

Incorporating conjunctions in Polish

Agnieszka Patejuk

aep@ipipan.waw.pl

Introduction While coordinating conjunctions typically occur between items that are being joined, incorporating conjunctions may¹ be placed inside the conjunct they are supposed to precede. In (1) two sentences are coordinated, but the conjunction *zaś* follows the word *dołączony* ‘attached’ which belongs to the second conjunct – it is an adjunct of *protokół* ‘report’, the object (OBJ) of *otoczono*, a form of the verb *otoczyć* ‘surround’ – see the f-structure in (13).²

- (1) Tekst był jawny, dołączony **zaś** tajny protokół otoczono szczelną zasłoną milczenia.
text was public attached.ACC but secret.ACC report.ACC surrounded airtight curtain silence

‘The text was public, but the attached report was shrouded in secrecy.’ (The National Corpus of Polish, NKJP)

Similarly, while subordinating conjunctions are expected to be the first element of a subordinate clause (CP), followed by the clause, incorporating subordinating conjunctions may be placed instead inside the relevant clause. In (2) the subordinating conjunction *bowiem* follows the word *podstawowym* ‘main’ which belongs to the subordinate clause – it is an adjunct of *zadaniem* ‘duty’, the predicative complement (XCOMP-PRED) of *jest*, a form of the verb *BYĆ* ‘be’ – see the f-structure in (14).

- (2) Takim zawodem [...] jest zawód lekarza weterynarii, podstawowym **bowiem** jego zadaniem jest ochrona.
such profession is profession doctor veterinary main.INST since his duty.INST is protection.NOM

‘Such a profession is the profession of veterinarian, since protection is his main duty.’ (NKJP)

Though in both examples above the conjunction follows the first word of the relevant clause, the structural distance is greater – the conjunction is embedded as an adjunct of an argument (subject or predicative complement, respectively) of the main verb of the relevant clause (coordinate or subordinate). Moreover, contrary to what is often assumed, incorporating conjunctions may be far more distant both in terms of linear order (number of words) and syntactic distance (the number of dependents) – incorporating conjunctions may be embedded deep in the clause, inside any of its dependents.

More data Let us discuss some more examples which illustrate the use of incorporating conjunctions in Polish, starting with coordinating conjunctions. In (3) the conjunction *zaś* is placed after the third word of the second conjunct – the verb *wyraziły* ‘expressed’, which is preceded by its subject, *władze sowieckie* ‘Soviet authorities’, so the linear distance is two phrases. There are also 3 words before the conjunction *natomiast* in (4), but it is embedded deeper, so the structural distance is greater – *natomiast* follows the object (OBJ, *ich* ‘their’) of the infinitival complement (XCOMP, *wyszukiwać* ‘seek’) of the head verb of the second conjunct (*trzeba* ‘need’), preceded by negation (*nie*) – see the f-structure in (15).

- (3) uzyskał zwolnienie wszystkich zakładników, władze sowieckie wyraziły **zaś** zgodę na ich powrót.
achieved release all hostages authorities Soviet expressed but consent on their return

‘He achieved the release of all hostages, whereas the Soviet authorities agreed to their return.’ (NKJP)

- (4) należy karać tych chrześcijan, którzy są oskarżeni przed władzą, nie trzeba ich **natomiast** wyszukiwać.
should penalise those Christians who are accused before authority NEG need them but seek.INF

‘We should penalise those Christians who were accused by the authorities, but we should not search for them.’ (NKJP)

Let us now proceed to subordinating conjunctions, illustrated using examples with *bowiem*. In (5) *bowiem* is preceded by 4 words: the adjunct prepositional phrase (*od dawna* ‘long’), the main verb of the subordinate clause (*był* ‘was’) and its subject (*on* ‘he’), which it follows – the linear distance is therefore 3 phrases. In (6) the linear distance is two words, but the structural distance is different – *bowiem* is embedded inside the first phrase of the subordinate clause, it follows the numeral object (*wielu* ‘many’) of the preposition functioning as an adjunct of the head of the subordinate clause, splitting the head of the numeral phrase (*wielu*) from the accompanying nominal (*wypadkach* ‘cases’) – see the f-structure in (16).

- (5) biskup uważał ich ciągle za swoich podwładnych, od dawna był on **bowiem** zwierzchnikiem szkół.
bishop considered them still for SELF subordinate from long was he since head school

‘The bishop still considered them to be his subordinates, since he has long been the head of schools.’ (NKJP)

- (6) Zrealizowanie zamierzenia spowoduje ogromne uciążliwości [...], w wielu **bowiem** wypadkach kolej jest jedynym środkiem lokomocji.
realising plan cause great inconvenience in many since cases railway is only

środkiem lokomocji.

means transport

‘Realising this plan will cause great inconvenience, since in many cases the rail is the only means of transport.’ (NKJP)

Complex interactions Finally, let us investigate more interesting examples where the incorporating coordinating conjunction is located inside a subordinate clause which is an adjunct of one of the conjuncts joined by the conjunction.

- (7) Komisja przedstawiła swoje stanowisko, jeśli **natomiast** mamy to przedstawić Senatowi, musi zostać zaprezentowany projekt uchwały.
committee.NOM presented its view if but have.1.PL this.ACC present.INF senate must be.INF presented draft resolution

‘The committee has presented its view, but, if we are to present it to the senate, the draft of the resolution must be presented.’ (NKJP)

¹Though some prescriptive rules claim that they must occur in the non-standard position, corpus data convincingly shows that this is not the case – there are numerous non-incorporating instances of conjunctions discussed in this paper. The existence of such variation provides additional motivation for having a consistent, parallel representation of both uses – the interpretation is the same in spite of presence or absence of incorporation.

²The f-structure of the item to which the conjunction attaches is marked in red, while the f-structure built by the conjunction is marked in blue.

- (8) trudno wydawać pismo w regularnych odstępach czasu, jeśli się **zaś** tego nie czyni, nie zdobywa się stałego
hard publish.INF magazine in regular intervals time if REFL but this NEG do NEG gain REFL regular
odbiorcy
reader

‘It is hard to publish the magazine in regular intervals, but if you do not, you do not gain regular readers.’ (NKJP)

In (7) *natomiast*, which joins sentences headed by verbs *przedstawiła* ‘presented’ and *musi* ‘must’, is placed inside the subordinate clause introduced by *jeśli* ‘if’, an adjunct of *musi*. However, (7) would also be grammatical if *jeśli* and *natomiast* were not adjacent, for instance: *jeśli mamy to natomiast*, with the verb and the object of its infinitival complement between.

In (8) the incorporating conjunction *zaś* joining sentences headed by *trudno* ‘hard’ and *zdobywa* ‘gain’ is embedded in the adjunct subordinate clause (also introduced by *jeśli*) which depends on *zdobywa*. The conjunctions are separated by the impersonal marker *się* – a dependent of *czyni* ‘do’ (the main predicate of the subordinate clause). In (8) *zaś* could also follow *tego*, the object of *czyni*: *jeśli się tego zaś nie czyni* – the f-structure in (17) corresponds to this possibility.

Analysis and formalisation Based on the data presented above, the analysis (implemented in XLE) aims to capture two basic insights: first, that the position of an incorporating conjunction is always following some other category – it may be the first word of the relevant phrase, coordinate or subordinate, but it may also be more distant both in terms of linear distance as well as structural distance. Secondly, that the incorporating conjunction is embedded inside the category that it follows – and it builds its f-structure from that non-standard position. As a result, while in terms of c-structure the position of incorporating conjunctions is non-standard, the obtained f-structures are fully parallel to corresponding f-structures with standard conjunctions – following the way they are interpreted. Drawing on these insights, the idea behind the formal analysis of this phenomenon is to minimally affect the organisation of the rest of the grammar, putting the burden of producing the appropriate analysis on elements related to incorporating conjunctions.

The first insight is formalised using a metarule macro *METARULEMACRO-INK* defined in (9) – it allows incorporating conjunctions, both coordinating (*CONJINK*) and subordinating (*COMPINK*), to attach after *CAT* – the variable corresponding to any category defined in the grammar. For example, when *CAT* is an NP, the output of (9) are the rules $NP \rightarrow NP\ CONJINK$ and $NP \rightarrow NP\ COMPINK$ – note that, since there is no f-structure annotation, all right-hand side elements are co-heads.

(9) $METARULEMACRO-INK(CAT) \equiv CAT\ \{ CONJINK\ | COMPINK\ }$

The second, main part of the analysis is formalised in the lexical entries of incorporating conjunctions – they implement the insight that such conjunctions, despite being embedded inside the conjunct or inside the subordinate phrase, build their f-structure bottom up, possibly far higher in the f-structure than the place where they are located in terms of c-structure. This is captured in templates in (11) and (12) with the help of constraints relying on inside-out functional uncertainty, both of which use the *GF* variable defined in (10):

(10) $GF \equiv \{ SUBJ|OBJ|OBJ_{\theta}|OBL|XCOMP|XCOMP-PRED|COMP|ADJUNCT \}$

(11) $CONJ-INK(P) \equiv$
 $(\in GF^* \uparrow) = \%G$
 $(\%G\ COORD-FORM) = P$
 $\neg(ADJUNCT\ \%G)$

(12) $COMP-INK(P) \equiv$
 $(COMP\ GF^* \uparrow) = \%B$
 $(\%B\ PRED) = 'P < (\%B\ COMP) >'$

The template in (11) is called inside lexical entries of coordinating incorporating conjunctions such as *zaś* and *natomiast* – the *P* parameter corresponds to the lemma of the conjunction. The first line of (11), $(\in GF^* \uparrow) = \%G$, is a definition of the path in which the f-structure of the incorporating conjunction is to be built – it uses an inside-out equation coupled with functional uncertainty, allowing the conjunction to build its structure going up the path consisting of any sequence (including zero) of *GF* defined in (10), with an obligatory set element at the very end of the path. This path is assigned to the *%G* variable used in the remaining constraints: the second constraint, $(\%G\ COORD-FORM) = P$, introduces *P*, the conjunction’s lemma, as the value of the *COORD-FORM* attribute (coordinating conjunction form) in *%G* path, which is the f-structure which contains the set with the conjuncts – yielding the hybrid f-structure typical of coordination in LFG. Finally, the third constraint, $\neg(ADJUNCT\ \%G)$, ensures that the “landing” site of the f-structure introduced by the incorporating conjunction is not inside the *ADJUNCT* grammatical function – its value is also a set, so it would also satisfy the condition of having a set element at the end of the path in *%G*. However, this simple negative constraint precludes this.

The idea behind the template in (12), which handles subordinating incorporating conjunctions such as *bowiem*, is similar but the details are different – this is because subordinating conjunctions, unlike coordinating conjunctions, introduce a *PRED* attribute, which takes a *COMP* attribute containing the clause. As a result, the first constraint in (12), $(COMP\ GF^* \uparrow) = \%B$, defines an inside-out path which passes through any sequence (including zero) of *GF* and ends with a *COMP* grammatical function – this path is assigned to *%B* variable. It is used in the following constraint, $(\%B\ PRED) = 'P < (\%B\ COMP) >'$, which introduces the *PRED* value of the subordinating conjunction, whose value is another variable, *P* – its lemma.

The interaction of the metarule in (9) and the lexical entries of incorporating conjunctions, which call templates in (11) and (12), results in creating appropriate dependencies in f-structure despite the non-standard placement of such conjunctions – the functional uncertainty used in (11) and (12) is constrained by the f-structure built by the rest of the sentence.

Conclusion This paper presented an implemented LFG analysis of incorporating conjunctions in Polish which, despite their non-standard c-structure placement, being potentially deeply embedded inside the structure of a different constituent, successfully accounts for their f-structure, maintaining parallelism with sentences without incorporation.

