A Gapping Analysis of Lexicalised Comparative Constructions

Agnieszka Patejuk

Institute of Computer Science, Polish Academy of Sciences

Abstract

This paper is devoted to lexicalised comparative constructions as dependents of verbs in Polish. It discusses how such constructions can be analysed, offering a gapping-inspired analysis (formalised and implemented in XLE).

1   Introduction

Polish abounds in sentences containing lexicalised comparisons – lexicalised in the sense that the given predicate lexically requires certain forms to appear in the comparison if it is used (usually it is optional). In (1) the verb CHŁONĄĆ 'absorb' requires the nominal GĄBKA 'sponge' to appear in the comparison, while in (2) the comparison used with the verb PASOWAĆ 'fit' must involve the nominal KWIATEK 'flower' and the prepositional phrase featuring the nominal KOŻUCH 'fur coat':

(1)   Chłonęła   wszystko   jak gąbka.
       absorbed.F everything.ACC like sponge.NOM
       'She absorbed everything like a sponge.'   (http://nkjp.pl)

(2)   Halloween   pasuje do naszej kultury   jak kwiatek   do kożucha.
       Halloween.NOM fits   to our   culture.GEN like flower.NOM to fur coat.GEN
       'Halloween fits our culture like a flower (fits) a fur coat.' (literal)
       = it is out of place   (http://nkjp.pl)

It is worth noting that in these examples the phrases following the comparative element JAK correspond to respective dependents of the given verb: in (1) *gąbka* corresponds to the implicit subject of CHŁONĄĆ, while in (2) *kwiatek* corresponds to the subject of PASOWAĆ, *Halloween*, and *do kożucha* corresponds to its oblique dependent, *do naszej kultury*. Moreover, there is a morphosyntactic correspondence in these examples – the respective elements belong to the same syntactic categories (*Halloween* and *kwiatek* are nominal phrases, *do naszej kultury* and *do kożucha* are prepositional phrases) and have the same values of appropriate morphosyntactic features (case, preposition form).

The question is how such comparative elements should be analysed. One of the major issues is the categorial status of the word *jak* in examples such as (1) and (2): is it a preposition, a conjunction, a complementiser or something else?

1.1   PP analysis

The widely adopted view in Polish linguistics (Kallas 1986) is to assume that *jak gąbka* 'like a sponge' in (1) is a prepositional phrase involving the preposition JAK which takes a nominative complement, see the f-structure in (3).[1] Such an analysis is very elegant as it provides a minimal representation without positing any unmotivated elements (which are not represented overtly).

---

[1]Is is perhaps worth mentioning that an analogous analysis was adopted in the English ParGram grammar for sentences such as *John swims like a fish*, where *like a fish* expresses a comparison.

(3)
$$
\begin{bmatrix}
\text{PRED} & \text{`ABSORB}\langle\boxed{1},\boxed{2}\rangle\text{'} \\[4pt]
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\[12pt]
\text{OBJ} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{`EVERYTHING'} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\[18pt]
\text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{`LIKE}\langle\boxed{3}\rangle\text{'} \\ \text{OBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{`SPONGE'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \end{bmatrix} \right\}
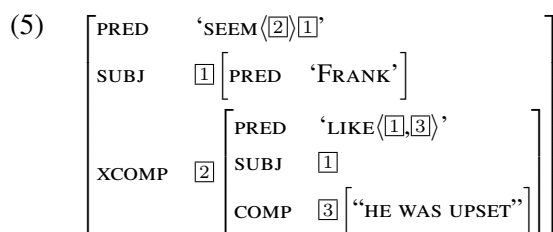\end{bmatrix}
$$
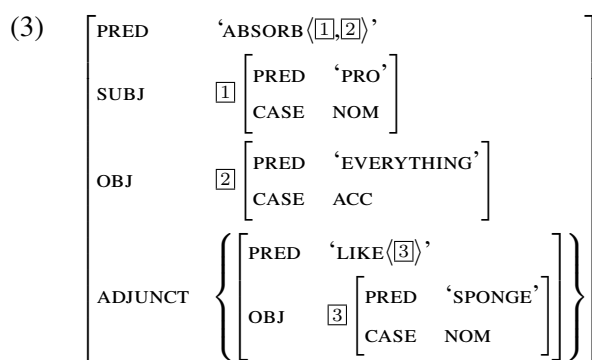
However, it seems that such an analysis could not be used to account for examples such as (2), where the comparison involves 2 elements rather than one (a nominal phrase and a prepositional phrase) – normally, Polish prepositions take only one nominal complement marked for the case required by the preposition. In predicative constructions, the number of arguments of the preposition depends on whether the predicative complement is analysed as an open (XCOMP(-PRED)) or closed (PREDLINK, Butt et al. 1999) grammatical function – under the open analysis the preposition is assumed to have a subject. However, this aspect has no bearing on the issue discussed, because these comparative constructions are not predicative.

Therefore it seems unmotivated to introduce lexical entries for prepositions that would require two (or more) dependents, a nominal and a prepositional phrase in case of (2). Even if such entries were created, interpreting constructions with such special prepositions (taking 2 or more complements) would pose a problem (which grammatical functions should be assigned, on what grounds).

There exist analyses where the comparative element in certain comparative constructions (not lexicalised) is analysed as a preposition – this is the case in the copy raising analysis of English provided in Asudeh and Toivonen 2006. According to this analysis, *like he was upset* in (4) is analysed as an open predicative complement (XCOMP) of SEEM, as shown in the f-structure in (5). The word *like* is a predicative two-argument preposition which takes a subject (SUBJ, structure-shared with the subject of SEEM) and a closed complement (COMP) which contains the rest of the comparison – its subject is assumed to be co-referent with the subject of LIKE (structure-shared with SEEM, as described above).

(4)   Frank seemed like he was upset.

(5)
$$
\begin{bmatrix}
\text{PRED} & \text{`SEEM}\langle\boxed{2}\rangle\boxed{1}\text{'} \\[4pt]
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{`FRANK'} \end{bmatrix} \\[12pt]
\text{XCOMP} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{`LIKE}\langle\boxed{1},\boxed{3}\rangle\text{'} \\ \text{SUBJ} & \boxed{1} \\ \text{COMP} & \boxed{3}\begin{bmatrix} \text{``HE WAS UPSET''} \end{bmatrix} \end{bmatrix}
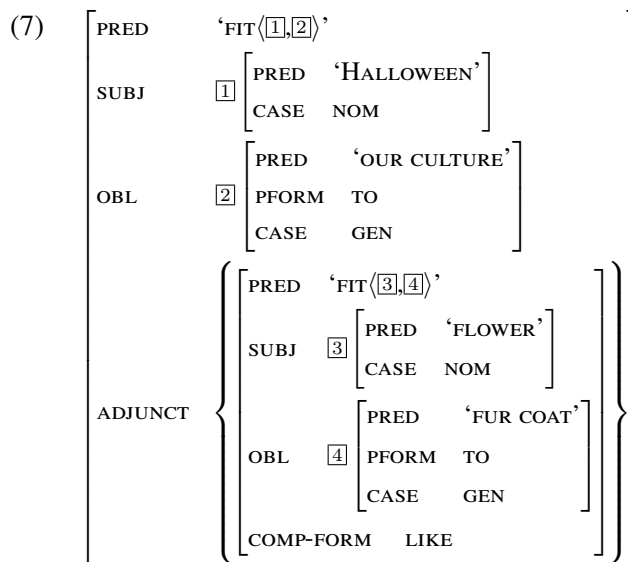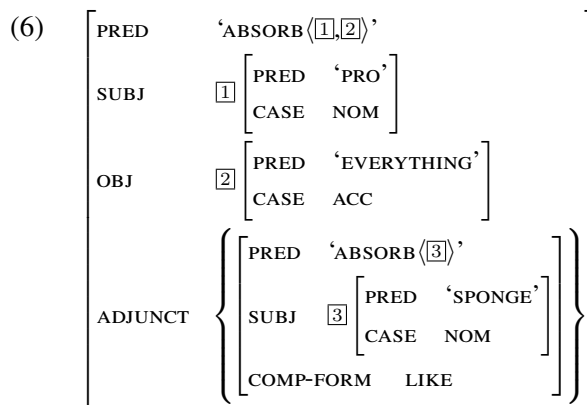\end{bmatrix}
$$

Though such an analysis is a theoretical possibility, there seems to be no evidence that would support adopting it for Polish comparative constructions such as

in (1)–(2), because these do not have the properties of copy raising. Also, there are no Polish prepositions that would take the COMP grammatical function.

1.2 CP analysis

Due to the way comparative constructions such as in (1) and (2) are used and interpreted, the alternative is to treat JAK as a complementiser introducing a comparative clause whose dependents are the phrases following JAK. Though such an analysis is appealing from the perspective of representation (and semantics), the fundamental problem is the absence of an overt predicate in the comparison. A potential solution is to treat such comparative constructions as an instance of gapping (see Patejuk and Przepiórkowski 2017 and references there), where the predicate of the gapped comparative clause is reconstructed on the basis of the predicate from the main clause: see (6), which corresponds to (1), and (7), which is the f-structure for (2).[2]

(6)
$$
\begin{bmatrix}
\text{PRED} & \text{`ABSORB}\langle\boxed{1},\boxed{2}\rangle\text{'} \\
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\
\text{OBJ} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{`EVERYTHING'} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\
\text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{`ABSORB}\langle\boxed{3}\rangle\text{'} \\ \text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{`SPONGE'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{COMP-FORM} & \text{LIKE} \end{bmatrix} \right\}
\end{bmatrix}
$$

(7)
$$
\begin{bmatrix}
\text{PRED} & \text{`FIT}\langle\boxed{1},\boxed{2}\rangle\text{'} \\
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{`HALLOWEEN'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\
\text{OBL} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{`OUR CULTURE'} \\ \text{PFORM} & \text{TO} \\ \text{CASE} & \text{GEN} \end{bmatrix} \\
\text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{`FIT}\langle\boxed{3},\boxed{4}\rangle\text{'} \\ \text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{`FLOWER'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBL} & \boxed{4}\begin{bmatrix} \text{PRED} & \text{`FUR COAT'} \\ \text{PFORM} & \text{TO} \\ \text{CASE} & \text{GEN} \end{bmatrix} \\ \text{COMP-FORM} & \text{LIKE} \end{bmatrix} \right\}
\end{bmatrix}
$$

---

[2]For the sake of simplifying the representation, the f-structures presenting the CP gapping analysis uniformly assume that the complementiser is a co-head contributing the COMP-FORM attribute and that the entire comparative clause is an adjunct of the main predicate. However, these can be modified (CP as an argument, complementiser as the head) without affecting the general analysis.

Such an analysis seems to be supported by the fact that it is possible to use an overt instance of the predicate in the comparison (though it is typically omitted since it is the same as in the main clause), see (8) and (9).

(8) chłoniesz wiedzę   jak gąbka   chłonie   wodę.
     absorb.2.SG knowledge.ACC like sponge.NOM absorb.3.SG water.ACC
     'You absorb knowledge like a sponge absorbs water.'   (Google)

(9) tak pasuje do otoczenia   jak kwiatek   pasuje do kożucha.
     so fit.3.SG to surrounding.GEN like flower.NOM fit.3.SG to fur coat.GEN
     'He fits the surroundings like a flower fits a fur coat.' (literal)
     = he does not fit   (Google)

Note that the verb forms used in corresponding clauses do not need to be identical – they may have different agreement features: *chłoniesz* is a second person singular form, while *chłonie* is a third person singular form. The same holds in the case of typical instances of gapping in Polish (Patejuk and Przepiórkowski 2017).

2   More complex data and interactions

The following sections present data supporting the analysis which involves the use of a gapped predicate in the comparative clause.

2.1   Numeral phrases

While there seem to be good reasons to analyse comparative constructions in examples such as (2) as gapped clauses, it appears that comparisons such as in (1) could minimally be analysed as prepositional phrases. This would lead to a split analysis of comparative constructions, resulting in different, non-parallel syntactic representation. This is not a problem if making such a distinction is well-motivated, it seems, however, that examples such as (10) and (11) provide evidence that comparisons involving one argument should be treated as an instance of a clause rather than a prepositional phrase, arguing in favour of a unified gapping analysis.

(10) Kobieta   wyglądała jak siedem   nieszczęść.
     woman.NOM looked   like seven.ACC disasters.GEN
     'Woman looked like seven disasters.' (literal)
     = she looked miserable   (http://nkjp.pl)

(11) wygląda jak pół   dupy   zza   krzaka
     looks   like half.ACC arse.GEN behind bush
     'It looks like a half of an arse from behind a bush.' (literal)
     = it looks stupid, bad   (http://nkjp.pl)

The clausal analysis, see (12) which corresponds to (10), is supported by the fact that the element inside the comparison is a non-agreeing numeral phrase, where the numeral head (*siedem* 'seven', *pół* 'half') takes accusative case and the accompanying nominal (*nieszczęść* 'disasters', *dupy* 'arse') is marked for genitive case.[3]

---

[3]See Przepiórkowski 2006 for discussion of PÓŁ as a non-agreeing numeral.

(12)
```
┌                                                              ┐
│ PRED      'LOOK⟨1⟩'                                          │
│                                                              │
│           ┌ PRED   'WOMAN' ┐                                 │
│ SUBJ    1 │ CASE   NOM     │                                 │
│           └                ┘                                 │
│                                                              │
│           ⎧ ┌ PRED      'LOOK⟨2⟩'                    ┐ ⎫     │
│           ⎪ │                                        │ ⎪     │
│           ⎪ │            ┌ PRED   'SEVEN⟨3⟩'      ┐  │ ⎪     │
│           ⎪ │            │ CASE   ACC            │  │ ⎪     │
│ ADJUNCT   ⎨ │ SUBJ    2  │                       │  │ ⎬     │
│           ⎪ │            │        ┌ PRED  'DISASTER' ┐│ ⎪   │
│           ⎪ │            │ OBJ  3 │ CASE  GEN        ││ ⎪   │
│           ⎪ │            └        └                 ┘┘  │ ⎪ │
│           ⎪ │                                        │ ⎪     │
│           ⎩ │ COMP-FORM    LIKE                      │ ⎭     │
│             └                                        ┘        │
└                                                              ┘
```
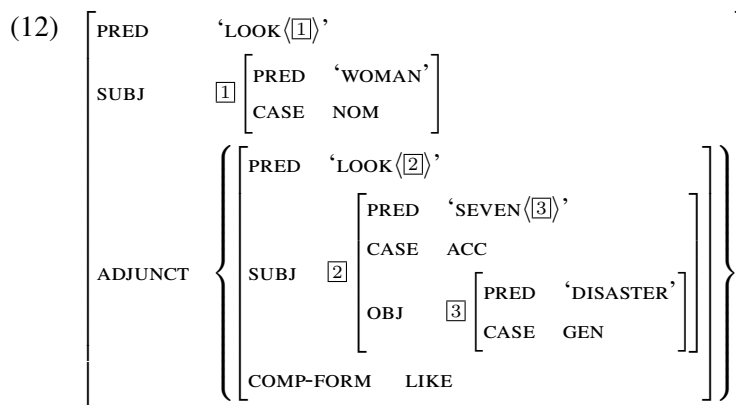
As discussed in the literature devoted to case assignment and agreement in Polish (e.g. Przepiórkowski 1999, Przepiórkowski and Patejuk 2012), such case marking is characteristic of non-agreeing numeral subjects in Polish – expected in clauses.

An alternative solution would be to allow, alongside nominative nominals, accusative non-agreeing (taking a genitive nominal object) numeral phrases with JAK as the preposition. As a result, the putative preposition would be required to allow as its complement (typically OBJ) nominative nominals and accusative non-agreeing numerals, both of which perfectly match subjecthood criteria in Polish, while disallowing non-numeral accusative nominals, which cannot act as a subject in Polish. Such a coincidence seems to strongly suggest that a generalisation is being missed, namely that the comparison should be analysed as a clause and the numeral as its subject – even in cases where there is only one dependent following JAK.

2.2 Passive

There are interesting interactions between comparative constructions and passive voice – it seems that the comparative dependent undergoes passivisation in the same way as the main predicate. In (13) the object of the verb WYDOIĆ 'milk', *nas* 'us', is marked for accusative case (structural case in the absence of negation) and the comparison includes the phrase *krowę* 'cow' – also an accusative form. When the verb is passivised, as in (14), the form of KROWA 'cow' inside the comparison is nominative, which is the same as the case of the implicit subject of the main verb.

(13) Gdy już   nas   wydoją jak krowę.
      When already us.ACC milk   like cow.ACC
      'When they have already milked us like a cow.' (literal)
      = they have milked us dry   (http://nkjp.pl)

(14) Zostaniesz wydojony   jak krowa.
      become   milk.PASS.NOM like cow.NOM
      'You will be milked like a cow.' (literal)
      = you will be milked dry   (Google)

Another piece of evidence is provided by examples (15) and (16). This pair features the predicate PILNOWAĆ 'watch, keep an eye on somebody' which is different from

WYDOIĆ 'milk' discussed above in that its complement is not marked for structural case as in (13), but for lexical genitive case, see (15). Still, in Polish being marked for lexical case (not being accusative) does not preclude such an object from undergoing passivisation, as shown in (16).

(15) Pilnujcie jej   jak oka   w głowie.
     watch   she.GEN like eye.GEN in head
     'Watch her like an eye in the head.' (literal)
     = watch her closely   (http://nkjp.pl)

(16) będzie pilnowany   jak oko   w głowie
     will   watch.PASS.NOM like eye.NOM in head
     'He will be watched like an eye in the head.' (literal)
     = he will be watched closely   (http://nkjp.pl)

As in (13)–(14), the dependents in the comparisons in (15)–(16) seem to undergo passivisation in the same way as the corresponding dependents of the main clause – they are marked for case according to the same rules that apply to the dependents of the main predicate. The active predicate in (15) assigns lexical genitive case to its object and the dependent following JAK is also marked for genitive. Under passivisation the active object becomes the passive subject – in (16) the passive predicate assigns nominative case to its implicit subject and the dependent following JAK is also marked for nominative case.

It is worth highlighting that even if the resulting values of case are identical (as in the examples above), it would be wrong to assume that the respective values are copied from the main predicate to the comparison. Instead, they are assigned independently according to appropriate rules. As a result, the subject of (16) could be a non-agreeing accusative numeral, while the subject inside the comparison will remain nominative because it is a plain nominal:[4]

(17) pięciu   zakładników było pilnowanych   jak oko   w głowie
     five.ACC hostages.GEN were watch.PASS.ACC/GEN like eye.NOM in head
     'Five hostages were watched like an eye in the head.' (literal)
     = they were watched closely

These observations provide another argument supporting the clausal analysis of comparative constructions involving gapping – the f-structures in (18) and (19) correspond to (13) and (14), respectively.[5]

---

[4]The form *pilnowanych* 'watched' in (17) is glossed as ambiguous (accusative or genitive) because it may either agree with the accusative numeral head or its genitive nominal dependent (see Przepiórkowski and Patejuk 2012 for discussion and analysis).

[5]For the sake of brevity and simplicity of representation, a flat analysis of passive is used in this paper, where the lexical verb is the main verb (contributing PRED), while the auxiliary is a co-head (without PRED). However, as discussed in Patejuk and Przepiórkowski (2014a), a raising analysis of passive should be used for Polish (the auxiliary is the main verb, the lexical verb is its complement).

(18)
$$\begin{bmatrix} \text{PRED} & \text{`MILK}\langle\boxed{1},\boxed{2}\rangle\text{'} \\ \text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{`WE'} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\ \text{ADJUNCT} & \left\{\begin{bmatrix} \text{PRED} & \text{`MILK}\langle\boxed{3},\boxed{4}\rangle\text{'} \\ \text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & \boxed{4}\begin{bmatrix} \text{PRED} & \text{`COW'} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\ \text{COMP-FORM} & \text{LIKE} \end{bmatrix}\right\} \end{bmatrix}$$

(19)
$$\begin{bmatrix} \text{PRED} & \text{`MILK}\langle\text{NULL},\boxed{1}\rangle\text{'} \\ \text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{ADJUNCT} & \left\{\begin{bmatrix} \text{PRED} & \text{`MILK}\langle\text{NULL},\boxed{2}\rangle\text{'} \\ \text{SUBJ} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{`COW'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{COMP-FORM} & \text{LIKE} \\ \text{PASSIVE} & + \end{bmatrix}\right\} \\ \text{PASSIVE} & + \end{bmatrix}$$

The alternative analysis treating the comparison as a prepositional phrase would require introducing additional entries to allow for accusative case in (13) and genitive in (15) when the main verb is active and for nominative case in (14) and (16) when the main verb is passive – it seems this would be another case of a missed generalisation.

3 Potential mismatches

Most of the examples discussed so far featured a direct correspondence between the dependents of the main verb and the dependents of the comparative construction in terms of syntactic categories and grammatical functions – the lists of arguments in PRED in the main clause and the comparative clause in (7) and (12) are exactly the same. The following sections discuss examples of different kinds of mismatches and their consequences for the analysis.

3.1 Different number of dependents

The number of dependents in the main clause and the comparison is not the same in (6) since there is no overt object in the comparison in (1). However, an implicit object (marked for accusative case – see the discussion of (21)) could be introduced

in the comparative clause, making the relevant argument lists parallel: see the resulting alternative f-structure for (1) provided in (20).

(20)
$$
\begin{bmatrix}
\text{PRED} & \text{`ABSORB}\langle\boxed{1},\boxed{2}\rangle\text{'} \\
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\
\text{OBJ} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{`EVERYTHING'} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\
\text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{`ABSORB}\langle\boxed{3},\boxed{4}\rangle\text{'} \\ \text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{`SPONGE'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & \boxed{4}\begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\ \text{COMP-FORM} & \text{LIKE} \end{bmatrix} \right\}
\end{bmatrix}
$$
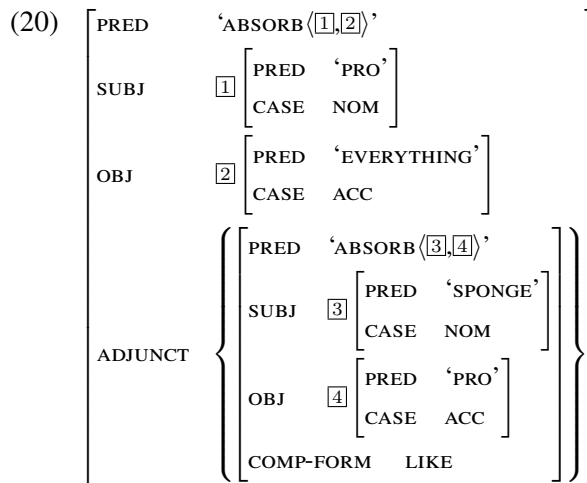
Such an analysis could be supported by the fact that there are examples where the overt object is present in the comparison – it must be a form of WODA 'water' marked for accusative case (the value of structural case appropriate for the object in the absence of negation) – see (8) where the predicate inside the comparison is represented overtly and (21), with the f-structure in (22), where the predicate is missing:

(21) Chłonęli Zachód   jak gąbka   wodę.
     absorbed West.ACC like sponge.NOM water.ACC
     'They absorbed the West like a sponge (absorbs) water.'   (http://nkjp.pl)

(22)
$$
\begin{bmatrix}
\text{PRED} & \text{`ABSORB}\langle\boxed{1},\boxed{2}\rangle\text{'} \\
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\
\text{OBJ} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{`WEST'} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\
\text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{`ABSORB}\langle\boxed{3},\boxed{4}\rangle\text{'} \\ \text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{`SPONGE'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & \boxed{4}\begin{bmatrix} \text{PRED} & \text{`WATER'} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\ \text{COMP-FORM} & \text{LIKE} \end{bmatrix} \right\}
\end{bmatrix}
$$
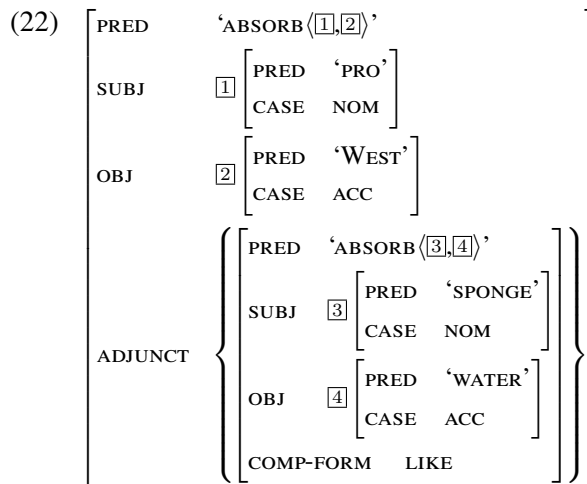
Still, there are cases where it is difficult to argue for the presence of implicit arguments – consider (23), where the main verb (LATAĆ 'fly, run') takes a prepositional phrase consisting of the preposition ZA and a nominal marked for instrumental case. By contrast, the comparison does not include such an argument in any way – neither lexical nor implicit.

314

(23) latał za nią   jak kot   z   pęcherzem
     ran after she.INST like cat.NOM with bladder
     'He ran after her like a cat with a bladder (attached to it as a prank).' (literal)
     = he was always chasing after her like mad   (http://nkjp.pl)

It seems that *z pęcherzem* in (23) can be analysed in three ways – as a dependent of *kot*, as an adjunct in the comparison or as its argument. In the first two cases, the argument lists of the predicates in the main clause and in the comparison would be distinct – see (24) where *z pęcherzem* is as an adjunct in the comparison. This suggests that the phenomenon that is being discussed is not a typical instance of gapping, because the predicates in the two clauses, the main clause and the comparison, may take a different number of arguments.

(24)
$$
\begin{bmatrix}
\text{PRED} & \text{'RUN}\langle\boxed{1},\boxed{2}\rangle\text{'} \\
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\
\text{OBL} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{'SHE'} \\ \text{PFORM} & \text{AFTER} \\ \text{CASE} & \text{INST} \end{bmatrix} \\
\text{ADJUNCT} & \left\{\begin{bmatrix} \text{PRED} & \text{'RUN}\langle\boxed{3}\rangle\text{'} \\ \text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{'CAT'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{ADJUNCT} & \left\{\begin{bmatrix} \text{PRED} & \text{'WITH}\langle\boxed{4}\rangle\text{'} \\ \text{OBJ} & \boxed{4}\begin{bmatrix} \text{PRED} & \text{'BLADDER'} \\ \text{CASE} & \text{INST} \end{bmatrix} \end{bmatrix}\right\} \\ \text{COMP-FORM} & \text{LIKE} \end{bmatrix}\right\}
\end{bmatrix}
$$

In the case where the prepositional phrase *z pęcherzem* is analysed as an argument in the comparison, it would be assigned the OBL grammatical function. Even though it happens to be the same grammatical function as the one assigned to *za nią*, these dependents clearly do not correspond to each other – they are not compared to each other as in other pairs discussed so far. Moreover, distinct constraints apply to these prepositional phrases: the main clause predicate requires the preposition ZA, while the comparison requires the preposition z. So, even if the number of arguments is assumed to be the same (by assuming that *z pęcherzem* is an argument), there is still the issue of different morphosyntactic requirements imposed on dependents of the predicates in the main clause and in the comparison. Again, this could not be accounted for under a standard gapping analysis whereby both predicates would be required to have the same requirements (possibly disjunctive).

3.2 Different constraints in dependents

Since in (23) the obliques in the main clause (*za nią*) and in the comparison (*z pęcherzem*) do not correspond to each other, let us discuss (25) and (26) where
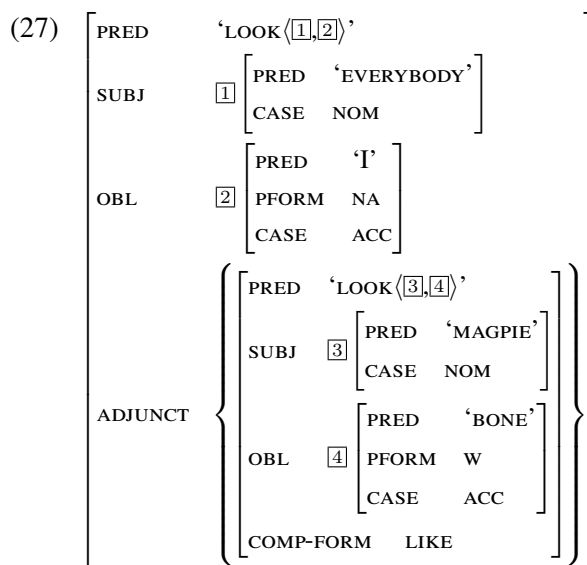
such a correspondence exists (corresponding phrases are compared to each other), but there are differences in mophosyntactic constraints imposed on them:

(25) wszyscy   na mnie patrzyli jak sroka   w kość.
      everybody at I.ACC looked like magpie.NOM in bone.ACC
      'Everybody looked at me like a magpie at a bone.' (literal)
      = they stared at me   (http://nkjp.pl)

(26) na premiera to on   sie   nadawał jak wół   do karety.
      for PM.ACC FOC he.NOM REFL fit   like ox.NOM to carriage/coach.GEN
      'He was fit to be a Prime Minister as an ox fits a carriage/coach.' (literal)
      = he was not fit to be a PM   (http://nkjp.pl)

In both examples all prepositional phrases (in the main clause and in the comparison) would be analysed as obliques. However, different prepositions are used in corresponding prepositional phrases. In (25) the main verb uses NA as the preposition, while the comparison must use W as the preposition – see the f-structure in (27).[6] Similarly, in (26) the main predicate features the preposition NA, while DO must be used in the comparison.

(27)

$$
\begin{bmatrix}
\text{PRED} & \text{'LOOK}\langle\boxed{1},\boxed{2}\rangle\text{'} \\[2mm]
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{'EVERYBODY'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\[4mm]
\text{OBL} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{'I'} \\ \text{PFORM} & \text{NA} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\[6mm]
\text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{'LOOK}\langle\boxed{3},\boxed{4}\rangle\text{'} \\ \text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{'MAGPIE'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBL} & \boxed{4}\begin{bmatrix} \text{PRED} & \text{'BONE'} \\ \text{PFORM} & \text{W} \\ \text{CASE} & \text{ACC} \end{bmatrix} \\ \text{COMP-FORM} & \text{LIKE} \end{bmatrix} \right\}
\end{bmatrix}
$$

The situation here is complicated by the fact that the main predicates in (25)–(26) have disjunctive subcategorisation requirements – the comparison in (25) can be used with NA or W as the preposition in the main predicate, or their coordination (phrases are coordinated). The possibility of coordination suggests that there is one valency schema for the verb PATRZYĆ 'look' which disjunctively specifies its oblique argument in the main clause (Patejuk and Przepiórkowski 2012).

This would pose a problem for the standard gapping account, whereby the predicates in the main clause and in the gapped clause are expected to be subject to the

---

[6]In (27) the values of PFORM are in Polish to show that they are distinct (unlike their translation).

same constraints. This is because the comparison is more restricted – it must use only one of the options allowed by the main predicate: in (25) only *w* is appropriate as the preposition in the comparison, it must not be NA (which is possible in the main clause). The case of (26) is analogous – the main predicate allows either NA or DO as the preposition, while only the latter can be used in the comparison.

Note that while a standard gapping analysis cannot account for such restrictions, the problem could easily be solved by stating such constraints in the lexicon – by requiring the appropriate preposition forms (and cases) to appear in the comparison.

3.3 Different grammatical functions of dependents

There is another type of mismatch between dependents of the main clause and the comparison – this is the mismatch in the assignment of grammatical functions to these dependents, which seems to pose another problem to typical gapping analyses.

(28) Pasuje on   do reszty jak świni   siodło.
     fits   he.NOM to rest.GEN like pig.DAT saddle.NOM
     'He fits the rest like a saddle (fits) a pig.' (literal)
     = he is out of place   (http://nkjp.pl)

In (28) the main clause contains a nominative subject (*on* 'he') and a prepositional (*do reszty* 'to rest') oblique (OBL), while the comparison contains the corresponding nominative subject (*siodło* 'saddle') and a dative nominal phrase (*świni* 'pig'), typically analysed as an oblique object (OBJ$_\theta$) in Polish, see (29).

(29)
$$
\begin{bmatrix}
\text{PRED} & \text{'FIT}\langle\boxed{1},\boxed{2}\rangle\text{'} \\
\text{SUBJ} & \boxed{1}\begin{bmatrix}\text{PRED} & \text{'PRO'} \\ \text{CASE} & \text{NOM}\end{bmatrix} \\
\text{OBL} & \boxed{2}\begin{bmatrix}\text{PRED} & \text{'REST'} \\ \text{PFORM} & \text{TO} \\ \text{CASE} & \text{GEN}\end{bmatrix} \\
\text{ADJUNCT} & \left\{\begin{bmatrix}\text{PRED} & \text{'FIT}\langle\boxed{3},\boxed{4}\rangle\text{'} \\ \text{SUBJ} & \boxed{3}\begin{bmatrix}\text{PRED} & \text{'SADDLE'} \\ \text{CASE} & \text{NOM}\end{bmatrix} \\ \text{OBJ}_\theta & \boxed{4}\begin{bmatrix}\text{PRED} & \text{'PIG'} \\ \text{CASE} & \text{DAT}\end{bmatrix} \\ \text{COMP-FORM} & \text{LIKE}\end{bmatrix}\right\}
\end{bmatrix}
$$

Unlike in (25)–(26), where the main predicate has disjunctive subcategorisation requirements (the given argument can be filled by category A, category B or their coordination), this is not the case with the verb PASOWAĆ 'fit' – it can either take the prepositional phrase (with DO as the preposition, requiring a genitive nominal) or a dative nominal phrase, but not their coordination. This is why there must be four valency schemata for this verb – two where the categories in the main clause and

in the comparison match and two where these do not match (oblique in the main clause as opposed to dative indirect object in the comparison, as in (29), or the other way round; subjects are assumed to match in both cases).

Unlike in (23), where under the analysis which treats *z pęcherzem* as an oblique (rather than an adjunct) the dependents have the same grammatical function (but different morphosyntactic properties) yet they do not correspond to one another (they are not compared), in (28) such mismatched dependents (different grammatical function, different morphosyntactic constraints) do correspond to each other and are comparable. While in (23) there is no way to express the fact that these phrases do not correspond to one another, in (28) there is no way to express the opposite – that such mismatched phrases do correspond to each other.

A potential solution to this problem would be to introduce an explicit representation of the correspondence between the respective dependents of the main predicate and the comparison. This can be achieved by introducing a dedicated attribute, for instance REF, in the f-structure of the relevant dependent inside the comparison – its value would be structure-shared with the f-structure of the corresponding dependent from the main clause. If the REF attribute is absent, such a dependent does not have a counterpart in the main clause. The equations in (30),[7] placed in the lexical entry of the main verb, represent the correspondences in (28): between the subjects and between the main clause oblique and the OBJ$_\theta$ inside the comparison. The resulting extended f-structure representation of (28) is provided in (31).

(30) $(\uparrow$ SUBJ$)= ($%C SUBJ REF$)$
      $(\uparrow$ OBL$)= ($%C OBJ$_\theta$ REF$)$

(31)
$$
\begin{bmatrix}
\text{PRED} & \text{`FIT}\langle\boxed{1},\boxed{2}\rangle\text{'} \\[4pt]
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\[12pt]
\text{OBL} & \boxed{2}\begin{bmatrix} \text{PRED} & \text{`REST'} \\ \text{PFORM} & \text{TO} \\ \text{CASE} & \text{GEN} \end{bmatrix} \\[18pt]
\text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{`FIT}\langle\boxed{3},\boxed{4}\rangle\text{'} \\[4pt] \text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{`SADDLE'} \\ \text{CASE} & \text{NOM} \\ \text{REF} & \boxed{1} \end{bmatrix} \\[14pt] \text{OBJ}_\theta & \boxed{4}\begin{bmatrix} \text{PRED} & \text{`PIG'} \\ \text{CASE} & \text{DAT} \\ \text{REF} & \boxed{2} \end{bmatrix} \\[14pt] \text{COMP-FORM} & \text{LIKE} \end{bmatrix} \right\}
\end{bmatrix}
$$

---

[7]%C is a variable which corresponds to an element of the ADJUNCT set, $(\uparrow$ ADJUNCT $\in)=$ %C, as defined in (35). See §4 for a detailed discussion of the formalisation.

3.4 Voice mismatch

Let us consider the following pair of examples. (33) seems to show that, unlike in examples discussed in §2.2, the comparative construction does not passivise when the main verb does – *jak kura pazurem* used in active (32) remains in passive (33).

(32) Nawet małe dzieci potrafią nabazgrolić jak (pijana) kura
even small children.NOM can scribble.INF like drunk.NOM hen.NOM
pazurem.
claw.INST
'Even small children can scribble like a drunk hen with a claw.' (literal)
= they can write in a spidery scrawl (Google)

(33) Na karteczce było coś nabazgrolone jak (*pijana)
on card was something.NOM scribble.PPAS.NOM like drunk.NOM
kura pazurem.
hen.NOM claw.INST
'Something was scribbled on the card like a hen with a claw.' (literal)
= it was written in a spidery scrawl (Google)

Such voice mismatch could be approached using the mechanism for representing the correspondence between the dependents of the main predicate and the comparison proposed in §3.3. In (32) the active main clause subject (*małe dzieci* 'small children') corresponds to the active subject in the comparison (*pijana kura* 'drunk hen'), while in (33) the active subject *kura* would correspond to NULL in the passive main clause (there is no oblique agent) – see the f-structure in (34).

(34)
$$\begin{bmatrix} \text{PRED} & \text{`SCRIBBLE}\langle \text{NULL}, \boxed{1} \rangle\text{'} \\ \text{SUBJ} & \boxed{1} \begin{bmatrix} \text{PRED} & \text{`SOMETHING'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{`SCRIBBLE}\langle \boxed{2}, \boxed{3} \rangle\text{'} \\ \text{SUBJ} & \boxed{2} \begin{bmatrix} \text{PRED} & \text{`HEN'} \\ \text{CASE} & \text{NOM} \\ \text{REF} & \text{NULL} \end{bmatrix} \\ \text{OBL}_{inst} & \boxed{3} \begin{bmatrix} \text{PRED} & \text{`CLAW'} \\ \text{CASE} & \text{INST} \end{bmatrix} \\ \text{COMP-FORM} & \text{LIKE} \end{bmatrix} \right\} \\ \text{PASSIVE} & + \end{bmatrix}$$
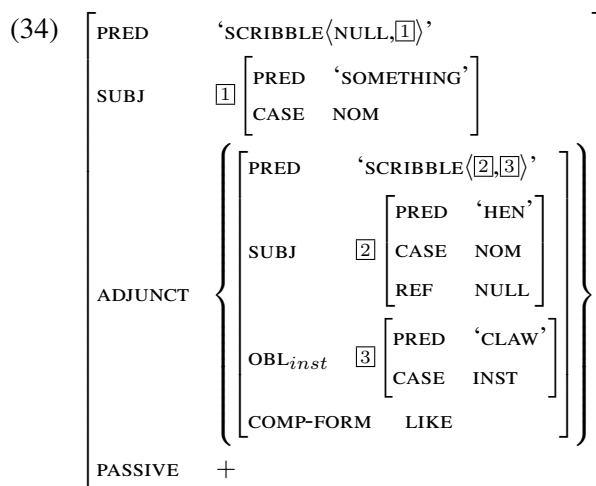
The other solution is to assume that the phrase *jak kura pazurem* 'like hen with a claw' in (33) is a lexicalised fixed expression roughly meaning *in an ugly way* (of writing) – a multiword expression functioning as one unit and thus not having internal structure. This approach could be supported by the difference in modification patterns – in (32), where the predicate is active, the comparative dependent

319

may be modified with *pijana* 'drunk', which suggests that it has internal structure and therefore it could be analysed under the gapping analysis. By contrast, such modification is not possible in (33) under passive voice in the main clause.

4   Formalisation

Since the comparative constructions discussed in this paper are lexicalised, the relevant requirements are formalised in the lexicon – in the lexical entry of the relevant predicate taking a comparative dependent.[8]

First, (35) assigns the path to the comparative construction to the variable %C: it is an element of the verb's adjunct set (due to the assumption stated in fn. 2, which, as explained there, could be changed).

(35) $(\uparrow \textsc{adjunct} \in)= \%\textsc{c}$

Subsequently, this variable is used in the specification of the PRED of the gapped comparative clause – the constraint in (37) states that the comparison includes the subject and an oblique (it corresponds to (2), repeated in (39)), while in (38) it has a subject and an indirect object $\textsc{obj}_\theta$ (it covers (28), repeated in (40)).[9] (36) states the PRED value of the main verb (it takes a subject and an oblique in both examples).

(36) $(\uparrow \textsc{pred})=$'$\textsc{fit}{<}(\uparrow \textsc{subj}),(\uparrow \textsc{obl}){>}$'

(37) $(\%\textsc{c pred})=$'$\textsc{fit}{<}(\%\textsc{c subj}),(\%\textsc{c obl}){>}$'

(38) $(\%\textsc{c pred})=$'$\textsc{fit}{<}(\%\textsc{c subj}),(\%\textsc{c obj}_\theta){>}$'

(39)   Halloween   pasuje do naszej kultury   jak kwiatek   do kożucha.
       Halloween.NOM fits   to our   culture.GEN like flower.NOM to fur coat.GEN
       'Halloween fits our culture like a flower (fits) a fur coat.' (literal)
       = it is out of place   (http://nkjp.pl)

(40)   Pasuje on   do reszty   jak świni   siodło.
       fits   he.NOM to rest.GEN like pig.DAT saddle.NOM
       'He fits the rest like a saddle (fits) a pig.' (literal)
       = he is out of place   (http://nkjp.pl)

Next, the %C variable is used when imposing lexicalised constraints such as the form of the relevant argument inside the comparison. (41) states constraints on (37) – it requires that the subject must be a form of FLOWER, while the oblique must be a form of FUR COAT (inside a prepositional phrase). The constraints in (42) match

---

[8]As mentioned in §1, in most cases the comparison is optional. The formalisation presented here models cases where the lexicalised comparison is used – as a result, the comparison is required and its dependents must satisfy all the constraints specified in the lexicon. However, unless the comparison is required (as in §5), the relevant lexical entry contains additional constraints making it possible to use a given predicate without such a comparison (not presented here for reasons of space).

[9]The constraints in (37) and (38) use the explicit lemma (FIT) for clarity, but the %STEM variable can be used in XLE – it supplies the stem of the lexical entry.

the PRED in (38): the subject must be a form of SADDLE, the indirect object must be a form of PIG. The constraints in (38) and (42) are accompanied by the constraint in (30), repeated in (43), which provides the mapping between the dependents of the comparison and the main clause using the REF attribute (as explained in §3.3).

(41) $(\%c$ SUBJ PRED$)=_c$ FLOWER   (42) $(\%c$ SUBJ PRED$)=_c$ SADDLE
       $(\%c$ OBL PRED$)=_c$ FUR COAT $(\%c$ OBJ$_\theta$ PRED$)=_c$ PIG

(43) $(\uparrow$ SUBJ$)= (\%c$ SUBJ REF$)$
       $(\uparrow$ OBL$)= (\%c$ OBJ$_\theta$ REF$)$

The $\%c$ variable is also used for imposing appropriate morphosyntactic constraints on the relevant dependents in the comparison – these include restricting the number of the dependent and, more importantly, imposing appropriate valency requirements such as the appropriate preposition form or case, especially structural case.

(44) $(\%c$ OBL PFORM$)=_c$ TO   (46) $(\%c$ OBJ$_\theta$ CASE$)=_c$ DAT

(45) $(\%c$ OBL CASE$)=_c$ GEN   (47) @(STRCS $(\%c$ SUBJ$))$

The constraints in (44)–(45) apply to the oblique (OBL) inside the comparison – see the corresponding PRED in (37) and the example in (39). (44) makes sure that the preposition form is TO, while (45) ensures that the nominal in the prepositional phrase is marked for genitive case.

     The constraint in (46) is a simple constraint on case, which corresponds to the PRED in (38) and example (40) – the indirect object (OBJ$_\theta$) inside the comparison is required to bear dative case – it is a lexical case, as opposed to structural case.

     The template STRCS handles structural case assignment to the subject in the grammar – if the element in the path given as its argument is a noun, nominative case is required, if it is a non-agreeing numeral, accusative case is required (see §2.1 and references therein). Since the template call in (47) contains the path which points to the subject of the comparison, it will impose relevant constraints there, independently of the case of the main clause subject (see the discussion of (17) in §2.2). Since in both examples considered above the subject is a noun (FLOWER, SADDLE), it must be marked for nominative case. However, accusative case will be required in (10) because the subject there is a non-agreeing numeral (SEVEN).

     Let us consider the formalisation of examples involving passivisation discussed in §2.2. (50) is the the full lexical entry of the verb WYDOIĆ 'milk' – it contains all the constraints[10] needed to account for examples (13)–(14), repeated as (48)–(49).

(48) Gdy już   nas   wydoją jak krowę.
     When already us.ACC milk   like cow.ACC
     'When they have already milked us like a cow.' (literal)
     = they have milked us dry   (http://nkjp.pl)

---

[10]In (50) STRCO is a template handling structural case assignment to the object – it takes the path to the object as its argument and assigns an appropriate case (see Patejuk and Przepiórkowski 2014b).

(49) Zostaniesz wydojony   jak krowa.
     become   milk.PASS.NOM like cow.NOM
     'You will be milked like a cow.' (literal)
     = you will be milked dry   (Google)

The first disjunct of (50) handles active voice, see (48), while the second one handles the passive, see (49). Both disjuncts provide an appropriate mapping between dependents of the main clause and the comparison using the REF attribute.

(50) *wydoić*   V $(\uparrow$ ADJUNCT $\in)=$ %C
     $[\neg(\uparrow$ PASSIVE) $\neg($%C PASSIVE)
     $(\uparrow$ PRED)$=$'MILK$<(\uparrow$ SUBJ),$(\uparrow$ OBJ)$>$'
     @(STRCS $(\uparrow$ SUBJ))
     @(STRCO $(\uparrow$ OBJ))
     (%C PRED)$=$'MILK$<$(%C SUBJ),(%C OBJ)$>$'
     @(STRCS (%C SUBJ))
     @(STRCO (%C OBJ))
     (%C OBJ PRED)$=_c$ COW
     $(\uparrow$ SUBJ)$=$ (%C SUBJ REF)
     $(\uparrow$ OBJ)$=$ (%C OBJ REF)]
     $\vee$
     $[(\uparrow$ PASSIVE)$=_c +$ (%C PASSIVE)$= +$
     $(\uparrow$ PRED)$=$'MILK$<$NULL,$(\uparrow$ SUBJ)$>$'[11]
     (%C PRED)$=$'MILK$<$NULL,(%C SUBJ)$>$'
     (%C SUBJ PRED)$=_c$ COW
     @(STRCS $(\uparrow$ SUBJ))
     @(STRCS (%C SUBJ))
     $(\uparrow$ SUBJ)$=$ (%C SUBJ REF)]

Note that (50) is a fully expanded lexical entry for the verb WYDOIĆ in the use which uses a lexicalised comparison. In principle, it should be possible to equivalently use the active part of (50) in conjunction with a lexical rule for handling passive voice.[12]

Finally, the second part of the formalisation of the proposed analysis is creating c-structure rules necessary to build the relevant structures. (51) is a rule which builds the gapped clause – it consists of the co-head JAK, analysed as the complementiser, followed by a non-zero sequence of the DEP co-head, defined in (52).

(51) CP$_{compar}$ $\rightarrow$ COMP   DEP^{+}
                              $\uparrow=\downarrow$ $\uparrow=\downarrow$
                     $(\uparrow$ COMP-FORM)$=_c$ JAK

---

[11]If the oblique expressing the agent (OBL$_{ag}$) may be used with the passivised main predicate (it seems to be a possibility), the following disjunctive specification of PRED should be used instead: $[(\uparrow$ PRED)$=$'MILK$<$NULL,$(\uparrow$ SUBJ)$>$' $\vee$ $(\uparrow$ PRED)$=$'MILK$<(\uparrow$ OBL$_{ag}$),$(\uparrow$ SUBJ)$>$']. If it is possible, it would not be a problem that the main predicate takes an OBL$_{ag}$ while the comparison does not (it uses NULL).

[12]However, the statements related to the PASSIVE attribute would need to be added on top of this.

(52) DEP ≡ {   NP | PP | ... }
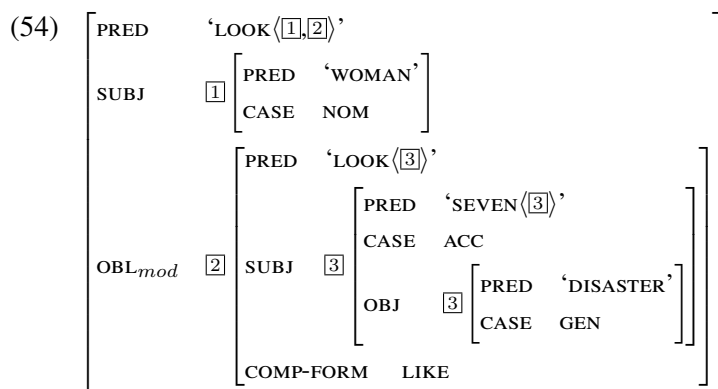   (↑ {SUBJ|OBJ|OBJ$_\theta$|...})=↓   (↑ OBL)=↓

According to (52), DEP corresponds[13] to various dependents used in the grammar – the annotation of each right hand side disjunct builds the partial f-structure, for instance it states that an NP is a subject, an object or an indirect object (this is restricted using case constraints imposed by lexical entries – see (46), (47) and (50)). Because of the co-head annotation of DEP in (51), partial structures of relevant DEPs are merged into one f-structure in (51) and matched against the constraints in the lexical entry of the verb (constraints of appropriate lemmata, morphosyntactic constraints).

5 Potential alternative analysis of JAK

Upon closer scrutiny the representation of (10) provided in (12) seems to pose an interesting potential problem to the current analysis. Unlike in most other cases discussed so far, *jak siedem nieszczęść* is a required dependent of the main verb – without it the sentence becomes ungrammatical, as shown in (53):

(53) Kobieta   wyglądała *(jak siedem   nieszczęść).
      woman.NOM looked   like seven.ACC disasters.GEN

In Polish, there are predicates such as WYGLĄDAĆ 'look', ZACHOWYWAĆ SIĘ 'behave' or CZUĆ SIĘ 'feel' that obligatorily take a dependent which expresses manner – in some approaches these are modelled as arguments (listed in PRED), see (54),[14] while in other approaches these are treated as obligatory adjuncts (which, according to the most prominent definition of adjuncts, is a contradiction of terms), see (12). Whichever analysis is adopted, such a dependent expressing manner must be present since its absence results in ungrammaticality.

(54)
$$
\begin{bmatrix}
\text{PRED} & \text{'LOOK}\langle\boxed{1},\boxed{2}\rangle\text{'} \\
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{'WOMAN'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\
\text{OBL}_{mod} & \boxed{2}\begin{bmatrix}
\text{PRED} & \text{'LOOK}\langle\boxed{3}\rangle\text{'} \\
\text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{'SEVEN}\langle\boxed{3}\rangle\text{'} \\ \text{CASE} & \text{ACC} \\ \text{OBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{'DISASTER'} \\ \text{CASE} & \text{GEN} \end{bmatrix} \end{bmatrix} \\
\text{COMP-FORM} & \text{LIKE}
\end{bmatrix}
\end{bmatrix}
$$

---

[13]Since ≡ is used in (52) instead of →, DEP is a metacategory – the left-hand side of this rule, DEP, does not appear in the c-structure, the relevant right-hand side category is used instead. While this is equivalent to having the right-hand side of (52) instead of DEP in (51), it makes rules more readable and less complicated.

[14]The OBL$_{mod}$ grammatical function is used in Polish for arguments expressing manner.

This is precisely the problem with alternative representations of (10) given in (12) and (54): in both cases the dependent expressing manner (analysed as an adjunct or an argument, respectively) is only present in the main clause – realised as a comparison (*jak siedem nieszczęść*).[15] However, there is no such dependent in the comparison itself, though it has the same predicate as the main clause and it has the same requirements (presence of a manner dependent, be it an argument or an adjunct).

There seems to be a potential solution to this problem, though it involves changing the treatment of JAK – so far it has been analysed as a complementiser. Such comparative constructions could instead be analysed as relative clauses, where the entire clause is the manner dependent of the main verb, while JAK is the relative element inside the relative clause, filling its manner dependent position, see (55). Alternatively, a free relative-like analysis could be adopted, where JAK is at the same time a dependent of the main clause predicate and a dependent of the comparative clause – see (56), inspired by the account of free relatives in Dalrymple and Kaplan 2000. Under both variants of the relative clause solution the requirements of both predicates taking a manner dependent are satisfied.

(55)
$$
\begin{bmatrix}
\text{PRED} & \text{'LOOK}\langle\boxed{1},\boxed{2}\rangle\text{'} \\[2pt]
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{'WOMAN'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\[10pt]
\text{OBL}_{mod} & \boxed{2}\begin{bmatrix}
\text{PRED} & \text{'LOOK}\langle\boxed{3},\boxed{4}\rangle\text{'} \\[2pt]
\text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{'SEVEN}\langle\boxed{5}\rangle\text{'} \\ \text{CASE} & \text{ACC} \\ \text{OBJ} & \boxed{5}\begin{bmatrix} \text{PRED} & \text{'DISASTER'} \\ \text{CASE} & \text{GEN} \end{bmatrix} \end{bmatrix} \\[10pt]
\text{OBL}_{mod} & \boxed{4}\begin{bmatrix} \text{PRED} & \text{'JAK'} \end{bmatrix} \\[4pt]
\text{CLAUSE-TYPE} & \text{REL}
\end{bmatrix}
\end{bmatrix}
$$

(56)
$$
\begin{bmatrix}
\text{PRED} & \text{'LOOK}\langle\boxed{1},\boxed{2}\rangle\text{'} \\[2pt]
\text{SUBJ} & \boxed{1}\begin{bmatrix} \text{PRED} & \text{'WOMAN'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\[10pt]
\text{OBL}_{mod} & \boxed{2}\begin{bmatrix}
\text{PRED} & \text{'JAK'} \\[2pt]
\text{ADJUNCT} & \left\{\begin{bmatrix}
\text{PRED} & \text{'LOOK}\langle\boxed{3},\boxed{2}\rangle\text{'} \\[2pt]
\text{SUBJ} & \boxed{3}\begin{bmatrix} \text{PRED} & \text{'SEVEN}\langle\boxed{4}\rangle\text{'} \\ \text{CASE} & \text{ACC} \\ \text{OBJ} & \boxed{4}\begin{bmatrix} \text{PRED} & \text{'DISASTER'} \\ \text{CASE} & \text{GEN} \end{bmatrix} \end{bmatrix} \\[10pt]
\text{OBL}_{mod} & \boxed{2} \\[2pt]
\text{CLAUSE-TYPE} & \text{REL}
\end{bmatrix}\right\}
\end{bmatrix}
\end{bmatrix}
$$

---

6 Conclusion

This paper presented an analysis of Polish data involving lexicalised comparative constructions with JAK. It presented arguments supporting a unified clausal, gapping-like analysis of lexicalised comparative dependents of verbs based on attested data from Polish. On the basis of the discussion of possible mismatches between the main clause and the comparison, it argued that a standard gapping analysis where the gapped comparative clause uses the same predicate and imposes the same requirements is not capable of accounting for all the data presented in this paper. For this reason it proposes a gapping-inspired formalisation of the analysis of gapped comparisons, which relies on the use of appropriate constraints in the lexicon. Adopting such an approach makes it possible to account for mismatches that could not be handled by a standard gapping analysis or would be problematic. The formalisation of the analysis proposed in this paper was successfully implemented and tested in XLE as a part of a general grammar of Polish.

However, it must be noted that the examples of mismatch between the main verb and the comparison are very rare. This paper is based on the analysis of 147 valency schemata[16] in which the comparison involves more than one lexicalised dependent. In the vast majority of inspected schemata there is a perfect match between the dependents of the main verb and the comparison – it might therefore be worth exploring the possibility of adapting the analysis of gapping described in Patejuk and Przepiórkowski 2017 to such comparisons, hoping for a unified account of gapping (perhaps leaving out the extremely rare genuine mismatches). While the current paper does not investigate instances of non-lexicalised comparisons with JAK (which are very productive), these could be taken into account under such a unified account.

Furthermore, it seems that the current account of lexicalised comparisons could be extended so as to use Optimality Theory marks to account for instances of word play where some of the constraints imposed on the lexicalised comparison are deliberately violated – it seems that such violations do not simply result in ungrammaticality, which would call for a more refined solution.

Finally, though Polish abounds in other types of comparative constructions, for instance featuring NICZYM 'like' (which is sometimes treated as a high-register variant of JAK) or NIŻ 'than', these are outside of the scope of this paper – it is focused exclusively on lexicalised comparisons involving JAK.

References

Asudeh, Ash and Toivonen, Ida. 2006. Expletives and the syntax and semantics of copy raising. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG'06 Conference*, University of Konstanz, Germany: CSLI Publications.

---

[16]The schemata were taken from Walenty (Przepiórkowski et al. 2014), a valency dictionary of Polish. It is freely available online at http://walenty.ipipan.waw.pl/. At the time of writing, the schemata containing comparisons are not analysed in Walenty as described in this paper (no grammatical functions such as SUBJ or OBJ, no mapping between corresponding dependents).

Butt, Miriam and King, Tracy Holloway (eds.). 2014. *The Proceedings of the LFG'14 Conference*, Stanford, CA, CSLI Publications.

Butt, Miriam, Niño, María-Eugenia and Segond, Frédérique. 1999. *A Grammar Writer's Cookbook*. Stanford, CA: CSLI Publications.

Dalrymple, Mary and Kaplan, Ronald M. 2000. Feature Indeterminacy and Feature Resolution. *Language* 76(4), 759–798.

Kallas, Krystyna. 1986. Syntaktyczna charakterystyka wielofunkcyjnego *JAK*. *Polonica* XII, 127–143.

Patejuk, Agnieszka and Przepiórkowski, Adam. 2012. A Comprehensive Analysis of Constituent Coordination for Grammar Engineering. In *Proceedings of the 24rd International Conference on Computational Linguistics (COLING 2012)*.

Patejuk, Agnieszka and Przepiórkowski, Adam. 2014a. In favour of the raising analysis of passivisation. In Butt and King (2014), pages 461–481.

Patejuk, Agnieszka and Przepiórkowski, Adam. 2014b. Structural Case Assignment to Objects in Polish. In Butt and King (2014), pages 429–447.

Patejuk, Agnieszka and Przepiórkowski, Adam. 2017. Filling the gap. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG'17 Conference*, Stanford, CA: 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. 2006. O inherentnej liczbie mnogiej liczebników ćwierć, pół i półtora. *Poradnik Językowy* 9, 78–87.

Przepiórkowski, Adam, Hajnicz, Elżbieta, Patejuk, Agnieszka, Woliński, Marcin, Skwarski, Filip and Świdziński, Marek. 2014. Walenty: Towards a comprehensive valence dictionary of Polish. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk and Stelios Piperidis (eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation, LREC 2014*, pages 2785–2792, ELRA, Reykjavík, Iceland.

Przepiórkowski, Adam and Patejuk, Agnieszka. 2012. The puzzle of case agreement between numeral phrases and predicative adjectives in Polish. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG'12 Conference*, pages 490–502, Stanford, CA: CSLI Publications.