LEXICO-SEMANTIC COORDINATION IN POLISH

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

Abstract

On the basis of data extracted from the largest currently available corpus of Polish, this paper discusses a variety of coordination under which, contrary to previous (mostly implicit) assumptions, particular conjuncts may correspond to distinct grammatical functions at the level of f-structure as long as they represent the same restricted semantic class (*wh*-words, *n*-words and items expressing quantifiers). Moreover, it demonstrates that dependents coordinated in this way may additionally belong to entirely different f-structures (depend on distinct heads) and it offers a formal analysis which was successfully implemented in a large XLE grammar of Polish.

1 Introduction

It was assumed for a long time that coordinated items should belong to the same c-structure category. When coordination of unlikes came to the attention of LFG, the new assumption was that it is possible to coordinate different categories but the coordinate structure bears the same grammatical function as a whole (Dalrymple and Lødrup 2000 discusses an example from Sag et al. 1985 where a nominal is coordinated with a clause and together they correspond to the object grammatical function). However, over the years it was noticed in different formalisms that, under certain circumstances, it is possible to coordinate dependents which bear different grammatical functions. This phenomenon was first discussed in Sannikov 1979, 1980 on the basis of Russian data, its existence was mentioned (though largely disregarded) in Mel'čuk 1988 and later a dependency-like analysis was provided for Polish by Kallas (1993); other analyses include Chaves and Paperno 2007 for Russian in the framework of HPSG (where this phenomenon is referred to as 'hybrid coordination') and quite recently Gazdik 2010 and 2012 in LFG for French and Hungarian.

This paper presents attested examples selected from abundant data extracted from the National Corpus of Polish (NKJP, Przepiórkowski et al. 2010, 2012; `http://nkjp.pl/`) and it shows how generalisations stemming from presented data were formalised and implemented in an LFG grammar of Polish (Patejuk and Przepiórkowski, 2012). It demonstrates, providing relevant corpus evidence, that it is possible to coordinate dependents which correspond to grammatical functions belonging to various levels in the f-structure (particular conjuncts depend on different heads), and it employs different formal representations of lexico-semantic coordination, monoclausal vs biclausal, depending on which items are involved in such coordination. Finally, it provides some discussion of controversial issues.

2 Data and generalisations

Though most often examples of lexico-semantic coordination include question words (*wh*-words), this phenomenon is by no means limited to these:

(1) czy   komukolwiek, kiedykolwiek i   do czegokolwiek przydał się
    PART anybody.DAT anytime   and for anything   come in handy
    poradnik
    guide
    'Has a(ny) guide ever come in handy to anybody for anything?'   (NKJP)

(2) Obiecać można wszystko   i   wszystkim.
    promise may   everything.ACC and everyone.DAT
    'One may promise everything to everyone.'   (NKJP)

In (1) all coordinated items contain a pronoun which expresses an existential quantifier. This is the only similarity: particular conjuncts belong to different categories (noun phrase, adverbial phrase and a prepositional phrase, respectively) and bear distinct grammatical functions: indirect object (OBJ$_\theta$), adjunct (ADJ) and oblique object (OBL), respectively. In (2) both conjuncts contain a pronoun expressing a universal quantifier: the first corresponds to the direct object (OBJ) while the other is the indirect object (OBJ$_\theta$). Unlike in the previous example, both conjuncts happen to belong to the same category (noun phrase).

It is also possible to coordinate phrases containing pronouns which belong to another semantic class, namely *n*-words:

(3) nic   i   nikogo   nie może tłumaczyć.
    nothing.NOM and nobody.GEN NEG can   excuse
    'Nothing can excuse anybody.'   (NKJP)

This example is interesting because particular conjuncts not only correspond to distinct grammatical functions but they also belong to different predicates: the first conjunct (*nic*) is the subject (SUBJ) of the main clause verb *może*,[1] while the other (*nikogo*) is the object (OBJ) of *tłumaczyć* in the embedded infinitival clause (XCOMP). There are further, more sophisticated examples of coordination where conjuncts depend on different heads:

(4) Skąd i   jakie otrzymujemy informacje?
    whence and what receive   information
    'What information and from where do we receive?'   (NKJP)

In (4) both conjuncts are modifiers though they depend on different heads: the first conjunct is an adjunct (ADJ) of the verb (*otrzymujemy*), the other modifies the verb's object (*informacje*). Furthermore, it is possible that one lexico-semantic conjunct may be the head of the other one:

(5) ile   i   czego znaleźli.
    how much.ACC and what.GEN found
    'How much, and what, did they find?'   (NKJP)

---

[1]As a result of structure sharing under raising, it is also the subject of *tłumaczyć* at the same time.

Polish numeral phrases are headed by a numeral while the accompanying nominal is analysed as its dependent.[2] In (5) the first conjunct is a numeral, analysed as the object (OBJ) of the verb (*znaleźli*), while the other is the object of the numeral (*ile*) – together they constitute a complete numeral phrase with the following f-structure representation:

(6)
$$\begin{bmatrix} \text{PRED} & \text{`HOW\_MUCH}\langle\boxed{1}\rangle\text{'} \\ \text{OBJ}\ \boxed{1} & \begin{bmatrix} \text{PRED} & \text{`WHAT'} \end{bmatrix} \end{bmatrix}$$

Conjuncts taking part in lexico-semantic coordination in (4) and (5) belong to yet another semantic class, namely *wh*-words. Let us consider one more example featuring such conjuncts:

(7)　Nie　wiadomo było, czy　　*(i)　kiedy wróci.
　　　NEG know　　was　whether and when returns
　　　'It was not clear whether and when he would return.'　　　　　(NKJP)

At first glance (7) appears similar to previous examples as all conjuncts represent the same semantic class, *wh*-words in this case: the first conjunct is a question particle (*czy*), the other is an adverb (*kiedy*). The particle is analysed as a marker (marking interrogative clauses), the other conjunct is treated as an adjunct of the verb. There is a crucial difference, though: when the conjunction (*i* 'and') is removed, (7) becomes ungrammatical, while all other examples presented so far remain grammatical even if the conjunction is deleted. It is possible, however, to use a biclausal construction as an alternative to (7), with roughly the same meaning:

(8)　Nie　wiadomo było, [czy　　wróci] i　[kiedy wróci].
　　　NEG know　　was　whether returns and when　returns
　　　'It was not clear whether he would return and when he would return.'

This suggests that the representation of sentences such as (7), where the conjunction cannot be omitted without making the utterance ungrammatical, should be biclausal, i.e., based on the coordination of two clauses headed by the same main predicate. On the other hand, the remaining sentences, where the conjunction may be dropped without affecting the grammaticality of the utterance, will be analysed as essentially monoclausal, with only one occurrence of the main predicate in the representation.

　　Before proceeding to how lexico-semantic coordination is formalised in LFG and implemented in XLE, let us briefly summarise its properties: particular conjuncts bear distinct grammatical functions (arguments, adjuncts) or bear no grammatical function at all (as in the case of *czy*, the question marker), they may also belong to different levels of f-structure, sometimes even to different clauses (biclausal constructions featuring the question marker) as long as each conjunct represents the same semantic type (pronouns expressing a universal quantifier, existential

---

[2]This is the standard analysis in Polish linguistics, e.g., in the textbook of Saloni and Świdziński (2001); see also arguments for such a structure of Polish numeral phrases in Przepiórkowski and Patejuk 2012, Section 2 (in these proceedings).

quantifier, *n*-words or *wh*-words). Finally, particular conjuncts may correspond to different categories at the level of c-structure.

3   Formalisation and implementation

Lexical entries of items of a particular semantic type bear the attribute TYPE which may take one of four values: ANY (existential quantifier; cf. (1)), ALL (universal quantifier; cf. (2)), NEG (n-word; cf. (3)) or WH (question word; cf. (5), (4) and (7)). This feature has independent motivation: it is used in the grammar for the purposes of direct and embedded questions, free relatives and for handling negative concord. Simplified lexical entries of selected *n*-words are provided below:

(9)   a.   nic   N   ($\uparrow$ PRED)='NOTHING'
                          ($\uparrow$ TYPE)= NEG

      b.   nigdy   ADV   ($\uparrow$ PRED)='NEVER'
                          ($\uparrow$ TYPE)= NEG

Using parameterised c-structure rules, such elements are rewritten to phrases whose name contains, apart from category, a parameter whose value corresponds to its semantic type (represented below as a subscript in italics):

(10) $NP_{neg} \rightarrow$ { nic | nikt }

(11) $ADVP_{neg} \rightarrow$ { nigdy | nigdzie }

Parameters make it possible to use such semantic information at the level of c-structure without resorting to checking f-structure attributes (which is considerably more costly when it comes to measuring parser performance). It is possible to use parameters to ensure that certain categories in a given rule represent the same type:

(12) $XPextr_{type} \quad \rightarrow \quad XP_{type}$
                          $(\uparrow$ XPATH GF$^+)=\downarrow$

The rule in (12) is also independently motivated as it is used for the purposes of handling extraction. Its left-hand side rewrites to a disjunction of phrases of the same type; the XP category used in (12) is in fact a metacategory; its expansion rule is provided in (13), with the definition of allowed types given in (14):

(13) $XP_{type} \equiv \{NP|PP|ADVP|AP\}_{type}$

(14) $type \equiv$ { *all* | *any* | *wh* | *neg* }

The annotation attached to XP in (12) makes it possible for dependents representing relevant semantic types to appear at the level of c-structure outside the clause containing their f-structure head. There are two important elements of this annotation: XPATH, defined in (15), provides the extraction path, while GF, defined in (16), corresponds to grammatical functions which may be assigned:

(15) XPATH ≡ XCOMP*

(16) GF ≡ {SUBJ|OBJ|OBJ$_\theta$|OBL|ADJ ∈}

Together, these allow the dependent to be extracted from infinitival clauses:

(17) uśmiecha się   nieśmiało, bo   nikogo nie   chce krępować
      smiles   REFL shyly   because nobody NEG   wants intimidate
      'She smiles shyly as she does not want to intimidate anybody.'   (NKJP)

In (17) it is *nikogo* that undergoes extraction: even though it belongs at the level of c-structure to the main clause (with the verb *chce*), it is an argument of the embedded infinitival clause headed by *krępować*. Sometimes, however, it is possible to extract dependents from sentential complements:

(18) Kogo powiedziała, że nie   chce więcej widzieć?
      who   said   that NEG wants anymore see
      'Who did she say she does not want to see anymore?'

In (18) the *wh*-word *kogo* is placed in the main clause while in terms of f-structure it is an argument of the infinitival complement (*widzieć*) of the sentential complement (*chce*) of the main clause (*powiedziała*). To account for such data, the extraction path is extended for relevant items, namely for (phrases containing) *wh*-words:[3]

(19) XPATH ≡ COMP* XCOMP*

3.1 Monoclausal coordination

After particular conjuncts have been assigned appropriate functional annotation, they are fed into rules handling lexico-semantic coordination. The rule provided in (20) serves the purposes of handling sentences such as (1)–(4). Since the conjunction may be removed from these examples without any loss in grammaticality or any obvious change in meaning, these are assumed to have a monoclausal structure.

(20) XPlxm$_{type}$ →   XPextr$_{type}$ [,   XPextr$_{type}$]*   CONJ   XPextr$_{type}$
                          ↑=↓   ↑=↓   ↑=↓   ↑=↓

Let us see how the f-structure corresponding to (4), repeated in (21) below for convenience, is constructed in a stepwise manner.

(21) Skąd   i   jakie otrzymujemy informacje?
      whence and what receive   information
      'What information and where from do we receive?'   (NKJP)

---

[3]The provided extraction path is trivial since closer investigation of Polish extraction phenomena remains outside of the scope of this paper. To account for attested data, it may require certain adjustments, including imposing additional constraints on some of its parts.

Particular conjuncts build their own partial f-structures thanks to the rule provided in (12). It assigns each conjunct its own, independent grammatical function annotation and although this annotation is very general (it may in theory generate a path consisting of the extraction path and any sequence of grammatical functions), one must bear in mind that its output is constrained by the f-structure of the rest of the utterance. As a result, the rule may generate infinitely many structures, but only the following f-structures built by individual conjuncts may be unified with the rest:

(22) a. $\begin{bmatrix} \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHENCE'} \end{bmatrix} \right\} \end{bmatrix}$

  b. $\begin{bmatrix} \text{OBJ} & \begin{bmatrix} \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHAT'} \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix}$

Since all conjuncts in (20) bear the co-head annotation ('↑=↓'), unlike under the standard account of coordination (using the '↓∈↑' annotation), no set is created. Instead, f-structure fragments built by particular conjuncts, (22), are placed in one top-level f-structure:

(23) $\begin{bmatrix} \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHENCE'} \end{bmatrix} \right\} \\ \text{OBJ} & \begin{bmatrix} \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHAT'} \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix}$

Finally, (23) is unified with the f-structure of the rest of the utterance to yield the full f-structure provided in (24):[4]

(24) $\begin{bmatrix} \text{PRED} & \text{`RECEIVE}\langle\boxed{1},\boxed{2}\rangle\text{'} \\ \text{SUBJ } \boxed{1} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{NUM} & \text{PL} \\ \text{PERS} & 1 \end{bmatrix} \\ \text{OBJ } \boxed{2} & \begin{bmatrix} \text{PRED} & \text{`INFORMATION'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHAT'} \end{bmatrix} \right\} \end{bmatrix} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHENCE'} \end{bmatrix} \right\} \end{bmatrix}$

---

[4]Note that the f-structure provided in (24) includes an implicit subject (first person, plural).

3.2 Biclausal coordination[5]

A slightly different coordination rule, provided in (25)[6] below where the PART_{wh} category corresponds to the question particle *czy* (see the lexical entry in (26)), is designed for examples such as (7), repeated in (27) below, which are considered biclausal, as discussed above.

(25) $\quad \text{XPlxb}_{wh} \quad \rightarrow \quad \text{PART}_{wh} \quad [, \quad \text{XPextr}_{wh}]^* \quad \text{CONJ} \quad \text{XPextr}_{wh}$
$\qquad\qquad\qquad\qquad\quad\; \downarrow\in\uparrow \qquad\qquad \downarrow\in\uparrow \qquad\; \uparrow=\downarrow \qquad \downarrow\in\uparrow$

(26) $\quad$ czy $\quad \text{PART}_{wh} \quad (\uparrow \text{CLAUSE-TYPE})= \text{INT}$

(27) $\quad$ Nie wiadomo było, czy $\quad$ *(i) kiedy wróci.
$\qquad$ NEG know $\quad$ was whether and when returns
$\qquad$ 'It was not clear whether and when he would return.' $\qquad\qquad$ (NKJP)

To represent the fact that such utterances are not monoclausal, all conjuncts bear the set membership annotation ('$\downarrow\in\uparrow$'). As a result, partial f-structures constructed by individual conjuncts provided in (28) are placed inside a set, as shown in (29):

(28) $\quad$ a. $\quad \left[\text{CLAUSE-TYPE} \quad \text{INT} \right]$

$\qquad$ b. $\quad \left[\text{ADJ} \quad \left\{ \left[\text{PRED} \quad \text{'WHEN'} \right] \right\} \right]$

(29) $\quad \left\{ \left[\text{CLAUSE-TYPE} \quad \text{INT} \right], \left[\text{ADJ} \quad \left\{ \left[\text{PRED} \quad \text{'WHEN'} \right] \right\} \right] \right\}$

When the structure in (29) is unified with the f-structure of the rest of the utterance provided in (30), a biclausal coordinate structure results, as in (31):[7]

(30) $\quad \left[\begin{array}{ll} \text{PRED} & \text{'RETURN}\langle\boxed{1}\rangle\text{'} \\ \\ \text{SUBJ} \; \boxed{1} & \left[\begin{array}{ll} \text{PRED} & \text{'PRO'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 3 \end{array} \right] \end{array} \right]$

---

[5]The type of lexico-semantic coordination described here is referred to as *biclausal* coordination despite the fact that the rule provided in (25) may generate structures containing more than two clauses. It seems, however, that examples illustrating this phenomenon tend to feature two conjuncts, leading to a biclausal representation, hence the name.

[6]This rule accounts for examples where the question particle is the first conjunct. Such examples seem to be most frequent; there exist, however, examples in which *czy* serves as the last conjunct:

(i) $\quad$ Będą sprawdzać kto i $\quad$ czy miał zezwolenie
$\qquad$ will check $\quad$ who and PART had permission
$\qquad$ 'They will check whether (they had permission) and who had permission.' $\qquad$ (NKJP)

Such cases may be handled by applying simple word order modifications to the rule in (25).

[7]Note that the implicit subject (third person, singular) in (31) is structure-shared: it belongs to both clauses at the same time.

(31)
$$
\left\{
\begin{bmatrix}
\text{PRED} & \text{`RETURN}\langle\boxed{1}\rangle\text{'} \\[6pt]
\text{SUBJ}\ \boxed{1} &
\begin{bmatrix}
\text{PRED} & \text{`PRO'} \\
\text{NUM} & \text{SG} \\
\text{PERS} & 3
\end{bmatrix} \\[18pt]
\text{CLAUSE-TYPE} & \text{INT}
\end{bmatrix},
\begin{bmatrix}
\text{PRED} & \text{`RETURN}\langle\boxed{1}\rangle\text{'} \\[6pt]
\text{SUBJ} & \boxed{1} \\[6pt]
\text{ADJ} & \left\{\begin{bmatrix} \text{PRED} & \text{`WHEN'} \end{bmatrix}\right\}
\end{bmatrix}
\right\}
$$

The structure in (31) is biclasal because of the interaction of properties of its partial f-structures: (29) is a set and the PRED attribute in (30) is a distributive feature. When these structures are combined, (30) distributes to particular elements of the set in (29): 'copies' of (30) are merged with respective elements of (29), the question particle *czy* and the adjunct *kiedy* ('when'), and the resulting structures are enclosed in a set, as in (31).

3.3 Argument saturation under the biclausal analysis

The biclausal analysis of certain instances of lexico-semantic coordination forces the introduction of some changes to relevant rules in order to account for independent argument saturation in coordinated clauses. While modifications are not required by examples such as (27) because there is an intransitive predicate whose only argument is shared (the implicit subject in (31)), argument saturation turns out to be an issue with sentences such as the following:

(32) czy *(i) ile   będzie mogła zarobić tego typu placówka?
      PART and how much AUX be able earn   such   institution
      'Will such an institution be able to earn and how much will it be able to earn?'   (NKJP)

In (32) the interrogative particle (*czy*) is coordinated with one of the arguments of the verb *zarobić*, which expresses a two-place predicate, taking a subject and a direct object. The former is overt (*placówka*) and it is shared by both coordinated clauses. According to the analysis provided above, dependents coordinated under biclausal lexico-semantic coordination belong to different clauses. As a result, *ile* may only fill the object grammatical function of one of the clauses. To avoid the violation of the completeness principle, the object of the other clause must also be filled in some way. This can be achieved using the following statement to handle implicit argument saturation:

(33) PRODROP $\equiv$ $((\uparrow \text{SUBJ PRED})=\text{`PRO'})$
                       $((\uparrow \text{OBJ PRED})=\text{`PRO'})$
                       $\ldots$
                       $((\uparrow \text{GF PRED})=\text{`PRO'})$

The statement provided above consists of a set of equations optionally (they are enclosed in brackets) filling a given grammatical function with an implicit argument (represented as the PRO value of the PRED attribute). The last line of (33) is to be

treated as an abbreviation for all other appropriate grammatical functions, as defined in (16), with the exception of adjuncts.

It must be noted that the place of attachment of such statements is of importance – attaching (33) to the entire coordinate structure would give rise to a shared implicit dependent, which could cause violations of the uniqueness condition – a given grammatical function could be filled with a lexical dependent, leading to a clash with the implicit argument attempting to fill the same slot. For this reason (33) must not be placed inside the rule adding conjuncts to a set, it should instead be placed so that implicit arguments attach inside individual clauses. To achieve this, (33) should be attached at an intermediate level, so that its partial f-structure is merged with the f-structure fragment built by a given conjunct:

(34) $\text{XPextrbicl}_{type} \quad \rightarrow \quad \text{XPextr}_{type}$
$$\uparrow=\downarrow$$
$$\text{PRODROP}$$

Furthermore, care must be taken in order to ensure that conjuncts with prodrop statements are only used with biclausal lexico-semantic coordination. One of possible means to this end is to introduce additional categories for biclausal conjuncts exclusively, as in (34) above and (35) below:

(35) $\text{PARTbicl}_{type} \quad \rightarrow \quad \text{PART}_{type}$
$$\uparrow=\downarrow$$
$$\text{PRODROP}$$

Finally, the rule provided in (25) must be rewritten, replacing XPextr and PART categories with XPextrbicl and PARTbicl, respectively:

(36) $\text{XPlxb}_{wh} \quad \rightarrow \quad \text{PARTbicl}_{wh} \quad [,\ \text{XPextrbicl}_{wh}]^{*} \quad \text{CONJ} \quad \text{XPextrbicl}_{wh}$
$$\qquad\qquad\qquad\quad \downarrow\in\uparrow \qquad\qquad \downarrow\in\uparrow \qquad\quad \uparrow=\downarrow \qquad \downarrow\in\uparrow$$

Let us now construct the f-structure representing (32) stepwise to see the modifications discussed above at work. First, individual conjuncts construct their partial f-structures using (12):[8]

(37) a. $\begin{bmatrix} \text{CLAUSE-TYPE} & \text{INT} \end{bmatrix}$

　　　 b. $\begin{bmatrix} \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'HOW\_MUCH'} \end{bmatrix} \end{bmatrix}$

Subsequently, optional implicit arguments are added as a result of attaching (33) inside particular conjuncts, (34) and (35). The f-structure fragment provided in (38) shows how an implicit argument fills the object grammatical function in the f-structure which contains the question particle:

(38) $\begin{bmatrix} \text{CLAUSE-TYPE} & \text{INT} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \end{bmatrix} \end{bmatrix}$

---

[8]Note that (37b) is simplified: the implicit object of the numeral is not represented.

Next, conjuncts are added to a set using the modified rule handling biclausal lexico-semantic coordination provided in (36):

(39)
$$\left\{ \begin{bmatrix} \text{CLAUSE-TYPE} & \text{INT} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`HOW\_MUCH'} \end{bmatrix} \end{bmatrix} \right\}$$

Finally, the partial f-structure built by lexico-semantic coordination is merged with the f-structure of the rest of the utterance. The following full f-structure results:

(40)
$$\left\{ \begin{bmatrix} \text{PRED} & \text{`EARN}\langle\boxed{1},\boxed{2}\rangle\text{'} \\ \text{SUBJ}\ \boxed{1} & \begin{bmatrix} \text{PRED} & \text{`INSTITUTION'} \end{bmatrix} \\ \text{OBJ}\ \boxed{2} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \end{bmatrix} \\ \text{CLAUSE-TYPE} & \text{INT} \end{bmatrix}, \begin{bmatrix} \text{PRED} & \text{`EARN}\langle\boxed{1},\boxed{3}\rangle\text{'} \\ \text{SUBJ} & \boxed{1} \\ \text{OBJ}\ \boxed{3} & \begin{bmatrix} \text{PRED} & \text{`HOW\_MUCH'} \end{bmatrix} \end{bmatrix} \right\}$$

While the lexical subject is shared by both clauses (as indicated by appropriate structure sharing of relevant f-structure fragments), the object of the first clause (it contains the question particle, the first lexico-semantic conjunct) is filled with an implicit argument, while the object of the other clause is filled with a lexical argument, the second lexico-semantic conjunct.

4 Issues

While previous sections discussed key facts concerning lexico-semantic coordination, the aim of this section is to address potential doubts as to the standing of this phenomenon as a variety of coordination, as well as some less obvious (though important) issues and, finally, possible extensions.

4.1 Is this coordination?

Since lexico-semantic coordination is a potentially very surprising variety of coordination, it seems natural to question whether it is indeed an instance of coordination. While typical tests such as agreement seem inapplicable, there is fortunately some other potentially convincing evidence.

First, it is possible to use such constructions with items which are unambiguous and uncontroversial conjunctions in Polish:

(41)   [kto oraz kiedy] miałby płacić za postawiony budynek
       who and when   should pay   for erected   building
       'Who and when would be supposed to pay for the erected building?' (NKJP)

(41) features *oraz* ('and'), an entirely unambiguous conjunction as there is no other available interpretation of this word.

Furthermore, it is possible to find examples where a preconjunction is used, as in 'both... and...' coordinate structures:

(42) A jest i   co,   i   gdzie eksportować.
      and is and what and where export
      'There (certainly) is what and where to export to.'   (NKJP)

While all examples presented so far featured conjoining, *and*-type conjunctions (mostly *i*), there exist examples with alternative conjunctions:

(43) [kto lub czego] będzie w Wikipedii szukał.
      who or what   will   in Wikipedia seek
      'Who will seek what in Wikipedia?'   (NKJP)

While the word *lub* ('or') is not perfectly unambiguous, its other interpretation, the imperative form of the verb *lubić* 'like', is not an option in this context, leaving the conjunction interpretation. The LFG account of coordination using an alternative conjunction is exactly the same as for phrases coordinated using a conjoining conjunction and consequently the same convention was used for lexico-semantic coordination with such conjunctions. However, as in the case of more standard coordination, there is a difference in the semantics, perhaps less evident under lexico-semantic coordination. For this reason, it is important to record the shape of the conjunction involved, which is discussed in more detail in Section 4.2.

4.2 Representing the conjunction

The f-structures provided in Section 3 did not include the contribution of the annotation of the conjunction in any way. As mentioned in Section 4.1, the form of the conjunction, namely whether it belongs to the conjoining or the alternative type, is of importance from the perspective of semantics. Such information may be provided using a dedicated attribute, COORD-FORM for instance:

(44) a. i   CONJ (↑ COORD-FORM)= AND

      b. lub CONJ (↑ COORD-FORM)= OR

When conjunctions annotated in this way are used with rules such as (20), the rule handling monoclausal lexico-semantic coordination, the conjunction is represented in the top-level f-structure. The relevant fragment corresponding to the entire lexico-semantic coordinate phrase (including the conjunction) from example (4), repeated later as (21), is provided below:

(45)
$$
\begin{bmatrix}
\text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'WHENCE'} \end{bmatrix} \right\} \\
\text{OBJ} & \begin{bmatrix} \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'WHAT'} \end{bmatrix} \right\} \end{bmatrix} \\
\text{COORD-FORM} & \text{AND}
\end{bmatrix}
$$

When this fragment is unified with the f-structure of the rest of the utterance, the following structure results:[9]

---

[9]The f-structure of the implicit subject is simplified in (46) and the following examples.

(46)
$$
\begin{bmatrix}
\text{PRED} & \text{`RECEIVE}\langle\boxed{1},\boxed{2}\rangle\text{'} \\
\text{SUBJ } \boxed{1} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \end{bmatrix} \\
\text{OBJ } \boxed{2} & \begin{bmatrix} \text{PRED} & \text{`INFORMATION'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHAT'} \end{bmatrix} \right\} \end{bmatrix} \\
\text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHENCE'} \end{bmatrix} \right\} \\
\text{COORD-FORM} & \text{AND}
\end{bmatrix}
$$

Such a representation is potentially vulnerable to interference caused by dependent sharing whereby a single phrase is shared by more than one head, as in the following modified version of (4):

(47) [Skąd i jakie] [otrzymujemy lub kradniemy] informacje?
whence and what receive or steal information
'What information and where from do we receive or steal?'

When, as in (47), verbs are coordinated, the conjunction is represented at the same level as the set containing particular verbal heads. The structure provided below represents (47) with the exclusion of the lexico-semantic coordinate phrase (it corresponds to the following fragment: *[otrzymujemy lub kradniemy] informacje*):

(48)
$$
\begin{bmatrix}
\left\{ \begin{bmatrix} \text{PRED} & \text{`RECEIVE}\langle\boxed{1},\boxed{2}\rangle\text{'} \\ \text{SUBJ } \boxed{1} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \end{bmatrix} \\ \text{OBJ } \boxed{2} & \begin{bmatrix} \text{PRED} & \text{`INFORMATION'} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \text{PRED} & \text{`STEAL}\langle\boxed{1},\boxed{2}\rangle\text{'} \\ \text{SUBJ} & \boxed{1} \\ \text{OBJ} & \boxed{2} \end{bmatrix} \right\} \\
\text{COORD-FORM} \quad \text{OR}
\end{bmatrix}
$$

While *i* ('and') would be (accidentally) unproblematic as the element conjoining verbs because the same conjunction is used in the lexico-semantic coordinate phrase (see (45) and (46)), using *lub* ('or') as the conjunction in the coordinate verb phrase results in the clash ($\neq$) of values of COORD-FORM in the top-level f-structure:

(49)
$$
\begin{bmatrix}
\left\{ \begin{bmatrix} \text{PRED} & \text{`RECEIVE}\langle\boxed{1},\boxed{2}\rangle\text{'} \\ \text{SUBJ } \boxed{1} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \end{bmatrix} \\ \text{OBJ } \boxed{2} & \begin{bmatrix} \text{PRED} & \text{`INFORMATION'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHAT'} \end{bmatrix} \right\} \end{bmatrix} \\ \text{ADJ } \boxed{3} & \left\{ \begin{bmatrix} \text{PRED} & \text{`WHENCE'} \end{bmatrix} \right\} \end{bmatrix}, \begin{bmatrix} \text{PRED} & \text{`STEAL}\langle\boxed{1},\boxed{2}\rangle\text{'} \\ \text{SUBJ} & \boxed{1} \\ \text{OBJ} & \boxed{2} \\ \text{ADJ} & \boxed{3} \end{bmatrix} \right\} \\
\text{COORD-FORM} \quad \text{AND}\neq\text{OR}
\end{bmatrix}
$$

In (49) conjunctions used in two coordinate phrases, lexico-semantic (*i*) and verbal (*lub*), set conflicting COORD-FORM values, AND and OR, respectively, represented as inequality: AND$\neq$OR.

A related problem is caused by the embedding of coordination within lexico-semantic coordination, as in the two examples below:

(50) Nigdy nie   wiadomo, [[kto lub co], skąd   i   kiedy] zaatakuje.
     never NEG know   who or what whence and when   attacks
     'You never know who or what, where from and when may attack.'   (NKJP)

(51) kombinowaniem [kto, kogo, kiedy i   jak], [z   kim   przeciw komu] albo
     plotting   who whom when and how with whom against whom or
     [od kogo i   za co]
     from whom and for what
     '[...] plotting about who, whom, when and how, with whom against whom
     or from whom and for what [...]'   (NKJP)

Two varieties of coordination are involved in (50): the first conjunct of lexico-semantic coordination is at the same time a regular coordinate NP (both its elements bear the subject grammatical function), while the remaining lexico-semantic conjuncts are adjuncts (ablative and temporal). This is unproblematic representationally, because the conjunction is represented inside the coordinate NP.

Example (51), is considerably more interesting as it presents embedded lexico-semantic coordination: two edge conjuncts are also instances of such coordination. The first conjunct contains a subject, an object and two adjuncts (temporal and manner), the middle conjunct features multiple *wh*-phrases (two obliques), and the last conjunct consists of another oblique coordinated with an adjunct. It is possible to construct a less complicated example, though:

(52) [Kto i   kogo] lub [kiedy i   gdzie] poznał?
     who and whom or   when and where met
     'Who did meet whom, or when and where?'

In the above example the first conjunct contains a subject and an object while the other conjunct consists of two coordinated adjuncts.

There is a potential solution to problems posed by such examples in terms of discourse functions. The value of a discourse function, a hybrid structure, would represent lexico-semantic coordination: it would gather particular conjuncts inside a set and, as under standard coordination, the conjunction would be represented outside the set. Furthermore, particular conjuncts would be structure-shared with relevant parts of the main f-structure. In this way, the conjunction inside the lexico-semantic coordinate phrase would only be represented in the f-structure corresponding to the discourse function, making it impossible to conflict with the value of the COORD-FORM attribute (if present) of the top-level f-structure. Finally, embedding could be handled using standard coordination rules which give rise to embedded hybrid structures. Substantiating this general idea should be a matter of future work.

4.3 More types of lexico-semantic coordination?

Section 2 presented data focused on four types of items involved in lexico-semantic coordination, namely pronouns expressing an existential quantifier, (1), or a universal quantifier, (2), *n*-words, (3), and *wh*-words, (5)–(7). It seems, however, that the inventory of lexico-semantic types could be extended.

Kallas (1993) discusses the following example:[10]

(53) Jan pamięta tyle i takich oskarżeń.
John remembers that many and such accusations
'John remembers that many (of) such accusations.'   (Kallas, 1993, p. 53)

There are similar attested examples:

(54) my nie mogłybyśmy zapewnić naszym podopiecznym tylu i takich
we NEG could provide our charges so many and such
materiałów do pracy
resources for work
'We would not be able to provide our charges with so many (of) such work resources.'   (NKJP)

(55) Że będzie i jest tyle i takich afer?
that will and is that many and such scandals
'That there is and will be that many (of) such scandals?'   (NKJP)

In (53) and (54) particular conjuncts correspond to the object (*tyle* and *tylu*,[11] respectively) and the modifier of the object's object (*takich*). (55) shows lexico-semantic coordination of the subject and the modifier of the subject's object. It seems that the common feature of lexico-semantic conjuncts in the examples presented above is the fact that they belong to the class of demonstratives.

Another possible class is consistuted by free relatives. Recently, Citko and Gracanin-Yuksek (2012) discussed such coordination on the basis of data from Polish, English and Croatian. They provide the following Polish example:

(56) Jan je cokolwiek i kiedykolwiek Maria gotuje.
John eats whatever and whenever Mary cooks

In (56) an object (*cokolwiek*) and an adjunct (*kiedykolwiek*) are coordinated.

Finally, though Citko and Gracanin-Yuksek (2012) claim that there is 'a more general constraint that rules out two relative pronouns in a relative clause modifying a single head, regardless of whether the relative pronouns are coordinated or not' and provide example (57) in support of this claim, there are examples such as (58) which seem to be grammatical when coordination is used:

---

[10]The glosses and free translation in (53) are our own.

[11]In (54) the object is marked for genitive case, unlike in (53), as a result of object case assignment in the syntactic scope of negation. See Przepiórkowski 1999 for an extensive discussion of Genitive of Negation (GoN) in Polish.

(57) *student którego (i)   któremu Maria przedstawiła
      student who   and whom   Mary introduced

(58) człowiek, z   którym i   o   którym lubię mówić
      man   with whom and about whom like   talk
      'the man with whom and about whom I like to talk'

Incorporating examples such as the ones presented above (demonstratives, free relatives, relatives) in the analysis proposed in this paper should not pose any problems. Necessary changes would include extending the inventory of allowed semantic conjunct types in relevant rules and assigning demonstratives a type (DEM for instance). Relative pronouns and pronouns expressing an existential quantifier already bear appropriate types (REL and ANY, respectively) for independent reasons, namely for the purposes of handling relative clauses and free relatives. It must be noted, however, that such examples, especially ones including relative pronouns, are not as numerous, varied and productive as other examples presented in previous sections.

5   Conclusion

There is a growing interest in lexico-semantic coordination, also within LFG. In comparison to previous work, the main contributions of this paper include: showing that coordinated elements may be dependents of different heads, distinguishing between monoclausal and biclausal lexico-semantic coordination, and providing a relatively comprehensive analysis of lexico-semantic coordination in Polish.

However, some loose ends remain. The most pressing is the question of the representation of the conjunction, with some preliminary ideas suggested at the end of Section 4.2, but also the exact repertoire of semantic classes which may participate in lexico-semantic coordination, especially, the possibility of such coordination in free relatives and in ordinary relative clauses, mentioned in Section 4.3. Finally, the fundamental issue of why exactly such semantic classes make it possible to violate the overwhelming constraint that only the same grammatical functions can be coordinated has not been addressed. The present interest in lexico-semantic coordination will certainly continue to grow in the near future.

Acknowledgements

References

Chaves, Rui Pedro and Paperno, Denis. 2007. On The Russian Hybrid Coordination Construction. In Stefan Müller (ed.), *The Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, pages 46–64, Stanford: CSLI Publications.

Citko, Barbara and Gracanin-Yuksek, Martina. 2012. Wh-Coordination in Free Relatives. Abstract of a talk presented at the 43rd annual meeting of the North East Linguistic Society, City University of New York, 19–21 October 2012; `http://nels2012.commons.gc.cuny.edu/?page_id=14`.

Dalrymple, Mary and Lødrup, Helge. 2000. The grammatical functions of complement clauses. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG'00 Conference*, University of California, Berkeley.

Gazdik, Anna. 2010. Multiple Questions in French and Hungarian: An LFG Account. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG'10 Conference*, pages 249–269, Ottawa, Canada.

Gazdik, Anna. 2012. *Multiple Questions in French and Hungarian. A Lexical-Functional Analysis with Special Emphasis on the Syntax-Discourse Interface*. Ph. D.thesis, Université Paris 7.

Kallas, Krystyna. 1993. *Składnia współczesnych polskich konstrukcji współrzędnych*. Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika.

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

Patejuk, Agnieszka and Przepiórkowski, Adam. 2012. Towards an LFG parser for Polish: An exercise in parasitic grammar development. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC 2012*, pages 3849–3852, ELRA, Istanbul, Turkey.

Przepiórkowski, Adam. 1999. *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Ph.D. dissertation, Universität Tübingen, Germany.

Przepiórkowski, Adam, Bańko, Mirosław, Górski, Rafał L. and Lewandowska-Tomaszczyk, Barbara (eds.). 2012. *Narodowy Korpus Języka Polskiego [Eng.: National Corpus of Polish]*. Warsaw: Wydawnictwo Naukowe PWN.

Przepiórkowski, Adam, Górski, Rafał L., Łaziński, Marek and Pęzik, Piotr. 2010. Recent Developments in the National Corpus of Polish. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation, LREC 2010*, ELRA, Valletta, Malta.

Przepiórkowski, Adam and 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.

Sag, Ivan A., Gazdar, Gerald, Wasow, Thomas and Weisler, Steven. 1985. Coordination and How to Distinguish Categories. *Natural Language and Linguistic Theory* 3, 117–171.

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

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

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