

**STRUCTURAL CASE ASSIGNMENT TO OBJECTS
IN POLISH**

Agnieszka Patejuk and Adam Przepiórkowski
Institute of Computer Science, Polish Academy of Sciences

Proceedings of the LFG14 Conference

Miriam Butt and Tracy Holloway King (Editors)

2014

CSLI Publications

<http://csli-publications.stanford.edu/>

Abstract

The aim of this paper is to present the details of an analysis of structural case assignment to Polish objects. While the analysis is implemented in the XLE system, its technical details are explained here using the easier vocabulary of theoretical LFG. A comparison to two earlier accounts of the same phenomenon demonstrates that the current analysis is conceptually and representationally simpler than either.

1 Introduction

This paper presents an analysis of structural case assignment¹ to objects implemented in an XLE (Crouch et al. 2011) grammar of Polish, *POLFIE* (Patejuk and Przepiórkowski 2012b). While the basic generalisations are well-known (Przepiórkowski 1999, 2000), there are various ways of encoding them, and the account presented below is leaner than our previous attempts (Przepiórkowski 1999, Patejuk and Przepiórkowski 2011).

The current analysis of the basic facts is presented in §2, while §3 extends the analysis so as to account for long-distance Genitive of Negation. The following sections discuss extensions aimed at covering various other phenomena, including multiple objects requiring structural case marking (§4), so-called *feature indeterminacy* (§5) and structural case assignment to selected conjuncts under unlike category coordination (§6). Then, §7 compares this analysis to our previous accounts. Finally, §8 concludes the paper.

2 Basics

In Polish, as in various other languages, the case of some arguments is stable and does not depend on the syntactic environment or the form of the verb, while the case of other arguments is not stable in this sense. This difference, widely discussed in generative linguistics since the early 1980s (Rouveret and Vergnaud 1980, Vergnaud 1982), also in the context of Slavic (Babby 1980a,b), is often referred to as *lexical* (or *inherent*) versus *structural* case assignment (cf. Przepiórkowski 1999 and numerous references therein).

A typical illustration of lexically assigned case is the dative case of an argument of *POMAGAĆ* – the argument is marked for the dative regardless of the presence of

[†]Many thanks to the two anonymous reviewers for their comments, which led to various improvements of this paper. The work reported here was partially financed by the projects NEKST (<http://nekst.ipipan.waw.pl/>) and CLARIN-PL (<http://www.clarin-pl.eu/en/>).

¹Note that, while we use the usual term “case assignment”, the analysis is formalised in terms of case checking, i.e., using constraining equations such as $(\uparrow \text{OBJ CASE}) =_c \text{ACC}$, and not defining equations such as $(\uparrow \text{OBJ CASE}) = \text{ACC}$. However, it could probably be formalised equivalently using defining constraints (with appropriate disjunctions).

negation (no negation in (1), negation present in (2)) and regardless of the form of the head (gerund in (3)):²

- (1) Oni pomagają ludziom.
they help people.DAT
'They help people.' (NKJP)
- (2) Nie pomagają ludziom będącym w potrzebie.
NEG help people.DAT being in need
'They don't help people in need.' (NKJP)
- (3) Chodzi o pomaganie ludziom.
walks about helping people.DAT
'It is about helping people.' (NKJP)

By contrast, the value of case assigned to Polish objects requiring structural case depends on factors such as the presence of negation and the form of the head which assigns case:

- (4) Poczytam książkę.
read.1.SG book.ACC
'I'll read a book.' (NKJP)
- (5) Nie poczytają książki czy gazety.
NEG read.3.PL book.GEN OR newspaper.GEN
'They won't read a book or a newspaper.' (NKJP)
- (6) Proponuję też poczytanie książki.
suggest.1.SG also reading book.GEN
'I also suggest reading a book.' (NKJP)

Taking into consideration only uncontroversially verbal forms (gerunds uniformly assign genitive case to their structurally-cased objects, see (6)), here is the first approximation of the basic rules of structural case assignment to objects in Polish:

- If the verb is not negated, as in (4), whose f-structure is provided in (7), its structurally-cased object must bear the accusative case (here: *książkę*).
- If the verb is negated, as in (5), whose f-structure is presented in (8), i.e., contains the attribute NEG whose value is +, its structurally-cased object bears the genitive case (both conjuncts in *książki czy gazety* are in the genitive); this phenomenon is known in Slavic linguistics as genitive of negation (GoN).³

²Almost all examples in this paper are attested, mostly extracted from the National Corpus of Polish (NKJP; <http://nkjp.pl/>; Przepiórkowski et al. 2010, 2012).

³See Neidle 1982, 1988 for early LFG work on Russian case, but note that Russian GoN facts differ from Polish considerably.

$$(7) \left[\begin{array}{l} \text{PRED} \quad \text{'READ}(\underline{1},\underline{2})\text{' } \\ \text{SUBJ} \quad \underline{1} \left[\begin{array}{l} \text{PRED} \quad \text{'PRO'} \end{array} \right] \\ \text{OBJ} \quad \underline{2} \left[\begin{array}{l} \text{PRED} \quad \text{'BOOK'} \\ \text{CASE} \quad \text{ACC} \end{array} \right] \end{array} \right]$$

$$(8) \left[\begin{array}{l} \text{PRED} \quad \text{'READ}(\underline{1},\underline{2})\text{' } \\ \text{SUBJ} \quad \underline{1} \left[\begin{array}{l} \text{PRED} \quad \text{'PRO'} \end{array} \right] \\ \text{OBJ} \quad \underline{2} \left\{ \left[\begin{array}{l} \text{PRED} \quad \text{'BOOK'} \\ \text{CASE} \quad \text{GEN} \end{array} \right], \left[\begin{array}{l} \text{PRED} \quad \text{'NEWSPAPER'} \\ \text{CASE} \quad \text{GEN} \end{array} \right] \right\} \\ \text{NEG} \quad + \end{array} \right]$$

Such structural case assignment generalisations could be formalised using the constraint in (9), which should be placed in lexical entries of verbs with structurally-cased objects. This is a disjunctive constraint, with disjuncts *AFFIRMATIVE* and *NEGATIVE* defined in (10) and (11), respectively.

$$(9) \text{STRCASE} \equiv [\text{AFFIRMATIVE} \vee \text{NEGATIVE}]$$

$$(10) \text{AFFIRMATIVE} \equiv [\neg(\uparrow \text{NEG}) \wedge (\uparrow \text{OBJ CASE}) =_c \text{ACC}]$$

$$(11) \text{NEGATIVE} \equiv [(\uparrow \text{NEG}) =_c + \wedge (\uparrow \text{OBJ CASE}) =_c \text{GEN}]$$

The negated existential constraint ($\neg(\uparrow \text{NEG})$) in (10) ensures that there is no negation in the f-structure of the head, while the second conjunct in (10) requires the object to be marked for accusative case; this corresponds to examples such as (4) with f-structure in (7). By contrast, (11) checks that the verb is negated ($(\uparrow \text{NEG}) =_c +$) and checks that its object bears genitive case; these requirements are satisfied by sentences such as (5), whose f-structure is given in (8).

3 GoN in Verb Chains

The previous section dealt with the simplest cases of genitive of negation; the following examples illustrate some additional complexity:

$$(12) \text{Nie chcesz} \quad \text{poczytać Kodeksu.}$$

NEG want.2.SG read.INF Code.GEN

‘You don’t want to read the Code.’ (NKJP)

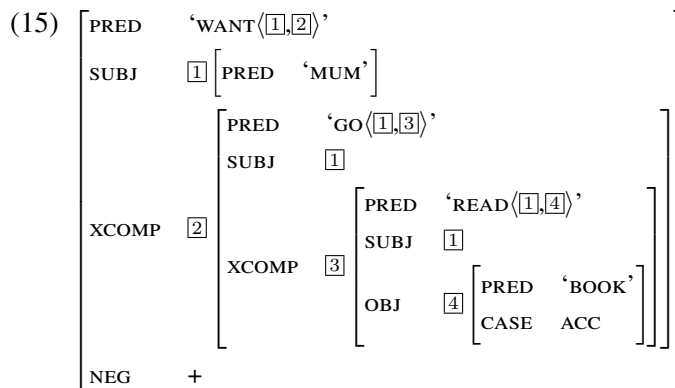
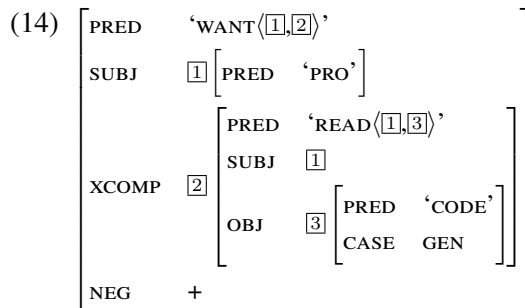
$$(13) \text{Mama nie chce} \quad \text{iść} \quad \text{poczytać książkę.}$$

mum NEG want.3.SG go.INF read.INF book.ACC

‘Mum doesn’t want to go and read a book.’ (NKJP)

These examples show that the generalisations discussed in the previous section, formalised in (9)–(11), need some refinement in the context of verb chains, i.e., constructions with control or raising verbs combining with (open) infinitival arguments. In particular, the case assignment constraints defined in (9)–(11) are too simple because – since negation is not local to the predicate subcategorising for the object in (12) – they would reject the genitive object (*Kodeksu*) there. As discussed at length in Przepiórkowski 2000, it is not the case, however, that genitive is always obligatory in such environments: in (13), the object is marked for accusative case (*książkę* ‘book’), even though there is verbal negation present higher in the structure of the sentence (at the main verb *chce* ‘wants’). Rather, the generalisation seems to be that local negation on the verb results in the obligatory genitive on structurally-cased objects, while higher negation in such verb chains leaves some optionality as to whether the lower object should occur in the accusative or in the genitive, with semantic and — structural or linear – distance factors perhaps weighing for one or the other.

The f-structures provided below show where exactly negation is present in the verb chain (it is marked as the NEG attribute with value ‘+’): (14) corresponds to (12), (15) provides a representation of (13).

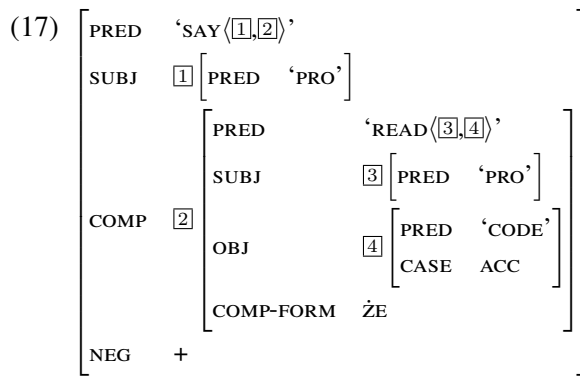


While example (12) shows that verbal negation may work at a distance in Polish, this long-distance impact is not unconstrained:

- (16) (Wcale) nie powiedziałeś, że poczytasz Kodeks/*Kodeksu.
 not at all NEG said.2.SG that read.2.SG code.ACC/GEN

‘You did not say that you will read the Code.’

The above example shows that verbal negation may not be “transferred” to lower sentential (closed) clauses. In LFG there is a convenient way of distinguishing between the two embedded clause types, infinitival (open) and sentential (closed): the former correspond to the xCOMP grammatical function, while the latter are represented in f-structure as COMP; this is the main difference between (17), the f-structure representation of the grammatical version of (16), where the object must be marked for accusative case, and (14), which corresponds to (12), where genitive is possible.



These observations are formalised via constraints in (18)–(24):

- (18) $\text{STRCASE} \equiv [\text{AFFIRMATIVE} \vee \text{NEGATIVE}]$
- (19) $\text{AFFIRMATIVE} \equiv [\neg((\text{XCOMP}^* \uparrow) \text{NEG}) \wedge (\uparrow \text{OBJ CASE}) =_c \text{ACC}]$
- (20) $\text{NEGATIVE} \equiv [\text{ANYNEG} \wedge \text{NEGTYPE}]$
- (21) $\text{ANYNEG} \equiv ((\text{XCOMP}^* \uparrow) \text{NEG}) =_c +$
- (22) $\text{NEGTYPE} \equiv [\text{LOCNEG} \vee \text{NONLOCNEG}]$
- (23) $\text{LOCNEG} \equiv [(\uparrow \text{NEG}) =_c + \wedge (\uparrow \text{OBJ CASE}) =_c \text{GEN}]$
- (24) $\text{NONLOCNEG} \equiv [\neg(\uparrow \text{NEG}) \wedge (\uparrow \text{OBJ CASE}) \in_c \{\text{ACC}, \text{GEN}\}]$

First of all, (18) is the top-level case assignment rule STRCASE , repeated from (9), to be placed in lexical entries of verbs with structurally-cased objects. Its disjuncts are, however, redefined as in (19) and (20) in order to take long-distance GoN into account, together with its optionality, as described above.

The first disjunct of (18), AFFIRMATIVE defined in (19), is a slightly modified version of (10): it handles the situation where the object must be marked for accusative case since verbal negation is not present in the relevant domain at all, neither locally to the verb directly subcategorising for the object nor higher in the verb chain. This is achieved by using an inside-out path which makes it possible to reach into any

number of successive predicates subcategorising for *xcomp*, an infinitival complement, and check if any of these predicates is negated. The statement in (19) handles sentences such as (4), whose f-structure is given in (7).

The second disjunct of (18), *NEGATIVE* defined in (20), handles situations where verbal negation is present. It is a conjunction of two statements: *ANYNEG* and *NEGTYPE*. The former, *ANYNEG*, ensures that verbal negation is present at some level, either locally or non-locally (higher within the verb chain). The latter, *NEGTYPE*, assigns case to the object depending on what type of negation is present (local or non-local to the predicate subcategorising for the structurally-cased object).

The statement *LOCNEG* defined in (23) is identical to (11) – it assigns genitive case to the object when the verb subcategorising for the object is negated, i.e., this is the obligatory genitive of negation clause. It is applied in examples such as (5), whose f-structure is given in (8) (negation is local to the predicate which subcategorises for the object).

The last statement, *NONLOCNEG* defined in (24), handles the case in which there is no local negation. However, the statement *ANYNEG* defined in (21) makes sure that there is negation at some level, so non-local negation must be present. In such environments the genitive of negation is optional – the object of the verb is assigned accusative (see sentence (13), and (15) for its f-structure) or genitive case (as in (12), with the corresponding f-structure in (14)).

3.1 Alternative Formalisation

Note that the constraint in (18) could be considerably simplified as follows:⁴

$$(25) \text{ STRCASE} \equiv [\text{AFFIRMATIVE} \vee \text{NEGATIVE}]$$

$$(26) \text{ AFFIRMATIVE} \equiv [\neg(\uparrow \text{NEG}) \wedge (\uparrow \text{OBJ CASE}) =_c \text{ACC}]$$

$$(27) \text{ NEGATIVE} \equiv [((\text{XCOMP}^* \uparrow) \text{NEG}) =_c + \wedge (\uparrow \text{OBJ CASE}) =_c \text{GEN}]$$

This constraint states that accusative case is possible whenever there is no local negation (this case is handled by (26)), while genitive case is possible whenever sentential negation is available somewhere in the verb chain, locally or non-locally (this is handled by (27)).

Though (25) is equivalent to (18) in that it allows for the same values of case in corresponding environments, the statement formalised in (18) might be better suited for use with OT (Optimality Theory) marks to, for instance, disprefer the use of accusative case when non-local negation is available. However, one of the templates called by (18) would have to be modified in order to achieve this: instead of using set-membership notation for accusative or genitive under non-local negation in (24) $((\uparrow \text{OBJ CASE}) \in_c \{\text{ACC}, \text{GEN}\})$, a disjunction of appropriate constraining equations would have to be used, preferring $(\uparrow \text{OBJ CASE}) =_c \text{GEN}$ or dispreferring $(\uparrow \text{OBJ CASE}) =_c \text{ACC}$.

⁴Thanks are due to the internal reviewer for pointing this out.

4 Multiple Objects Marked for Structural Case

There is also a subtle empirical difference concerning the (probably) only Polish verb which seems to have two non-subject arguments assigned case structurally, namely, *KOSZTOWAĆ* ‘cost’:

(28) *Codzienny transport kosztuje ją grosze.*
 daily.NOM transport.NOM costs she.ACC pennies.ACC
 ‘Daily transport costs her pennies.’ (Google)

(29) *Utrzymanie nie kosztuje jej majątku.*
 maintenance.NOM NEG costs she.GEN fortune.GEN
 ‘Maintenance does not cost her a fortune.’ (Google)

In (28), a form of the verb *KOSZTOWAĆ* combines with two accusative phrases: *ją* ‘her’ and *grosze* ‘pennies’. In (29), where the same verb is negated, both arguments occur in the genitive: *jej* ‘her’ and *majątku* ‘fortune’.

While the facts are a little more complex, as the genitive of negation on one or the other of these two arguments is perhaps not obligatory for some speakers,⁵ let us see how the analysis proposed in this paper can be modified so as to account for such cases: templates called by (18) defined in (19)–(24) assume that the argument bearing structural case is OBJ. However, these templates can be parameterised⁶ by replacing OBJ with a variable, call it GF, and placing this variable as a parameter of relevant templates, as shown below:

(30) $\text{STRCASE}(\text{GF}) \equiv [\text{AFFIRMATIVE}(\text{GF}) \vee \text{NEGATIVE}(\text{GF})]$

(31) $\text{AFFIRMATIVE}(\text{GF}) \equiv [\neg((\text{XCOMP}^* \uparrow) \text{NEG}) \wedge (\uparrow \text{GF CASE}) =_c \text{ACC}]$

(32) $\text{NEGATIVE}(\text{GF}) \equiv [\text{ANYNEG} \wedge \text{NEGTYPE}(\text{GF})]$

(33) $\text{ANYNEG} \equiv ((\text{XCOMP}^* \uparrow) \text{NEG}) =_c +$

(34) $\text{NEGTYPE}(\text{GF}) \equiv [\text{LOCNEG}(\text{GF}) \vee \text{NONLOCNEG}(\text{GF})]$

(35) $\text{LOCNEG}(\text{GF}) \equiv [(\uparrow \text{NEG}) =_c + \wedge (\uparrow \text{GF CASE}) =_c \text{GEN}]$

(36) $\text{NONLOCNEG}(\text{GF}) \equiv [\neg(\uparrow \text{NEG}) \wedge (\uparrow \text{GF CASE}) \in_c \{\text{ACC}, \text{GEN}\}]$

Once this change is introduced, the template *STRCASE* defined in (30) must be called with a parameter whose value corresponds to the grammatical function of the argument which takes structural case. Such a change makes it possible to make any number of calls to this template in the lexical entry of a verb – one for each argument assigned case structurally. This way, verbs such as *POCZYTAĆ* ‘read’ which take one object marked for structural case (see (4)–(5) and (12)–(13) for examples) make one call to (30) with an appropriate grammatical function as GF, as in (37).

⁵See Przeciorkowski 1999: 105, but more empirical research is needed to establish the facts.

⁶In fact, the currently implemented analysis contains exactly such parameterised templates.

(37) STRCASE(OBJ)

In order to account for case assignment with verbs such as *KOSZTOWAĆ* ‘cost’ (see (28)–(29) for examples), where two objects are marked for structural case, the lexical entry of the verb contains two calls to (30) with two different object grammatical functions, for instance OBJ and OBJ2, as in (38):

(38) STRCASE(OBJ) STRCASE(OBJ2)

When these template calls are made, they independently ensure that relevant objects bear values of structural case appropriate in the given syntactic context. The f-structure in (39) results for *Transport kosztuje ją grosze*, a slightly simplified version of (28), while (40) provides a representation of (29):

| | | | | | | | | | | | | | | | | | | | | | |
|------|---|------|---------------|------|------------------------|-----|------------------------------|------|--------------------------------|------|---|------|---------------|------|--------------------------|-----|------------------------------|------|----------------------------------|-----|---|
| (39) | <table style="border: none; width: 100%;"> <tr><td style="padding: 2px 10px;">PRED</td><td style="padding: 2px 10px;">‘COST(1,2,3)’</td></tr> <tr><td style="padding: 2px 10px;">SUBJ</td><td style="padding: 2px 10px;">1 [PRED ‘TRANSPORT’]</td></tr> <tr><td style="padding: 2px 10px;">OBJ</td><td style="padding: 2px 10px;">2 [PRED ‘SHE’ CASE ACC]</td></tr> <tr><td style="padding: 2px 10px;">OBJ2</td><td style="padding: 2px 10px;">3 [PRED ‘PENNY’ CASE ACC]</td></tr> </table> | PRED | ‘COST(1,2,3)’ | SUBJ | 1 [PRED ‘TRANSPORT’] | OBJ | 2 [PRED ‘SHE’ CASE ACC] | OBJ2 | 3 [PRED ‘PENNY’ CASE ACC] | (40) | <table style="border: none; width: 100%;"> <tr><td style="padding: 2px 10px;">PRED</td><td style="padding: 2px 10px;">‘COST(1,2,3)’</td></tr> <tr><td style="padding: 2px 10px;">SUBJ</td><td style="padding: 2px 10px;">1 [PRED ‘MAINTENANCE’]</td></tr> <tr><td style="padding: 2px 10px;">OBJ</td><td style="padding: 2px 10px;">2 [PRED ‘SHE’ CASE GEN]</td></tr> <tr><td style="padding: 2px 10px;">OBJ2</td><td style="padding: 2px 10px;">3 [PRED ‘FORTUNE’ CASE GEN]</td></tr> <tr><td style="padding: 2px 10px;">NEG</td><td style="padding: 2px 10px;">+</td></tr> </table> | PRED | ‘COST(1,2,3)’ | SUBJ | 1 [PRED ‘MAINTENANCE’] | OBJ | 2 [PRED ‘SHE’ CASE GEN] | OBJ2 | 3 [PRED ‘FORTUNE’ CASE GEN] | NEG | + |
| PRED | ‘COST(1,2,3)’ | | | | | | | | | | | | | | | | | | | | |
| SUBJ | 1 [PRED ‘TRANSPORT’] | | | | | | | | | | | | | | | | | | | | |
| OBJ | 2 [PRED ‘SHE’ CASE ACC] | | | | | | | | | | | | | | | | | | | | |
| OBJ2 | 3 [PRED ‘PENNY’ CASE ACC] | | | | | | | | | | | | | | | | | | | | |
| PRED | ‘COST(1,2,3)’ | | | | | | | | | | | | | | | | | | | | |
| SUBJ | 1 [PRED ‘MAINTENANCE’] | | | | | | | | | | | | | | | | | | | | |
| OBJ | 2 [PRED ‘SHE’ CASE GEN] | | | | | | | | | | | | | | | | | | | | |
| OBJ2 | 3 [PRED ‘FORTUNE’ CASE GEN] | | | | | | | | | | | | | | | | | | | | |
| NEG | + | | | | | | | | | | | | | | | | | | | | |

5 Feature Indeterminacy

We have assumed so far that the value of the CASE attribute is atomic. It has been demonstrated, however, that such an assumption is problematic for the proper treatment of case syncretism, where two morphological cases expressed by a given form are both active in some construction. Within LFG, such feature indeterminacy is discussed in Dalrymple and Kaplan 2000 and in Dalrymple et al. 2009 – see these papers for references to previous work on this issue.

Consider the following constructed examples, which illustrate the problem:

- (41) *Marysia lubi ale też boi się Marka.*
 Marysia.NOM likes but also be afraid REFL Marek.ACC/GEN
 ‘Marysia likes as well as fears Marek.’
- (42) *Marysia lubi ale nie szanuje Marka.*
 Marysia.NOM likes but NEG respects Marek.ACC/GEN
 ‘Marysia likes but does not respect Marek.’

The first verb in (41), *lubi* ‘likes’, takes an object marked for structural case while the second verb, *boi się* ‘fears, is afraid’, takes an object marked for the genitive case lexically. Example (42) is similar, with the only difference being that the genitive is assigned structurally, by the negated verb *nie szanuje* ‘does not respect’.

The solution to the problem of different case requirements on a shared dependent of coordinated verbs which is proposed in Dalrymple et al. 2009 is to use a complex *CASE* attribute which has as many subattributes as there are cases in a given language. Each of these subattributes can take one of the two values: positive (when the relevant case value is realised in a given environment) or negative (when the relevant case value is not possible). This way, a given word can be marked for more than one case at the same time, as shown in the simplified f-structure for the word *Marka* provided in (45). According to this f-structure, the noun is marked for accusative and genitive case at the same time, as in examples (41)–(42). This should be contrasted with the situation when the value of case is atomic – see the two separate f-structures for *Marka* in (43) and (44), where the noun is marked for accusative or genitive case exclusively, respectively.

$$(43) \begin{bmatrix} \text{PRED} & \text{'MAREK'} \\ \text{CASE} & \text{ACC} \end{bmatrix} \quad (44) \begin{bmatrix} \text{PRED} & \text{'MAREK'} \\ \text{CASE} & \text{GEN} \end{bmatrix} \quad (45) \begin{bmatrix} \text{PRED} & \text{'MAREK'} \\ \text{CASE} & \begin{bmatrix} \text{ACC} & + \\ \text{DAT} & - \\ \text{GEN} & + \\ \text{INST} & - \\ \text{LOC} & - \\ \text{NOM} & - \\ \text{VOC} & - \end{bmatrix} \end{bmatrix}$$

It is easy to extend the analysis of this paper so that it takes into account the approach to case of Dalrymple et al. 2009; the templates of §4 and their modified versions are given below:

$$(46) \text{ a. } \text{AFFIRMATIVE}(\text{GF}) \equiv [\neg((\text{XCOMP}^* \uparrow) \text{NEG}) \wedge (\uparrow \text{GF CASE}) =_c \text{ACC}] \quad (= (31))$$

$$\text{ b. } \text{AFFIRMATIVE}(\text{GF}) \equiv [\neg((\text{XCOMP}^* \uparrow) \text{NEG}) \wedge (\uparrow \text{GF CASE ACC}) =_c +]$$

$$(47) \text{ a. } \text{LOCNEG}(\text{GF}) \equiv [(\uparrow \text{NEG}) =_c + \wedge (\uparrow \text{GF CASE}) =_c \text{GEN}] \quad (= (35))$$

$$\text{ b. } \text{LOCNEG}(\text{GF}) \equiv [(\uparrow \text{NEG}) =_c + \wedge (\uparrow \text{GF CASE GEN}) =_c +]$$

$$(48) \text{ a. } \text{NONLOCNEG}(\text{GF}) \equiv [\neg(\uparrow \text{NEG}) \wedge (\uparrow \text{GF CASE}) \in_c \{\text{ACC}, \text{GEN}\}] \quad (= (36))$$

$$\text{ b. } \text{NONLOCNEG}(\text{GF}) \equiv [\neg(\uparrow \text{NEG}) \wedge (\uparrow \text{GF CASE } \{\text{ACC}|\text{GEN}\}) =_c +]$$

(46b) and (47b), complex *CASE* counterparts of (31) and (35), respectively, are modified so that the value of relevant subattributes of *CASE*, *ACC* and *GEN*, is required to be positive (+), meaning that the object is marked for the relevant case. The modification introduced in (48b) with respect to (36), the version using a *CASE* attribute with atomic values, is slightly more complex. In (36) a set is used in the constraint to specify the possible values of *CASE*: $(\uparrow \text{GF CASE}) \in_c \{\text{ACC}, \text{GEN}\}$. This accounts for the optionality of *GoN* in this environment: the value of *CASE* must belong to the set containing *ACC* and *GEN* values. Once complex *CASE* attribute is adopted, it no longer makes sense to use set-based constraints: the subattributes can only take

boolean values (+ or –). Functional uncertainty is used instead in the path, i.e., $(\uparrow \text{GF CASE } \{\text{ACC|GEN}\}) =_c +$, which is equivalent to the disjunction: $[(\uparrow \text{GF CASE ACC}) =_c + \vee (\uparrow \text{GF CASE GEN}) =_c +]$. The effect of such a constraint is that the object can be marked for accusative case, genitive case or both – this last possibility accounts for feature indeterminacy in (41)–(42).

6 Unlike Category Coordination

Consider the following examples from Polish:

- (49) Doradził mu wyjazd i żeby nie wracał.
 advised him departure.ACC and that NEG come back
 ‘He advised him to leave and not to come back.’ (Kallas 1993: p. 92, ex. (48a))
- (50) (Wcale) nie doradził mu wyjazdu ani żeby nie wracał.
 not at all NEG advised him departure.GEN nor that NEG come back
 ‘He did not advise him to leave nor not to come back.’

In examples (49)–(50) the object is a coordinate phrase which consists of a nominal (a form of the noun WYJAZD ‘departure’) and a clause (*żeby nie wracał* ‘not to come back’). However, these examples differ consistently depending on whether negation (the negation marker NIE) is present – in (49) the nominal conjunct takes accusative case, while in (50) the object is marked for genitive case. Such variation is attributable to the fact that the verb DORADZIĆ ‘advise’ may take a structurally case-marked object.

To allow for clausal conjuncts in (49) and (50), a constraint checking that the object contains the appropriate complementiser (ŻEBY) must be added to the constraint handling structural case assignment. The template defined in (51) can be used for this purpose: it takes GF as the parameter and checks that the value of its COMP-FORM attribute is equal to ŻEBY:

$$(51) \text{ COMP-ŻEBY(GF)} \equiv (\uparrow \text{GF COMP-FORM}) =_c \text{ ŻEBY}$$

Intuitively, a disjunction of template calls to STRCASE defined in (30) and COMP-ŻEBY defined in (51) should account for unlike category coordination in (49) and (50):

$$(52) [\text{STRCASE(OBJ)} \vee \text{COMP-ŻEBY(OBJ)}]$$

This is not the case, however, because the disjunction is understood too early and, rather than being resolved for each element of the relevant f-structure (here: the object) individually, it is resolved once and applied to all its elements.⁷ As a result, (52) handles correctly simple cases (no coordination, coordination of elements bearing the same specification), but it cannot account for unlike category coordination witnessed in (49) and (50). In order to account for these examples, (52) must be rewritten so as to use off-path constraints, as in (53).

⁷See Patejuk and Przepiórkowski 2012a, 2014 for a detailed explanation of this problem.

$$(53) \left(\uparrow_{\text{OBJ}} \quad \text{PRED} \right) \\
\left[\neg((\text{XCOMP}^* \text{OBJ} \leftarrow) \text{NEG}) \wedge (\leftarrow \text{CASE}) =_c \text{ACC} \right] \vee \\
\left[((\text{XCOMP}^* \text{OBJ} \leftarrow) \text{NEG}) =_c + \wedge \right. \\
\left. \left[((\text{OBJ} \leftarrow) \text{NEG}) =_c + \wedge (\leftarrow \text{CASE}) =_c \text{GEN} \right] \vee \right. \\
\left. \left[\neg((\text{OBJ} \leftarrow) \text{NEG}) \wedge (\leftarrow \text{CASE}) \in_c \{ \text{ACC}, \text{GEN} \} \right] \right] \vee \\
(\leftarrow \text{COMP-FORM}) =_c \text{ŻEBY}$$

The constraint provided in (53) is an exact off-path equivalent of (the expanded version of) (52) – it assigns appropriate structural case or allows clauses containing a ŻEBY-type complementiser. If there is no negation at all, it assigns accusative case. If the verb assigning case is negated locally, genitive case is assigned. If there is no local negation but transferred negation is available, GoN is optional and the object may be assigned accusative or genitive case. Finally, clausal realisations of the object are also allowed (ŻEBY-type complementiser).

The f-structure in (54) below corresponds to (50):

$$(54) \left[\begin{array}{l} \text{PRED} \quad \text{'ADVISE'} \langle \boxed{1}, \boxed{2}, \boxed{3} \rangle \\ \text{SUBJ} \quad \boxed{1} \left[\text{PRED} \quad \text{'PRO'} \right] \\ \text{OBJ} \quad \boxed{2} \left[\begin{array}{l} \left[\begin{array}{l} \text{PRED} \quad \text{'DEPARTURE'} \\ \text{CASE} \quad \text{GEN} \end{array} \right], \\ \left[\begin{array}{l} \text{PRED} \quad \text{'COME_BACK'} \langle \boxed{4} \rangle \\ \text{SUBJ} \quad \boxed{4} \left[\text{PRED} \quad \text{'PRO'} \right] \\ \text{COMP-FORM} \quad \text{ŻEBY} \\ \text{NEG} \quad + \end{array} \right] \end{array} \right] \\ \left[\text{COORD-FORM} \quad \text{ANI} \right] \\ \text{OBJ}_\theta \quad \boxed{3} \left[\begin{array}{l} \text{PRED} \quad \text{'HE'} \\ \text{CASE} \quad \text{DAT} \end{array} \right] \\ \text{NEG} \quad + \end{array} \right]$$

6.1 Alternative Formalisation

Note that the fragment of (53) related to structural case assignment could be simplified as described in §3.1; the following constraint would result:

$$(55) \left(\uparrow_{\text{OBJ}} \quad \text{PRED} \right) \\
\left[\neg((\text{OBJ} \leftarrow) \text{NEG}) \wedge (\leftarrow \text{CASE}) =_c \text{ACC} \right] \vee \\
\left[((\text{XCOMP}^* \text{OBJ} \leftarrow) \text{NEG}) =_c + \wedge (\leftarrow \text{CASE}) =_c \text{GEN} \right] \vee \\
(\leftarrow \text{COMP-FORM}) =_c \text{ŻEBY}$$

7 Comparisons

7.1 Earlier LFG Analysis: Patejuk and Przepiórkowski 2011

The first XLE/LFG implementation of Polish structural case assignment is described in Patejuk and Przepiórkowski 2011; since this paper is only available in Polish, we

will briefly present here the part concerning structural case assignment to objects, ignoring the part concerned with subjects.

The c-structure of *POLFIE* (Patejuk and Przepiórkowski 2012b) was originally based on GFJP,⁸ a DCG⁹ grammar of Polish presented in Świdziński 1992; its implementation used by the *Świgr*a parser is described in Woliński 2004. Hence, the names of nonterminals in this early analysis follow those of *Świgr*a: ZDANIE means ‘sentence’, FF stands for *fraza finitywna* ‘finite phrase’, and FW – for *fraza wymagana* ‘required phrase’ (i.e., an argument). A considerably simplified sentential rule is given below:^{10,11}

```
ZDANIE --> (FW: ( ^ SUBJ)=! ),
           FF:  ^=!
           @(STRCS)
           @(STRCO) ,
           (FW: ( ^ OBJ)=! ) ,
           (FW: ( ^ OBJ_TH)=! ) ,
           (FW: ( ^ XCOMP)=!
           (! CAT)=c inf
           @(NEGTR)).
```

This rule states that a sentence (ZDANIE) consists of – in any order – an obligatory finite phrase (FF) and four optional arguments (FW): the subject, an object, a thematic object and an infinitival open complement. There are three template calls: STRCS for case assignment to the potentially structural subject (ignored here), STRCO for case assignment to the potentially structural object (explained in detail below), and NEGTR for the “transmission of negation” within the infinitival complement.

The idea of “transmission of negation” was proposed in Przepiórkowski and Świdziński 1997 in the context of GFJP, the aforementioned DCG grammar. In Patejuk and Przepiórkowski 2011, the template NEGTR is defined as follows:

```
NEGTR =
{
~( ^ NEG)
|
( ^ NEG)
(! NEG TR)=+
  {
  ( ^ NEG LOC)=c +
  |
```

⁸*Gramatyka formalna języka polskiego* (Świdziński 1992).

⁹*Definite Clause Grammar* (Warren and Pereira 1980).

¹⁰The relevant fragments of the grammar are presented in the XLE notation: ^ corresponds to ↑ in theoretical LFG notation, ! to ↓; ~ corresponds to ¬; @ marks template calls.

¹¹F-descriptions taken from Patejuk and Przepiórkowski 2011 are slightly modified here so as to comply with conventions adopted in the current analysis (e.g. lowercase values of attributes).

```

~(^ NEG LOC)
(^ NEG TR)=c +
} }

```

According to this template, if the higher phrase contains the attribute `NEG`, with either local negation ($(\uparrow \text{NEG LOC}) =_c +$) or transmitted negation ($(\uparrow \text{NEG TR}) =_c +$) already established, then the current phrase must also have this attribute and, within its value, the attribute `TR` (for ‘transmitted’) must be set to ‘+’. Since this template is present on all infinitival complements, it ensures the transmission of negation downwards in verb chains.

The template responsible for case assignment to structurally-cased objects is given below.

```

STRCO =
{~(^ RQSC)
|   {(^ RQSC)=c -
|   |
|   |   (^ RQSC)=c +
|   |   |   {~(^ NEG)
|   |   |   |   {(^ OBJ CASE)= acc
|   |   |   |   |
|   |   |   |   |   ~(^ OBJ)
|   |   |   |   |   (^ OBJ_TH CASE)= acc
|   |   |   |   |   }
|   |   |   |   }
|   |   |   |   (^ NEG)
|   |   |   |   |   {(^ NEG LOC)=c +
|   |   |   |   |   |   {(^ OBJ CASE)= gen
|   |   |   |   |   |   |
|   |   |   |   |   |   |   ~(^ OBJ)
|   |   |   |   |   |   |   (^ OBJ_TH CASE)= gen
|   |   |   |   |   |   |   }
|   |   |   |   |   |   }
|   |   |   |   |   |   ~(^ NEG LOC)
|   |   |   |   |   |   (^ NEG TR)=c +
|   |   |   |   |   |   |   {   {(^ OBJ CASE)= gen
|   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   ~(^ OBJ)
|   |   |   |   |   |   |   |   |   (^ OBJ_TH CASE)= gen
|   |   |   |   |   |   |   |   |   }
|   |   |   |   |   |   |   |   |   }
|   |   |   |   |   |   |   |   |   {(^ OBJ CASE)= acc
|   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   ~(^ OBJ)
|   |   |   |   |   |   |   |   |   |   (^ OBJ_TH CASE)= acc
|   |   |   |   |   |   |   |   |   |   }
}   }   }   }   }   }

```

This template relies on an additional attribute, apart from `LOC` and `TR` within values of `NEG`, namely, on `RQSC`, which is appropriate for verbs and stands for ‘required structural case (on some object)’. This template does not do anything if `RQSC` is absent or if its value is ‘-’ (lexical (or inherent) case). The action starts if the verb is lexically marked as `RQSC +`; if so:

- either the verb has no `NEG` attribute, in which case `OBJ` or – if there is no `OBJ` – `OBJ_TH` is assigned the accusative,
- or `NEG` is present, in which case:
 - either there is local negation, in which case `OBJ` or – if there is no `OBJ` – `OBJ_TH` is assigned the genitive case,
 - or there is no local negation but there is transmitted negation, in which case:
 - * either `OBJ` or – if there is no `OBJ` – `OBJ_TH` is assigned the genitive case,
 - * or `OBJ` or – if there is no `OBJ` – `OBJ_TH` is assigned the accusative case.

One problem with this analysis is that the above templates use defining case assignment equations rather than constraining equations, which is problematic for non-canonical (e.g. clausal or infinitival) realisations of `OBJ` or `OBJ_TH` – an example where a nominal marked for structural case can be coordinated with a clause was provided in (49). Unless appropriate disjunctions are used in `STRCO` in order to avoid assigning case to categories where it is undesired, the clause in (49) would be assigned case.

Second, the analysis is unnecessarily complex – it is repetitive, it does not use embedded templates. One complication within the above templates stems from the fact that – in the 2011 version of *POLFIE* – both `OBJ` and `OBJ_TH` were allowed to bear structural case, however, `OBJ_TH` could only be assigned structural case if `OBJ` was not present at the same time. In the current analysis the template `STRCASE` in (30) may be called with any object grammatical function, as determined by the lexical entry of the relevant verb. Furthermore, the analysis of Patejuk and Przepiórkowski 2011 requires the f-structure presence of two attributes whose role is purely technical: `NEG TR` and `RQSC`. In contrast, the analysis proposed in §3 only makes use of the boolean-valued `NEG` attribute which simply indicates whether the verb is (locally) negated or not.

Finally, examples such as (28)–(29) could not be properly accounted for on the analysis of Patejuk and Przepiórkowski 2011. The requirement for structural case is only marked on the verb (as `RQSC` attribute), so the usual situation where only one of non-subject arguments is assigned case structurally cannot be distinguished from the situation illustrated in (28)–(29), where both such arguments receive structural case. Besides, `STRCO` template can assign case to only one object – as a result, the case marking of the second object could be wrong.

7.2 HPSG Analysis: Przepiórkowski 1999

An even earlier analysis of structural case assignment to Polish objects which takes into account the optionality of long-distance GoN is the HPSG (Pollard and Sag 1994) analysis presented in Przepiórkowski 1999.¹² This HPSG analysis is based on an intuition similar to – but developed independently from – that of Dziwirek 1994, 1998, couched in the Relational Grammar framework (Perlmutter 1983). The main idea of this analysis is that the genitive-of-negation object is raised to the argument structure of the negated verb. The optionality of this raising mechanism reflects the optionality of long-distance GoN. The advantage of this analysis is that case assignment remains an intimately local affair between a verb and its arguments – once arguments are raised, no long-distance case assignment mechanisms are required such as those in the current analysis (see the non-local functional uncertainty constraints in (19) and (21) above). The price to be paid is, however, the postulation of some *ad hoc* constraints on this ‘clause union’, e.g. the constraint that arguments cannot be raised above the highest negated verb (Przepiórkowski 1999: 155, (5.209)). Other phenomena give at best mixed evidence for this stance (Przepiórkowski 1999: §5.2).

What is clear, on the other hand, is that a general ‘clause union’ analysis for Polish verb chains cannot be right. Consider the following example:¹³

- (56) Na studiach uczono ją/*jej nie ufać takim badaniom.
on studies taught.IMPS her.ACC/GEN NEG trust.INF such.DAT investigations.DAT
‘During the studies, she was taught not to trust such research.’

In this example, negation on the lower verb does not have any effect on the structurally-cased object of the higher verb: the argument of UCZYĆ ‘teach’ must occur in the accusative, not genitive, as negation only occurs on the lower verb (as shown in (57) for *Uczono ją nie ufać badaniom*, an abbreviated version of (56)). If both predicates were ‘unioned’, negation would be expected to apply to any arguments of these predicates, contrary to fact.¹⁴

¹²See also the two analyses in Przepiórkowski and Świdziński 1997 which, however, assume that GoN is always obligatory.

¹³On the basis of a sentence from NKJP.

¹⁴A similar argument may be constructed on the basis of the behaviour of *n*-words in Polish, e.g., the ungrammaticality of the following example from Przepiórkowski and Kupść 1997: 30, where the lower negation in a verb cluster cannot license a higher *n*-word:

- (i) *Nikt chciał nie kupić tego domu.
nobody.NOM wanted NEG buy.INF this.GEN house.GEN

(57)

| | | | | | | | | | |
|-------|---------|---|------|---------|----|------|-----------------|---|--|
| PRED | ‘TEACH(| [| 1, | 2, | 3] | ’ | | | |
| SUBJ | [| 1 | PRED | ‘PRO’ | ’ | | | | |
| OBJ | [| 2 | PRED | ‘SHE’ | ’ | | | | |
| | | | CASE | ACC | ’ | | | | |
| XCOMP | [| 3 | PRED | ‘TRUST(| [| 2, | 4] | ’ | |
| | | | SUBJ | [| 2 | ’ | | | |
| | | | OBJ | [| 4 | PRED | ‘INVESTIGATION’ | ’ | |
| | | | NEG | + | ’ | | | | |

Hence, we conclude that the current analysis is simpler and empirically more adequate than that of Przepiórkowski 1999, even if it gives up the overwhelming generalisation that case assignment is a strictly local phenomenon.

8 Conclusion

This paper demonstrated how Polish structural case assignment to objects may be formalised in LFG using functional uncertainty. The case assignment statements formalised here take into account situations where case assigned to the object is influenced by verbal negation which is not local to the predicate which subcategorises for this structurally-cased object. Furthermore, the analysis is conceptually simpler, and leads to simpler representations, than either the HPSG analysis proposed in Przepiórkowski 1999 or our earlier XLE/LFG analysis presented in Patejuk and Przepiórkowski 2011. Finally, it was demonstrated that the solution offered in this paper can be modified so as to take other phenomena into consideration, including feature indeterminacy and case assignment under unlike category coordination.¹⁵

References

- Babby, Leonard [H.]. 1980a. *Existential Sentences and Negation in Russian*. Ann Arbor, MI: Karoma Publishers.
- Babby, Leonard [H.]. 1980b. The Syntax of Surface Case Marking. In Wayne Harbert and Julia Herschensohn (eds.), *Cornell Working Papers in Linguistics*, volume 1, pages 1–32, Department of Modern Languages and Linguistics, Cornell University.
- Crouch, Dick, Dalrymple, Mary, Kaplan, Ron, King, Tracy, Maxwell, John and Newman, Paula. 2011. XLE Documentation. http://www2.parc.com/isl/groups/nlitt/xle/doc/xle_toc.html.

¹⁵Patejuk and Przepiórkowski 2014 (in this volume) discusses how control into one of the conjuncts can be coupled with structural case assignment to selected conjuncts.

- Dalrymple, Mary and Kaplan, Ronald M. 2000. Feature Indeterminacy and Feature Resolution. *Language* 76(4), 759–798.
- Dalrymple, Mary, King, Tracy Holloway and Sadler, Louisa. 2009. Indeterminacy by underspecification. *Journal of Linguistics* 45, 31–68.
- Dziwirek, Katarzyna. 1994. *Polish Subjects*. New York: Garland.
- Dziwirek, Katarzyna. 1998. Reduced Construction in Universal Grammar: Evidence from the Polish Object Control Construction. *Natural Language and Linguistic Theory* 16, 53–99.
- Kallas, Krystyna. 1993. *Składnia współczesnych polskich konstrukcji współrzędnych*. Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika.
- Neidle, Carol. 1982. Case Agreement in Russian. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*, MIT Press Series on Cognitive Theory and Mental Representation, pages 391–426, Cambridge, MA: The MIT Press.
- Neidle, Carol. 1988. *The Role of Case in Russian Syntax*. Dordrecht: Kluwer.
- Patejuk, Agnieszka and Przepiórkowski, Adam. 2011. Przypadki strukturalne w polskiej gramatyce LFG. In Mirosław Bańko and Dorota Kopcińska (eds.), *Różne formy, różne treści: Tom ofiarowany Profesorowi Markowi Świdzińskiemu*, pages 183–193, Warsaw: Uniwersytet Warszawski, Wydział Polonistyki.
- Patejuk, Agnieszka and Przepiórkowski, Adam. 2012a. A Comprehensive Analysis of Constituent Coordination for Grammar Engineering. In *Proceedings of the 24rd International Conference on Computational Linguistics (COLING 2012)*, pages 2191–2207.
- Patejuk, Agnieszka and Przepiórkowski, Adam. 2012b. Towards an LFG parser for Polish: An exercise in parasitic grammar development. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC 2012*, pages 3849–3852, ELRA, Istanbul, Turkey.
- Patejuk, Agnieszka and Przepiórkowski, Adam. 2014. Control into Selected Conjunctions. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG'14 Conference*, Stanford, CA: CSLI Publications, under review.
- Perlmutter, David (ed.). 1983. *Studies in Relational Grammar*, volume 1. University of Chicago Press.
- Pollard, Carl and Sag, Ivan A. 1994. *Head-driven Phrase Structure Grammar*. Chicago, IL: Chicago University Press / CSLI Publications.

- Przepiórkowski, Adam. 1999. *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Ph.D. dissertation, Universität Tübingen, Germany.
- Przepiórkowski, Adam. 2000. Long Distance Genitive of Negation in Polish. *Journal of Slavic Linguistics* 8, 151–189.
- Przepiórkowski, Adam, Bańko, Mirosław, Górski, Rafał L. and Lewandowska-Tomaszczyk, Barbara (eds.). 2012. *Narodowy Korpus Języka Polskiego [Eng.: National Corpus of Polish]*. Warsaw: Wydawnictwo Naukowe PWN.
- Przepiórkowski, Adam, Górski, Rafał L., Łaziński, Marek and Pęzik, Piotr. 2010. Recent Developments in the National Corpus of Polish. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation, LREC 2010*, ELRA, Valletta, Malta.
- Przepiórkowski, Adam and Kupść, Anna. 1997. Negative Concord in Polish. IPI PAN Research Report 828, Institute of Computer Science, Polish Academy of Sciences, Warsaw.
- Przepiórkowski, Adam and Świdziński, Marek. 1997. Polish Verbal Negation Revisited: A Metamorphosis vs. HPSG Account. Technical Report 829, Institute of Computer Science, Polish Academy of Sciences.
- Rouveret, Alain and Vergnaud, Jean-Roger. 1980. Specifying Reference to the Subject: French causatives and conditions on representations. *Linguistic Inquiry* 11(1), 97–202.
- Świdziński, Marek. 1992. *Gramatyka formalna języka polskiego*, volume 349 of *Rozprawy Uniwersytetu Warszawskiego*. Warsaw: Wydawnictwa Uniwersytetu Warszawskiego.
- Vergnaud, Jean-Roger. 1982. *Dépendances et Niveaux de Representation en Syntaxe*. Ph.D. dissertation, Université Paris 7.
- Warren, D. H. D. and Pereira, Fernando C. N. 1980. Definite Clause Grammars for Language Analysis — A Survey of the Formalism and a Comparison with Augmented Transition Networks. *Artificial Intelligence* 13, 231–278.
- Woliński, Marcin. 2004. *Komputerowa weryfikacja gramatyki Świdzińskiego*. Ph.D. dissertation, Institute of Computer Science, Polish Academy of Sciences, Warsaw.