the following data seem to argue for a prepositional status of PPCs. Here PPCs are selected by predicates which seem to always require PP complements headed by specific prepositions.

(7) a. Piotr pracuje w firmie, która *(doń) należy.
Piotr works in company that to_him belongs
‘Piotr works in a company that belongs to him.’

b. Anna namawia Jana na ślub, ale wciąż nie może *(nań)
Anna persuade Jan for marriage but still not can on_him
skutecznie wpłynąć.
successfully affect
‘Anna has been trying to persuade Jan to marry her, but she still cannot successfully move him.’

These observations, supported by the fact that the phonological form of each PPC contains a string identical to that of the corresponding preposition, possibly indicate that PPCs have the same syntactic and categorial status as the corresponding PPs.
PPCs as Pronouns? I

In the following examples PPCs seem to behave like ordinary personal pronouns controlled by referential NPs.

(8) a. Mariaₕ umówila się z Piotremₖ i czekała
Maria made an appointment RM with Piotr and waited
nań*ᵢ/j w mieście.
for_her / him in city
‘Maria made an appointment with Piotr and waited for him in the city.’

b. [Anna i Maria]ₖ zapewnily Janaₗ, że proₖ wierzą weń*ᵢ/j.
Anna and Maria assured Jan that pro believe in_each other / him
‘Anna and Maria assured Jan that they believe in him.’
(9) a. Anna_
_ i obiecała Janowi_
_ j PRO i troszczyć się oń *i / j.
Anna promised Jan PRO care RM for_her / him
‘Anna promised Jan to care for him.’

b. Jan_
_ i kazał Annie_
_ j PRO j troszczyć się oń i / *j.
Jan made Anna PRO care RM for_him / her
‘Jan made Anna care for him.’

The anaphoric reading (both reflexive and reciprocal) is not possible!

→

The reference data refer to semantic properties of PPCs and are not the decisive factor
in determining their categorial status and establishing their syntactic characteristics.
They merely indicate that the semantic representation of PPCs must contain an index.
This also applies to ordinary PPs.
Thus we will treat PPCs as syntactically saturated PPs.
No modification, coordination and occurrence in sentence initial position is allowed in structures involving PPCs:

(10) a. Jan użył cytatu pochodzącego od niego samego / *odeń
   Jan used citation coming from himself / from_him
   samego.
   self
   ‘Jan used one of his own citations.’

b. Anna zapytała Jana, czy ma czekać na niego i jego kolegę / *nań i jego kolegę.
   Anna asked Jan whether she should wait for him and his colleague / for_him and his colleague
   ‘Anna asked Jan whether she should wait for him and his colleague.’

c. Dla niego / *dlań Maria pracuje od roku.
   for him / for_him Maria works for year
   ‘For him, Mary has been working for one year.’

→ Unlike typical PPs, PPCs (as phonologically / prosodically deficient expressions) cannot be used in stressed positions.
Lexical Restrictions I

The set of prepositions which are able to contract with pronouns involves a very limited number of elements. Thus not every arbitrary preposition can occur in PPC, even if it does combine with genitive or accusative pronouns:

(11) Ps combining with genitive pronouns
a. koło niego ‘close to him’ → *kołoń
b. naprzeciw niego ‘opposite to him’ → *naprzeciweń
c. podczas niego ‘during him’ → *podczaseń

(12) Ps combining with accusative pronouns
a. między nich ‘between them’ → *międzyń
b. ponad niego ‘above him’ → *ponadeń
c. poprzez niego ‘through him’ → *poprzezeń
Lexical Restrictions II

According to dictionaries of contemporary Polish and considering corpus data, one can claim that only the oldest primary prepositions can contract with pronouns:

(13) a. bez niego ‘without him’ → bezeń
b. dla niego ‘for him’ → dlan
c. do niego ‘to him’ → don
d. o niego ‘about him’ → on

e. od niego ‘from him’ → oden
f. na niego ‘on him’ → nan
g. nad niego ‘above him’ → nadeń
h. pod niego ‘under him’ → podeń
i. przed niego ‘in front of him’ → przedeń
j. przez niego ‘owing him’ → przezeń
k. w niego ‘in him’ → wen
l. z niego ‘from him’ → zen

However: u niego ‘by him’ → *un
Summary of Empirical Observations

- Only genitive and accusative masculine (human, animal, inanimate) singular third person pronouns can contract with Ps.

- PPCs have the categorial status of a preposition.

- PPCs are valence-saturated but have an internal argument.

- Only particular prepositions can occur in PPCs.
Lexicalist Treatment

We propose to treat PPCs lexically, as morphological units and not as the result of any postlexical processes.

Other constraint-based approaches to similar phenomena:

- contraction of French prepositional forms à and de and the definite article le, (Abeillé, Bonami, Godard, and Tseng 2003), also (Baronian 2003);
- preposition-determiner contraction in German, (Winhart 1997);
- contracted auxiliaries in English (lexical rule), (Bender and Sag 2000);
- French pronominal clitics, (Abeillé, Godard, and Sag 1998) and (Miller and Sag 1997);
- Italian pronominal clitics, (Monachesi 1999);
- Polish complex verbs, (Borsley 1999);
- Polish verbal negation, (Kupść and Przepiórkowski 2002).
Theoretical Framework

We will use Head-Driven Phrase Structure Grammar in the tradition of (Pollard and Sag 1994) as a theoretical framework.

Advantages of HPSG: HPSG is a
- lexicalist (it offers the possibility to determine the properties of both words and phrases on the word level),
- comprehensive (it is possible to encode generalizations about all linguistic representation levels simultaneously, thereby accounting for a possible interaction between the particular levels),
- fully formalized (cf. (Richter 2000)),
- computer-applicable linguistic formalism.

Crucial property of HPSG: It is a non-derivational constraint-based grammar framework.
### An AVM Description of Linguistic Signs according to (Pollard and Sag 1994)

<table>
<thead>
<tr>
<th>PHONOLOGY</th>
<th>phonological structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNSEM</td>
<td>LOCAL</td>
</tr>
<tr>
<td></td>
<td>CONTENT</td>
</tr>
<tr>
<td></td>
<td>CONTEXT</td>
</tr>
<tr>
<td></td>
<td>NONLOCAL</td>
</tr>
<tr>
<td>DAUGHTERS</td>
<td>constituent structure</td>
</tr>
<tr>
<td>ARG-ST</td>
<td>argument structure</td>
</tr>
</tbody>
</table>

- **HEAD**: part of speech, etc.
- **VALENCE**: valence
- **CONTENT**: semantic structure
- **CONTEXT**: pragmatic information
- **NONLOCAL**: non-local dependencies (extraction)
- **LOCAL**: (phrases) or
- **PHONOLOGY**: phonological structure
- **SYNSEM**: part of speech, etc.
- **ARG-ST**: argument structure (words)
We propose to assume an underspecified lexical entry for each preposition being able to contract with pronouns, i.e., bez ‘without’, dla ‘for’, do ‘to’, o ‘about’, od ‘from’, na ‘on’, nad ‘above’, po ‘for’, pod ‘under’, przed ‘in front of’, przez ‘owing’, w ‘in’, z ‘from’, of the following form:

\[
\begin{align*}
\text{word} & \\
\text{MORPH} & \left\langle \left[ \text{PHON} \langle \text{na} \rangle, \ldots \right] \rightangle \\
\text{ARG-ST} & \left\langle \left[ \text{LOC} \mid \text{CAT} \mid \text{HEAD} \quad \text{noun} \right] \rightangle \\
\text{SYNS} & \left\langle \left[ \text{LOC} \mid \text{CAT} \mid \text{HEAD} \quad \text{prep} \quad \text{PFORM} \langle \text{na} \rangle \right] \rightangle
\end{align*}
\]
The Lexical Constraint for Licensing PPC

\[
\begin{align*}
\text{word} & \quad \text{ARG-ST} \quad \text{LOC} \quad \text{CONT} \quad \text{CAT} \quad \text{CASE} \quad \text{gen} \vee \text{acc} \quad \text{ppro} \quad \text{NUMBER} \quad \text{sing} \quad \text{GENDER} \quad \text{masc} \quad \text{PERSON} \quad 3rd
\end{align*}
\]

\[
\begin{align*}
\text{prepos} & \quad \text{SYNS} \quad \text{LOC} \quad \text{CAT} \quad \text{HEAD} \quad \text{PFORM} \quad \text{dla} \vee \text{do} \vee \text{o} \vee \text{od} \vee \text{na} \vee \text{nad} \vee \text{po} \vee \text{pod} \vee \text{przed} \vee \text{przez} \vee \text{w} \vee \text{z}
\end{align*}
\]
The phonological form of PPCs is licensed by the general principle describing phonological form of words:

$$\text{word} \rightarrow \left[ \begin{array}{c} \text{PHON } F(1 \oplus \ldots \oplus n) \\ \text{MORPH} \left\langle \left[ \text{PHON } 1 \right], \ldots, \left[ \text{PHON } n \right] \right\rangle \end{array} \right]$$

According to this principle, the PHON value of words (and thus PPCs) is assumed to be the result of applying a function $F$ to the concatenation of the PHON values of the word’s all morphological components.

The definition of the $F$ function ensures the right phonological form of words according to phonological constraints for a given language. In the case of Polish PPCs, the output of the $F$ function is

- the phonological form of a given preposition combined with the phonological form of $-ń$, if that preposition ends in a vowel, and
- the phonological form of that preposition combined with a phonetic variant of $-ń$, i.e., $-eń$, if that preposition ends in a consonant.
The Lexical Constraint for Licensing Modifying and Non-Modifying PPs
The Lexical Constraint Describing Prepositions Selecting Arbitrary Arguments

\[ \begin{align*}
\text{word} & \quad \left[ \begin{array}{c}
\text{LOC} \left[ \begin{array}{c}
\text{CAT} | \text{HEAD} | \text{CASE} \quad \text{nom} \lor \text{dat} \lor \text{inst} \lor \text{loc}
\end{array} \right]
\end{array} \right] \\
\text{ARG-ST} & \quad \left[ \begin{array}{c}
\text{CONT} \quad \text{ppro}
\end{array} \right]
\end{align*} \]

\[ \quad \lor \left[ \begin{array}{c}
\text{LOC} | \text{CONT} | \text{npro} \lor \text{ana}
\end{array} \right] \]

\[ \text{SYNS} | \text{LOC} | \text{CAT} | \text{HEAD} \quad \text{prep} \]

\[ \text{word} \quad \left[ \begin{array}{c}
\text{ARG-ST} \quad 1
\end{array} \right]
\]

\[ \text{MORPH} \left\langle \text{morph} \right\rangle \]

\[ \text{SYNS} | \text{LOC} | \text{CAT} | \text{VAL} | \text{COMPS} \quad 1 \]
According to (Saloni 1981), as well as many grammars of Polish, Polish third person personal pronouns inflect besides

- case (nominative, genitive, dative, accusative, instrumental and locative),
- number (singular and plural) and
- accentability (yes or not)

for postprepositionality (yes or not).

Thereby the grammatical category of postprepositionality relates to structural circumstances rather than to prosodic:

(14) a. Źle się to skończyło dla niej i niego / *jego.
    bad RM it ended for her and himpp / him non-pp
    ‘It came to a bad end for her and him.’

b. Ta wiadomość jest zła zarówno dla nas jak i niego / *jego
    this message is bad both for us like and himpp / him non-pp
    samego.
    self
    ‘This message is bad both for us and for him self.’
A Simplified Description of Postprepositional Third Person Pronouns

\[
\begin{array}{c}
\text{word} \\
\text{ARG-ST}\langle\rangle \\
\text{SYNS} | \text{LOC} \\
\text{CAT} | \text{HEAD} \left[ \text{noun} \right] \\
\text{CONT} \left[ \text{ppro} \right. \\
\text{INDEX} | \text{PERSON 3rd} \\
\end{array}
\]
Constraints Ensuring the Right PP Value of Objects Selected by Ps and Non-Ps

\[
\begin{align*}
\forall [\text{word}] & \quad \text{SYNS} | \text{LOC} | \text{CAT} | \text{HEAD prep} \\
\rightarrow & \quad \text{ARG-ST} \left< [\text{LOC} | \text{CAT} | \text{PP +}] \right>
\end{align*}
\]

\[
\begin{align*}
\forall [\text{word}] & \quad \text{ARG-ST} \left< [\text{LOC} | \text{CAT} | \text{PP +}] \right> \\
\land \text{MEMBER}(2, 1) & \quad \rightarrow \\
\text{SYNS} | \text{LOC} | \text{CAT} | \text{HEAD prep} \\
2[\text{LOC} | \text{CAT} | \text{PP +}]
\end{align*}
\]

Preposition-Pronoun Contraction in Polish – p.28
Summary and Outlook

We have discussed the prosodic, categorial, syntactic and semantic properties of preposition-pronoun contraction in Polish.

On the basis of empirical observations, the generalization has been made that PPCs are morphological units with prepositional status.

A lexicalist HPSG-based approach has been provided which accounts for both PPCs and full PPs in Polish.

In future work,

- the developed HPSG analysis will be implemented in TRALE;
- a corpus-based study on third person pronouns will be carried out in order to determine the number of pronominal forms exactly which can occur within PPCs in Polish and to verify the number of Ps being able to contract with pronouns.
- It might also be interesting to examine whether the analysis proposed here for Polish data can be applied to corresponding data in other languages, e.g., to da- / wo- / hier- expressions in German.
References


