Coordination of unlike grammatical functions

Agnieszka Patejuk[1,2]   **Adam Przepiórkowski**[1,3,4]
[1]Institute of Computer Science, Polish Academy of Sciences
[2]Faculty of Linguistics, Philology and Phonetics, University of Oxford
[3]Institute of Philosophy, University of Warsaw
[4]Wolfson College, University of Oxford
aep@ipipan.waw.pl   adamp@ipipan.waw.pl

Abstract

The aim of this paper is to propose a dependency analysis of coordination of unlike grammatical functions, as witnessed in Slavic and some neighbouring languages (including Romanian, Hungarian and West Armenian). In order to increase the practical impact of the analysis, the proposed representations adhere to Universal Dependencies, a syntactic corpus annotation scheme, though arguments are given for validity of such representations from the theoretical linguistic perspective.

1 Introduction

Coordination is a well-known and long-standing problem for dependency representations of natural language utterances, both in theoretical linguistics and in natural language processing. Representational devices beyond the usual dependency trees are proposed especially for the treatment of coordination in Lucien Tesnière's Dependency Syntax (1959, 2015), Richard Hudson's Word Grammar (1984, 1990, 2010), and Igor Mel'čuk's Meaning–Text Theory (1974, 1988, 2009). Also, the representation of coordination differs widely in different dependency corpora (Popel et al., 2013).

Coordination is also problematic for Universal Dependencies (UD; Nivre et al., 2016; http://universaldependencies.org/). In the current version 2 of the standard, each utterance may be represented by two dependency structures: the basic dependency tree and the enhanced representation, which does not have to be a tree. For example, the two representations of (1) (on one of its interpretations) are shown in (2).[1,2]

(1)   I wanted to buy fresh apples and oranges.

(2)
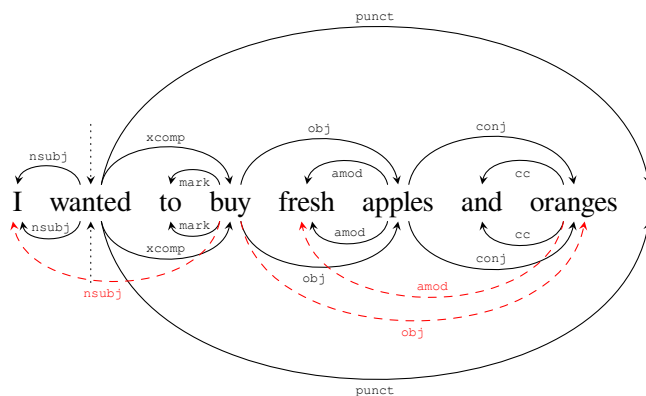
As is clear from the basic dependency tree (above the tokens), coordination is represented in UD as headed by the first conjunct, as in Mel'čuk's Meaning–Text Theory (MTT), but – unlike in that theory – all non-initial conjuncts are `conj` dependents of the initial conjunct, and the coordinating conjunction is a `cc` dependent of the following conjunct. This tree suffers from the usual deficiencies of dependency

---

[1]This example is based on examples given at http://universaldependencies.org/u/overview/enhanced-syntax.html. All URLs mentioned in this paper were last accessed on 1 April 2019.

[2]In drawing UD representations, the following conventions are adopted in this paper. The basic dependency tree is drawn above the word tokens and the enhanced dependency is drawn below the word tokens. Dependencies which differ between the two representations are drawn as dashed lines in red. The root is marked by a vertical dotted arrow.

trees: it does not represent the fact that *I* is not only the surface subject (`nsubj`; for nominal subject) of the matrix verb *wanted* but also the understood subject of the controlled verb *buy*, or the fact that the adjectival modifier (`amod`) *fresh* is understood here as referring to the whole coordinate structure, *apples and oranges*, rather than just to the first conjunct, *apples*. These deficiencies are corrected in the enhanced dependency structure (below the tokens), where – just as in Lexical Functional Grammar (LFG; Bresnan, 1982, Dalrymple, 2001), Head-driven Phrase Structure Grammar (HPSG; Pollard and Sag, 1987, 1994), and Hudson's Word Grammar (WG) – structure sharing in control constructions is represented explicitly, namely, by the additional `nsubj` dependency. Moreover, the additional `amod` dependency from *oranges* to *fresh* makes it clear that the adjective is shared by the two conjuncts. Finally, the additional `obj` (direct object) dependency from *buy* to *oranges* emphasises the symmetric nature of the two conjuncts with respect to the governing verb *buy*.

One problematic aspect of this representation of coordination, known to the UD community, concerns nested – i.e. immediately embedded – coordination: in the case of three conjuncts, A, B, C, the proposed representation does not distinguish between the flat structure (A, B, C), and the structure in which A and B are conjoined and the resulting coordination is conjoined with C, i.e., $((A, B), C)$.[3] Solutions to this problem are discussed in Przepiórkowski and Patejuk, 2019b. In this paper we deal with another phenomenon problematic for UD, namely, the possibility to coordinate different grammatical functions, as in the attested (3):[4]

(3)   [[What]$_{obj}$ and [when]$_{advmod}$] to *eat* to reduce insulin[5]

Such examples violate the overwhelming generalisation that only the same grammatical functions may be coordinated. Normally, languages satisfy this generalisation and attempts to coordinate phrases bearing different grammatical functions result in unacceptability, as in (4)–(5):

(4)   *I and an apple have already eaten.
      (intended meaning: I have already eaten an apple.)
(5)   *I have already eaten an apple and today.
      (intended meaning: I have already eaten an apple today.)

The utterance (4) is unacceptable as it involves a coordination of a subject, *I*, and a direct object, *an apple*. Similarly, (5) involves a coordination of a direct object, *an apple*, and an adjunct, *today*.

The assumption that all conjuncts must bear the same relation to the external head is also explicitly made in dependency grammars, e.g. (Hudson, 1984, 225):

> [W]e need to make sure that, in some sense, all the conjuncts in a coordinate structure have the same external relations… If we mix up conflicting external relations, the result is zeugma (e.g. *He came in {(a hurry) and (a taxi)}*, where the conjuncts require conflicting meanings of *in*), or sheer incoherence (e.g. *I ate potatoes and in the kitchen*).

This is also implicitly assumed in constituency- and constraint-based approaches, e.g., in LFG, where – in the f(unctional)-structure – the whole coordinate structure is the value of an attribute such as SUBJ(ect) or OBJ(ect), or belongs to the ADJ(uncts) set, i.e. where all conjuncts bear the same grammatical function.

There are, however, two classes of exceptions – both empirically constrained – to the generalisation that only the same grammatical functions may be coordinated. The first, sylleptic zeugma, is mentioned in the above quote from Hudson, 1984, 225. Such constructions, in which the two conjuncts evoke two different meanings of the head, have a metalinguistic feel and they are easy to distinguish from genuine coordination. We will not deal with zeugma here. The rest of this paper is devoted to the second class of exceptions, illustrated with the English (3). As discussed in §2, such constructions are robust especially in Slavic and they do not evoke different meanings of the head. §3 examines previous dependency approaches to such constructions, while §4 proposes a UD representation. Finally, §5 concludes the paper.

---

[3]`http://universaldependencies.org/u/dep/conj.html#nested-coordination`
[4]This phenomenon should be carefully distinguished from the coordination of unlike grammatical *categories*, relatively uncontroversial in contemporary linguistics (cf., e.g., Sag et al., 1985, Bayer, 1996, Dalrymple, 2017, but also Bruening and Al Khalaf, 2019 for dissent).
[5]`https://www.dietdoctor.com/what-and-when-to-eat-to-reduce-insulin`

2 Lexico-Semantic Coordination

The phenomenon in question is the so-called lexico-semantic coordination (Sannikov, 1979, 1980), also known as hybrid coordination (Chaves and Paperno, 2007). It occurs mainly in Slavic (incl. Bulgarian, Croatian, Polish and Russian) and in some neighbouring languages (Romanian, Hungarian, West Armenian), as well as – though significantly constrained – in English, French, German, Dutch, Italian and Spanish (Paperno, 2012, Lipták, 2012, Bîlbîie and Gazdik, 2012). In the case of these Germanic and Romance languages, the phenomenon seems to be limited to the coordination of optional *wh*-items (Gračanin-Yüksek, 2007, Lipták, 2012) – e.g., an adjunct and an optional argument – and often occurs in titles, as in (3) above. In the case of the "Slavic sprachbund", the phenomenon is much more robust.

First of all, as discussed in Patejuk and Przepiórkowski, 2012a,b and in Paperno, 2012, in Slavic such constructions are not limited to *wh*-items, although they particularly often involve such items; most of the examples in this paper are of this kind. But apart from *wh*-items, coordination of different grammatical functions may involve negative pronouns (so-called *n*-words; cf., e.g., the Russian (11) and the Polish (16)), certain items expressing existential or universal quantifiers (the latter illustrated in (6) below), and items belonging to a number of other pronominal or quantificational classes.[6] Second, the coordinated items may be obligatory arguments, as in (6) from the National Corpus of Polish (NKJP; Przepiórkowski et al., 2011, 2012), cited here after Patejuk and Przepiórkowski, 2012b, 463.[7]

(6) *Obiecać*   można [[wszystko]$_{obj}$ i   [wszystkim]$_{iobj}$].   (Polish)
   promise.INF may   everything.ACC and everyone.DAT
   'One may promise everything to everyone.'   (NKJP)

Hence, lexico-semantic coordination cannot easily be analysed via 'conjunction reduction', i.e., as some kind of ellipsis; arguments against such an analysis are reviewed in Patejuk, 2015, §5.4.[8] Third, such constructions are textually frequent and often occur in carefully edited texts; there is nothing marginal about them in the languages in which they occur.

Given that lexico-semantic coordination violates the 'same grammatical function' generalisation, it might seem that perhaps it does not involve coordination at all, i.e. that *i* 'and' in (6) and *and* in (3) are not really conjunctions here, but some homophonous elements of a different grammatical class. There are strong arguments against this view. First, in all languages which allow for joining different grammatical functions the joining element has the same form as a conjunction; on the homonymy view, this perfect synchronous correlation is somewhat surprising (even if it may be justified diachronically). Second, and more importantly, as shown in Patejuk and Przepiórkowski, 2012b and in Patejuk, 2015, other typical conjunctions may also occur in such constructions – for instance *lub* 'or', see (7), and *ani* 'nor', see (8).[9] It is worth noting that the conjunction *ani* has a special property – it is an *n*-word, so it requires negation. It retains this property when it combines unlike grammatical functions. As a consequence, the hypothesis that *ani* in (8) is not a conjunction seems like a typical missed generalisation: it combines two items just like conjunctions do, it has the same form as a conjunction, it also has the same properties with respect to negation as that very conjunction. Finally, lexico-semantic coordination may also occur with preconjunctions ('both… and…', 'not only… but also…'), as shown in (9), and it is possible to coordinate more than two items, see (10).

(7) …kto   lub czego   będzie w Wikipedii szukał.   (Polish)
   who.NOM or   what.GEN will   in Wikipedia seek
   '…who will seek what in Wikipedia.'   (NKJP)

(8) Rząd   USA *(nie) ujawnia, kogo   ani dlaczego umieścił na liście osób…   (Polish)
   government USA   NEG discloses who.ACC and why   put   on list   people
   'The US government does not disclose who and why they put on the list of people…'[10]

---

[6]See Przepiórkowski and Patejuk, 2014 and, especially, Patejuk, 2015, §5.8 for a comprehensive list of such classes in Polish.
[7]The labels `obj` and `iobj` reflect how this example would be annotated in Polish UD treebanks; cf. Patejuk and Przepiórkowski, 2018.
[8]For this reason we do not refer in this paper to dependency analyses of gapping, non-constituent coordination and the like.
[9]The asterisk before brackets in (8) means that the sentence is ungrammatical if the bracketed material is omitted.
[10]`http://wyborcza.pl/1,76842,15826586,Amerykanie_maja_tajna_liste__nielotow___Trafisz_na.html`

(9) …kiedy wyjawisz nie tylko kto,   ale i   dlaczego otrzymał awans.   (Polish)
      when disclose   not only   who.NOM but and why   received   promotion
      '…when you explain not only who, but also why got promoted.'[11]

(10) Kto,   kiedy i   dla kogo   napisał te   wiersze?   (Polish)
      who.NOM when and for who.GEN wrote   these poems
      'Who, when and for whom wrote those poems?'[12]

Hence, the combining words should be analysed as true conjunctions, and the phenomenon in question – as true coordination.

## 3 Theoretical Dependency Approaches

While coordination of unlike grammatical functions has attracted some attention in various linguistic theories, including Transformational Grammar (e.g. Lipták, 2012, Citko and Gračanin-Yüksek, 2013), Categorial Grammar (Paperno, 2012), HPSG (Chaves and Paperno, 2007, Bîlbîie and Gazdik, 2012) and LFG (Gazdik, 2010, Patejuk and Przepiórkowski, 2012a,b, Patejuk, 2015), to the best of our knowledge, it has not been given a serious analysis in dependency theories. Mel'čuk referred to lexico-semantic coordination in a couple of works, but only in passing: in an endnote in Mel'čuk, 1988 (and in Mel'čuk and Pertsov, 1987), then in the main text in Mel'čuk, 2009. Mel'čuk, 1988, 40 states that "There are more complicated cases of double dependency challenging the adequacy of D[ependency]-trees", illustrating this using example (11) together with representation in (12), both based on similar examples in Sannikov, 1979:

(11) Nikto   i   nikomu   ne pomogaet.   (Russian)
     nobody.NOM and nobody.DAT not helps
     'Nobody helps anybody.'

(12)
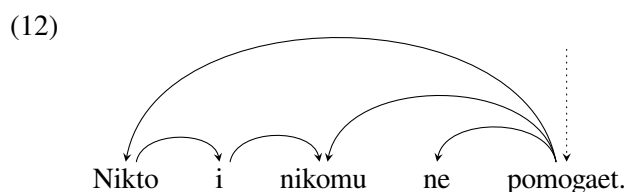
Mel'čuk, 1988 distances himself from this representation, claiming that "*nikomu* does not depend on *pomogaet* syntactically". Mel'čuk, 2009, 81 returns to lexico-semantic coordination and briefly considers the Russian sentence (13), which would receive a representation like (14) within his Meaning–Text Theory:

(13) Kto,   komu i   čem   pomog?   (Russian)
     who.NOM who.DAT and what.INS helped
   'Who helped whom with what?'

(14)
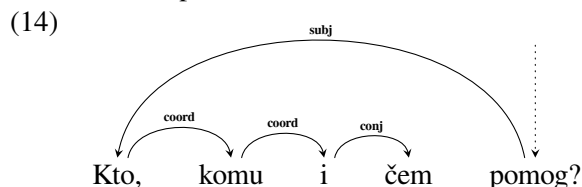
Noting that this representation loses information about grammatical functions of non-initial conjuncts, he proposes to "introduce some special [dependency relations] just for this very special construction: **coord-subj**, **coord-dir-obj**, **coord-indir-obj**, etc." So, presumably, (13) should be represented as (15).[13]

---

[11]https://www.hbrp.pl/b/dobry-system-awansow-pracownikow-to-awans-dla-calej-firmy/P1C5cBsu1

[12]Danielewiczowa, 1996, 85

[13]It is not clear to us that the two relations assigned to *komu* 'who.DAT' and *čem* 'what.INS' in Meaning–Text Theory are in fact indirect object and oblique object, as indicated in (15), but this does not matter for the argument in the main text.

(15)

$$\text{Kto,} \quad \text{komu} \quad \text{i} \quad \text{čem} \quad \text{pomog?}$$
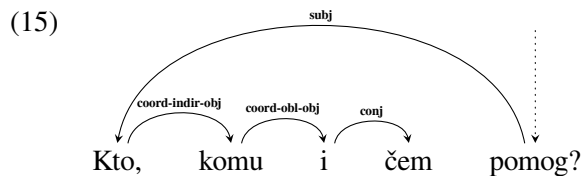
Apart from the problem of duplicating many syntactic relations as **coord**-relations, this suggestion is based on the assumption that all conjuncts in lexico-semantic coordination must be dependents of the same head. As demonstrated in Patejuk and Przepiórkowski, 2012b and Patejuk, 2015 on the basis of numerous examples such as the following, this assumption is false:

(16) Nic i nikogo nie może tłumaczyć. (Polish)
nothing.NOM and nobody.GEN NEG can excuse.INF
'Nothing may excuse anybody.' (NKJP)

(17) Czego i ile trzeba dostarczyć organizmowi? (Polish)
what.GEN and how much.ACC should provide.INF organism.DAT
'What – and how much – should one provide one's organism with?'[14]

(18) Jakie i kto może ponieść konsekwencje? (Polish)
what.ADJ.ACC and who.NOM can bear.INF consequences.ACC
'Who can suffer what consequences?'[15]

In (16), *nic* 'nothing.NOM' is the subject of the matrix verb *może* 'may', as well as the understood subject of the controlled verb *tłumaczyć* 'excuse', while *nikogo* 'nobody.GEN' is the direct object of the controlled verb (only). In (17), adopting the common and well-founded assumption in Polish structural and formal linguistics (e.g., Saloni and Świdziński, 2001) that numerals – not nouns – are heads of numeral phrases, the interrogative numeral *ile* 'how much.ACC' is the direct object of *dostarczyć* 'provide, supply', while *czego* 'what.GEN' is a dependent of this numeral.[16] Finally, in (18), *jakie* 'what.ACC' is the adjectival modifier of *konsekwencje* 'consequences', which is the object of *ponieść* 'suffer', which in turn is the infinitival complement of the main verb, *może* 'may', whose subject is *kto* 'who.NOM'.

On Mel'čuk's proposal, the **coord** dependency between conjuncts in (16)–(18) would not only have to encode the grammatical function (**coord-subj**, **coord-dir-obj**, **coord-indir-obj**, etc.), but also information about the actual head of each non-initial conjunct. It is not clear how this information could be encoded within a constrained set of dependency relations (52 in Mel'čuk, 2009). A potential solution, to be considered in more detail below, would be to devise a special labelling convention to account for this phenomenon, where – for each non-initial conjunct – a single dependency label would encode the entire dependency chain to this conjunct. However, this would result in a potentially infinite number of dependency labels.

4 Lexico-Semantic Coordination in UD

4.1 Proposal

How can coordination of unlike grammatical functions be represented in UD? Let us start with the simple (but attested) Polish example (19), where two different dependents of the same head are coordinated: the subject (*kto* 'who.NOM') and the object (*kogo* 'whom.ACC') of the verb *zdradził* 'betrayed'. Since there is no discussion of how to annotate lexico-semantic coordination in UD guidelines, the representation in (20) follows general guidelines related to coordination:
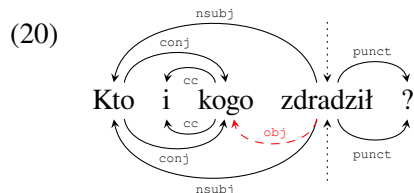
---

[14]https://vitalia.pl/forum22,446761,0_Czego-i-ile-trzeba-dostarczyc-organizmowi.html
[15]Patejuk, 2015, 99
[16]As discussed in §4.3, in UD the relation between the numeral and the noun is reversed, but the two conjuncts in (17) are still dependents of different heads.

(19) Kto   i   kogo   zdradził?   (Polish)
who.NOM and who.ACC betrayed
'Who betrayed whom?'[17]

(20)

In the basic dependency tree (edges above words), the first conjunct, *kto*, is the subject of *zdradził*, but the information that *kogo* is the object of *zdradził* is lost. This is because *kogo* is annotated as the second conjunct using the `conj` relation – since the basic representation must be a tree, there must not be another incoming relation (object). In effect, the coordination *kto i kogo* 'who.NOM and whom.ACC' is annotated as if it were the subject – which is not true in the case of the second conjunct – and there is no information that *zdradził* has an object. This problem is mitigated in the enhanced dependency representation (edges below words), which lacks such a restriction – the object dependency from *kogo* to *zdradził* is present in the graph, which shows that it is not its subject (despite the basic representation). As a result, appropriate grammatical functions are only provided in the enhanced dependency representation.

Note that it does not seem appropriate to think about the basic representation in (20) as an 'underspecified' version of the more detailed enhanced structure. On such an 'underspecification' view, the basic representation of ordinary coordinated subjects, as in *John and Mary arrived*, would also have to be considered 'underspecified', with the grammatical functions of non-initial conjuncts (here: *Mary*) to be ascertained only upon careful inspection of the enhanced representation. This view is not only questionable conceptually, but also untenable practically: as popular dependency parsers only deal with basic dependency trees – and are unable to learn from or parse with enhanced dependency representations – the lossy and misleading information about grammatical functions at the basic tree level translates into errors in downstream applications, especially those which rely on grammatical functions to extract information about who did what to whom.[18]

For these – conceptual and practical – reasons we propose the alternative basic UD representation in (21); while the enhanced dependency graph is the same as in (20), the basic dependency tree does not include the `conj` dependency between the two conjuncts, i.e. coordination is not fully represented at this level, but the much more important information about grammatical functions is: in (21) *kogo* is identified as the object of *zdradził* at both levels of dependency representation.

(21)
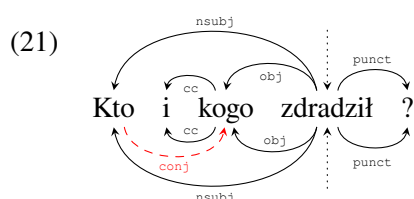
The more theoretical reason for preferring (21) to (20) as a representation of (19) is that coordination plays here a very different role than usual: it does not conjoin phrases which stand in the same syntactic and semantic relation to the head, but rather it joins elements which only have the same information structure status in the sentence. This difference between standard coordination and the coordination of unlike grammatical functions discussed in this section, since it is crucial for the syntactic and semantic interpretation of the sentence, should be represented in the basic dependency tree.

4.2 Potential Alternatives

Let us consider a potential alternative solution, inspired by the approach outlined in Mel'čuk, 2009 (see the discussion of (14)), which is aimed at saving the topology of the basic UD representation in (20) by enriching

---

the dependency label from the head to the coordinate structure so that it correctly represents grammatical functions of all conjuncts (instead of suggesting that the entire coordination is the subject), e.g.:

(22)

Kto i kogo zdradził ?
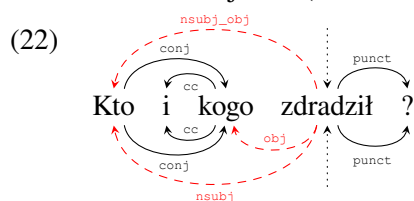
This solution suffers from the same problems as that proposed by Mel'čuk (2009): it greatly multiplies dependency labels (many different numbers and orders of grammatical functions[19] would have to be encoded) and it does not encode information about possibly different heads of particular conjuncts. Consider again (16), repeated below as (23):

(23) Nic   i   nikogo   nie może tłumaczyć.   (Polish)
     nothing.NOM and nobody.GEN NEG can   excuse.INF
   'Nothing may excuse anybody.'

On our proposal, its representation is given in (24) – coordination is not fully represented in the basic tree, but grammatical functions are:

(24)

Nic i nikogo nie może tłumaczyć .
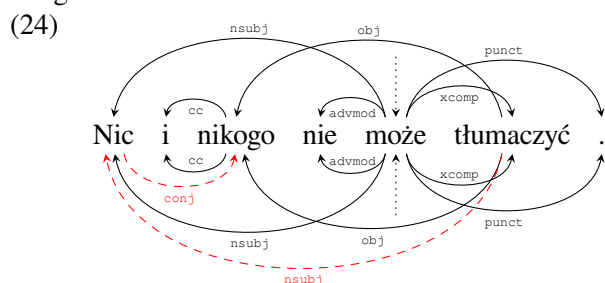
In contrast, the attempt to save the more standard basic representation (the one following from general guidelines) by labelling the dependency from *może* 'can, may' to *nic* 'nothing.NOM' as nsubj_obj (instead of nsubj) in (25) is misinformative, as it incorrectly suggests that *nikogo* 'nobody.GEN' is the direct object of *może* – rather than the object of *tłumaczyć* 'explain'.

(25)

Nic i nikogo nie może tłumaczyć .
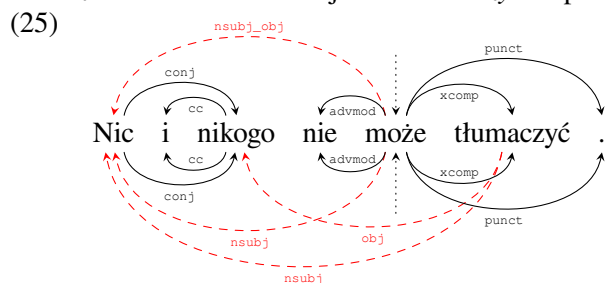
This problem could be approached in a way analogous to the earlier suggestion on how to modify Mel'čuk's account, namely, by providing – in the basic tree – full paths to non-initial conjuncts, as shown in (26), where the relation targeting *nic* is nsubj_xcomp:obj (because *nikogo* is the obj of xcomp, see (24)):

(26)

Nic i nikogo nie może tłumaczyć .

---
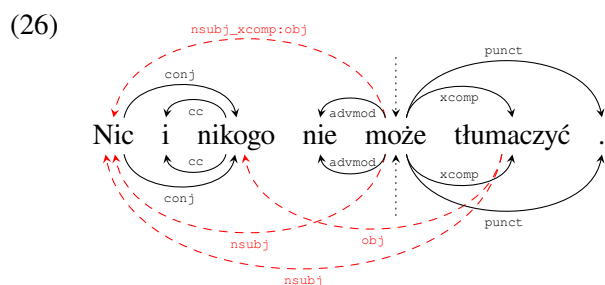
[19]For instance, assuming there are only instances of coordination with 2 conjuncts, each of which has a different grammatical function, this yields $n \times (n-1)$ labels, where n is the number of basic grammatical functions. For 3 conjuncts, there would be $n \times (n-1) \times (n-2)$ labels, and so on.

However, it is clear that such a solution involving full dependency paths as (parts of) dependency labels would further aggravate the issue of the number and complexity of dependency labels.[20] Moreover, in some cases such dependency paths may still be insufficient, e.g. in the case of a predicate with two or more `obl` dependents such that a dependent of one them participates in lexico-semantic coordination, as in (27):

(27) Kto   i   jakiej   bał   się napaści   tamtej   nocy?   (Polish)
     who.NOM.M and what.ADJ.GEN.F feared.3.SG.M RM assault.GEN.F that.GEN.F night.GEN.F
     'Who feared what assault on that night?'

Here, both *napaści* 'assault' and *tamtej nocy* 'that night' are genitive obliques, so the `obl` part of the hypothetical `nsubj_obl:det` dependency from the root verb *bał (się)* 'feared' to *kto* 'who' – the hypothetical head of the lexico-semantic coordinate structure *kto i jakiej* 'who.NOM.M and what.ADJ.GEN.F' – is ambiguous between these two oblique dependents, as shown in (28). Moreover, agreement facts do not help in resolving this ambiguity, because both obliques are feminine, singular, genitive – just like the adjective *jakiej*.[21]

(28)
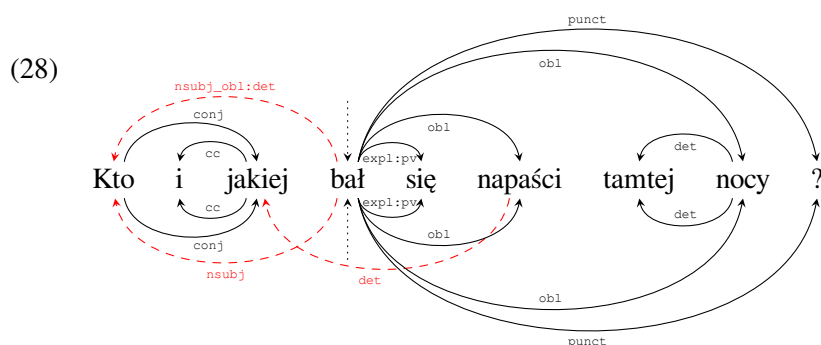
   Hence, we prefer the representation in (24) to hypothetical alternatives shown in (25) and (26) – the proposed solution ensures simple and accurate representation of grammatical functions of coordinated dependents using the standard repertoire of dependency labels, even if coordinated items depend on different heads. This is achieved at the cost of not fully representing the lexico-semantic coordination at the basic level, which is however fully represented in enhanced dependencies.

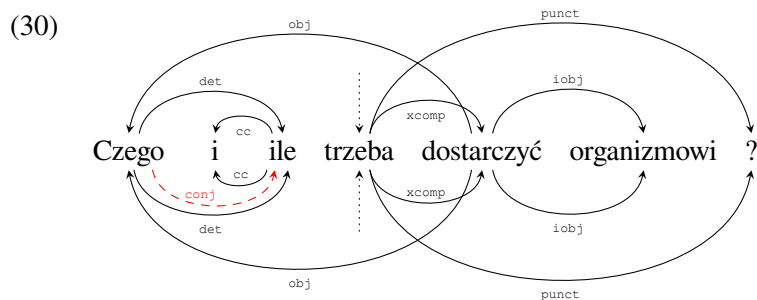### 4.3 Numeral Phrases: A Challenge for UD

Let us now return to (17), repeated below for convenience as (29), which poses an interesting additional challenge to UD:

(29) Czego   i   ile   trzeba dostarczyć organizmowi?   (Polish)
     what.GEN and how much.ACC should provide.INF organism.DAT
     'What – and how much – should one provide one's organism with?'
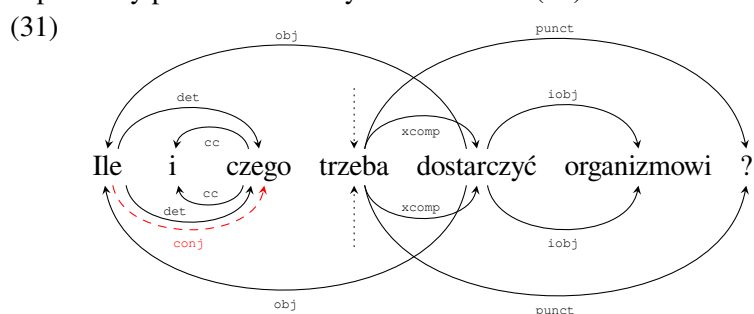
As mentioned above, in Polish – on the standard (non-UD) analysis – the numeral is the head (it receives case marking from the verb), while the accompanying noun is its dependent (it receives case from the numeral). However, following UD guidelines, this dependency relation is reversed: numerals are dependents of nominal heads, so the interrogative numeral *ile* 'how much' is a `det` dependent of *czego* 'what.GEN', which is in turn the direct object of *dostarczyć* 'provide, supply'. One potential problem with the UD representation arises at the level of enhanced dependencies, where *ile* is also a `conj` dependent of *czego*; as shown in (30), there are two different equidirectional dependency relations between these two tokens: `det` and `conj`.

---

[20]See Schuster et al., 2017, 130–131 for arguments against encoding paths in dependency labels in the context of the UD representation of gapping, the most important of which is that this would introduce an unbounded number of dependency relations.

[21]Though relations could be disambiguated by, for instance, adding indices, e.g. `obl1` and `obl2`, but this would further aggravate the problem of number and complexity of labels (resulting in `nsubj_obl1:det`, among others).

(30)

Czego i ile trzeba dostarczyć organizmowi ?

(dependency diagram with relations: punct, obj, iobj, det, xcomp, cc, cc, conj, det, xcomp, obj, iobj, punct)
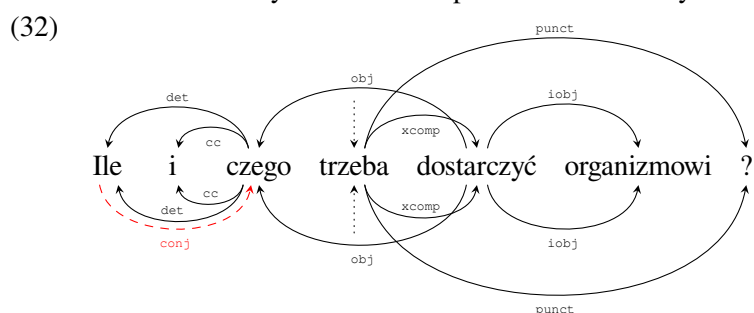
This problem arises regardless of which representation of lexico-semantic coordination – the one proposed here or the one arising from general UD guidelines – is chosen, because the enhanced representations are identical under both. Moreover, this problem is independent of the issue of headedness of Polish numeral phrases: if the UD analysis of numerals were reversed so that the numeral is the head and the noun is the dependent, the problem would resurface in examples such as (29) but with the order of conjuncts reversed (i.e. *Ile i czego trzeba dostarczyć organizmowi?* – also fully acceptable). This is illustrated in (31); the double dependency problem is exactly the same as in (30).

(31)

Ile i czego trzeba dostarczyć organizmowi ?

(dependency diagram with relations: obj, punct, det, iobj, cc, xcomp, cc, det, conj, xcomp, obj, iobj, punct)

Rather, the problem stems directly from the UD representation of coordination, which requires that the second conjunct is a `conj` dependent of the first conjunct: if the second conjunct is independently a non-`conj` dependent of the first conjunct, the double dependency inevitably arises. Note that this is not a fundamental problem, as – while unmet in UD so far[22] – such double dependencies in the enhanced representation do not seem to violate any deep UD principles, and they could be constrained by well-formedness conditions specifying which dependencies may co-occur this way.

Unfortunately, such constructions also present a more fundamental problem for the standard UD approach to coordination. Let us consider the representation of the sentence from (31) under the current proposal with the usual UD analysis of numeral phrases as headed by nouns:

(32)

Ile i czego trzeba dostarczyć organizmowi ?

(dependency diagram with relations: punct, det, obj, iobj, cc, xcomp, cc, det, xcomp, conj, obj, iobj, punct)

Unlike in (30), there is no problem of two equidirectional relations in (32), but the `det` edge targeting the first conjunct, *ile* 'how much', originates in the second conjunct, *czego* 'what.GEN', whose incoming edge originates in *dostarczyć* 'provide, supply'. At the level of basic tree lexico-semantic coordination is not represented (there is no `conj` relation between conjuncts), so the tree is well-formed and does not violate any fundamental UD principles.[23]

[22]However, Przepiórkowski and Patejuk, 2019a discuss a phenomenon which also calls for lifting this ban: Polish particle *się* (sometimes called 'reflexive marker') may have more than one function with respect to the same verb, for instance inherent and impersonal, which also requires two different relations between the same tokens.

[23]The `conj` relation is present in the enhanced representation, but this representation does not seem to violate any UD rules either: coordination is headed by the first conjunct, this first conjunct has its own a head (even if it is internal to coordination), and

However, an attempt to provide this example with a basic tree including the standard UD representation of coordination fails. Once *czego* is the `conj` dependent of *ile*, the dependency from the verb *dostarczyć* must target *ile*. But, if so, what should be the dependency label on this relation targeting the interrogative numeral? It cannot be `det`, as *ile* is a `det` dependent of *czego*, not of *dostarczyć*; verbs are not supposed to have `det` dependents at all. But it cannot be `obj` either, as this would mean that *ile* stands in the immediate `obj` relation to the verb, and *czego* perhaps does too, depending on the enhanced representation. As the enhanced representation does contain the secondary `obj` dependency from *dostarczyć* to *czego*, the `obj` label in the basic tree would in effect mean that the coordinate structure involves two conjuncts standing in the same `obj` relation to the verb, fully analogous to, for instance, *(Homer) likes donuts and burgers*. This problem is illustrated below.

(33)
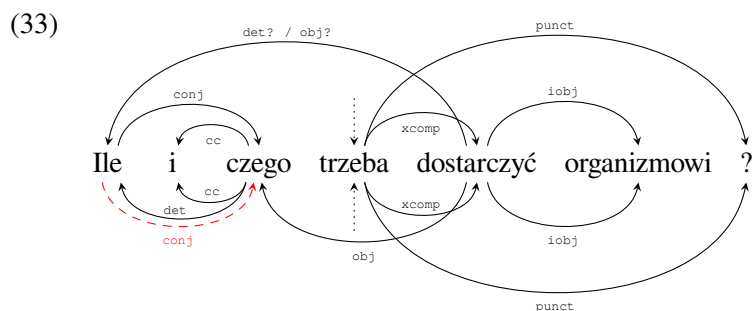
So, while the enhanced representation in (33) is the same as in (32), there seems to be no good solution for the basic representation in (33), which assumes that lexico-semantic coordination is represented in the basic tree. Once this assumption is given up, a reasonable representation becomes available, namely, the representation (32) advocated in this paper.

5   Conclusion

In this paper, we presented constructions which violate the principle that only the same grammatical functions may be coordinated – to the best of our knowledge, this is the first comprehensive discussion of such constructions within any dependency framework.

We showed that the few existing suggestions of how lexico-semantic coordination may be analysed in dependency approaches cannot account for the complex data without running into serious problems. Instead, we proposed a UD analysis of such constructions, which represents the vital information about grammatical functions of particular conjuncts on both levels: in basic trees and in enhanced representations. A feature of this representation is that lexico-semantic coordination is fully represented only at the enhanced level, which makes it possible to precisely specify different grammatical functions at the basic level. While non-standard, we believe that – given the very non-standard nature of coordination of unlike grammatical functions – this is an advantage of the proposed representation.

We also demonstrated that the phenomenon of lexico-semantic coordination necessitates giving up the assumption that there is at most one dependency relation from one token to another. (However, on our proposal, such double dependencies only occur at the enhanced level of representation, so they do not violate any deep UD principles.)

We hope that the above considerations will inspire other dependency work on the fascinating topic of coordination of unlike grammatical functions.

Acknowledgements

---

the `obj` edge to the second conjunct is secondary in a sense (it does not indicate the head of coordination).

References

Samuel Bayer. 1996. The coordination of unlike categories. *Language* 72(3):579–616.

Joan Bresnan, editor. 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: The MIT Press.

Benjamin Bruening and Eman Al Khalaf. 2019. Category mismatches in coordination revisited. *Linguistic Inquiry* Forthcoming.

Gabriela Bîlbîie and Anna Gazdik. 2012. Wh-coordination in Hungarian and Romanian multiple questions. *Empirical Issues in Syntax and Semantics* 9:19–36.

Rui Pedro Chaves and Denis Paperno. 2007. On the Russian hybrid coordination construction. In Stefan Müller, editor, *Proceedings of the HPSG 2007 Conference*, pages 46–64, Stanford, CA: CSLI Publications.

Barbara Citko and Martina Gračanin-Yüksek. 2013. Towards a new typology of coordinated *wh*-questions. *Journal of Linguistics* 49:1–32.

Mary Dalrymple. 2001. *Lexical Functional Grammar*. San Diego, CA: Academic Press.

Mary Dalrymple. 2017. Unlike phrase structure category coordination. In Victoria Rosén and Koenraad De Smedt, editors, *The Very Model of a Modern Linguist*, volume 8 of *Bergen Language and Linguistics Studies*, pages 33–55, Bergen: University of Bergen Library.

Magdalena Danielewiczowa. 1996. *O znaczeniu zdań pytajnych w języku polskim. Charakterystyka struktury tematyczno-rematycznej wypowiedzeń interrogatywnych*. Warsaw: Wydawnictwa Uniwersytetu Warszawskiego.

Anna Gazdik. 2010. Multiple questions in French and Hungarian: An LFG account. In Miriam Butt and Tracy Holloway King, editors, *The Proceedings of the LFG'10 Conference*, pages 249–269, Ottawa, Canada: CSLI Publications.

Martina Gračanin-Yüksek. 2007. *About Sharing*. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Richard Hudson. 1984. *Word Grammar*. Oxford: Blackwell.

Richard Hudson. 1990. *English Word Grammar*. Oxford: Blackwell.

Richard Hudson. 2010. *An Introduction to Word Grammar*. Cambridge University Press.

Anikó Lipták. 2012. Strategies of wh-coordination. *Linguistic Variation* 11:149–188.

Igor Mel'čuk. 1974. *Opyt teorii lingvističeskix modelej «Smysl ⇔ Tekst»*. Moscow: Nauka.

Igor Mel'čuk. 1988. *Dependency Syntax: Theory and Practice*. Albany, NY: The SUNY Press.

Igor Mel'čuk. 2009. Dependency in natural language. In Alain Polguère and Igor Mel'čuk, editors, *Dependency in Linguistic Description*, pages 1–110, Amsterdam: John Benjamins.

Igor Mel'čuk and Nikolaj Pertsov. 1987. *Surface Syntax of English. A Formal Model within the Meaning–Text Framework*. Amsterdam: John Benjamins.

Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajič, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty, and Daniel Zeman. 2016. Universal Dependencies v1: A multilingual treebank collection. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Tenth International Conference on Language Resources and Evaluation, LREC 2016*, pages 1659–1666, ELRA, Portorož, Slovenia: European Language Resources Association (ELRA).

Denis Paperno. 2012. *Semantics and Syntax of Non-Standard Coordination*. Ph.D. Thesis, University of California, Los Angeles.

Agnieszka Patejuk. 2015. *Unlike coordination in Polish: an LFG account*. Ph.D. Thesis, Instytut Języka Polskiego PAN, Cracow.

Agnieszka Patejuk and Adam Przepiórkowski. 2012a. A comprehensive analysis of constituent coordination for grammar engineering. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012)*, pages 2191–2207, Mumbai, India.

Agnieszka Patejuk and Adam Przepiórkowski. 2012b. Lexico-semantic coordination in Polish. In Miriam Butt and Tracy Holloway King, editors, *The Proceedings of the LFG'12 Conference*, pages 461–478, Stanford, CA: CSLI Publications.

Agnieszka Patejuk and Adam Przepiórkowski. 2018. *From Lexical Functional Grammar to Enhanced Universal Dependencies: Linguistically Informed Treebanks of Polish*. Warsaw: Institute of Computer Science, Polish Academy of Sciences.

Carl Pollard and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics, Volume 1: Fundamentals*. CSLI Lecture Notes, No. 13, Stanford, CA: CSLI Publications.

Carl Pollard and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago, IL: Chicago University Press / CSLI Publications.

Martin Popel, David Mareček, Jan Štěpánek, Daniel Zeman, and Zdeněk Žabokrtský. 2013. Coordination structures in dependency treebanks. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 517–527, Sofia, Bulgaria.

Adam Przepiórkowski, Mirosław Bańko, Rafał L. Górski, and Barbara Lewandowska-Tomaszczyk, editors. 2012. *Narodowy Korpus Języka Polskiego*. Warsaw: Wydawnictwo Naukowe PWN.

Adam Przepiórkowski, Mirosław Bańko, Rafał L. Górski, Barbara Lewandowska-Tomaszczyk, Marek Łaziński, and Piotr Pęzik. 2011. National Corpus of Polish. In Zygmunt Vetulani, editor, *Proceedings of the 5th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*, pages 259–263, Poznań, Poland.

Adam Przepiórkowski and Agnieszka Patejuk. 2014. Koordynacja leksykalno-semantyczna w systemie współczesnej polszczyzny (na materiale Narodowego Korpusu Języka Polskiego). *Język Polski* XCIV(2):104–115.

Adam Przepiórkowski and Agnieszka Patejuk. 2019a. From Lexical Functional Grammar to enhanced Universal Dependencies: The UD-LFG treebank of Polish, to appear in *Language Resources and Evaluation* (published online on 4 February 2019).

Adam Przepiórkowski and Agnieszka Patejuk. 2019b. Nested coordination in Universal Dependencies. In *Proceedings of SyntaxFest 2019*.

Ivan A. Sag, Gerald Gazdar, Thomas Wasow, and Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3:117–171.

Zygmunt Saloni and Marek Świdziński. 2001. *Składnia współczesnego języka polskiego*. Warsaw: Wydawnictwo Naukowe PWN, fifth edition.

Vladimir Z. Sannikov. 1979. Sočinitel'nye i sravnitel'nye konstrukcii: ix blizost', ix sintaksičeskoe predstavlenie I. *Wiener Slawistischer Almanach* 4:413–432.

Vladimir Z. Sannikov. 1980. Sočinitel'nye i sravnitel'nye konstrukcii: ix blizost', ix sintaksičeskoe predstavlenie II. *Wiener Slawistischer Almanach* 5:211–242.

Sebastian Schuster, Matthew Lamm, and Christopher D. Manning. 2017. Gapping constructions in Universal Dependencies v2. In Marie-Catherine de Marneffe, Joakim Nivre, and Sebastian Schuster, editors, *Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017)*, pages 123–132, Association for Computational Linguistics, Gothenburg, Sweden.

Natalia Silveira and Christopher Manning. 2015. Does Universal Dependencies need a parsing representation? An investigation of English. In Eva Hajičová and Joakim Nivre, editors, *Proceedings of the Third International Conference on Dependency Linguistics (DepLing 2015)*, pages 310–319, Uppsala.

Lucien Tesnière. 1959. *Éléments de Syntaxe Structurale*. Paris: Klincksieck.

Lucien Tesnière. 2015. *Elements of Structural Syntax*. Amsterdam: John Benjamins.