

Structure of conditional and (cor)relative clauses: New evidence from locality

Petr Biskup & Radek Šimík

Humboldt-Universität zu Berlin

1. Introduction and proposal

In this paper, we address the syntax and semantics of clausal arguments like (1) and adjuncts like (2) in two different syntactic positions: (a) left-adjoined and (b) integrated. All data are from Czech, unless indicated otherwise.

- (1) a. Co(koliv) mu dáš, utratí.
what(ever) him give.2SG spends
'What(ever) you give him, he'll spend (it).'
- b. Utratí, co(koliv) mu dáš.
spends what(ever) him give.2SG
'He'll spend what(ever) you give him.'
- (2) a. Když odejdeš, budu smutný.
when/if leave.2SG will.be.1SG sad
'When/If you leave, I'll be sad.'
- b. Budu smutný, když odejdeš.
will.be.1SG sad when/if leave.2SG
'I'll be sad when/if you leave.'

We put forth novel evidence from locality, suggesting that conditionals/(cor)relatives in the left-adjoined position are clausal (CP), denoting (sets of) propositions, and the corresponding integrated structures are nominal (NP), possibly embedded in a PP, denoting entities of the appropriate sort (individuals, situations, time intervals, etc.). More particularly, we propose that the underlying LFs of sentences like in (1)/(2) – independently of the (a)/(b) distinction – are as in (3)/(4):¹

¹ The shared LF implies that (a) and (b) are truth-conditionally equivalent, which seems correct to us.

- (3) [OP [_{CP} what(ever)_i you give him]] he spends [_{DP} the thing that you give him]_i.
- (4) [OP [_{CP} when_i you leave]] I'll be sad [_{PP} in [_{DP} that situation in which you leave]_i].

The *wh*-clauses are represented twice: once as conditional antecedents (arguments of a left-peripheral “conditional-generating” operator, assuming Kratzer’s 1979 semantics for conditionals) and once as definite descriptions, donkey-anaphoric to the entity introduced in the conditional by the *wh*-word. The way (a) and (b) differ is in their overt form: in the left-adjoined variant, the *wh*-clause spells out the conditional antecedent; in the integrated variant, the *wh*-clause spells out the definite description, essentially being a free relative, which are, rather uncontroversially, definite (ever since Jacobson 1995; see also Šimík to appear). Our analysis is, in a sense, a generalization of Hirsch’s (2016) proposal for ever free relatives, exemplified by (1b) with ‘ever’.² Hirsch proposes that ever free relatives are unconditionals (in the sense of Rawlins 2013) + donkey-anaphoric free relatives (at LF). What we propose here is that all *wh*-clauses are conditionals (incl. correlatives; Brasoveanu 2008; Arsenijević 2009) + donkey-anaphoric free relatives. A fuller exposition of this general claim will have to await another occasion.

In the present paper, our aim is more modest: We will show that left-adjoined *wh*-clauses are consistently transparent for A’-extraction, while their string-identical integrated versions are syntactic islands. This follows if the former are CPs and the latter are NPs. We propose a minimalist analysis employing phases and antilocality, in which the presence of the NP phase plays a crucial role in blocking extraction.

For reasons of space, we concentrate on syntax, leaving compositional semantics for another occasion (but see Hirsch 2016 and Šimík 2018 for analyses compatible with the present syntax). What we also leave open is the question how the LFs in (3) and (4) are arrived at, or more particularly, whether the *wh*-clause is represented in both positions or replaced by an unstructured variable in one of them and, if the former is the case, whether the relation is one of movement (or multidominance) or base-generated “matching” followed by ellipsis. Finally, let us point out that while all our data are from Czech, an initial informal survey suggests that the pattern extends to other Slavic languages (not reported here).

The rest of the paper is structured as follows. Section 2 is a systematic argument showing that left-adjoined *wh*-clauses are transparent, while integrated ones are not. Section 3 proposes an analysis of the observed extraction pattern. Section 4 concludes the paper.

2. Evidence from locality

We will show that the pattern schematized in (5), where XP has undergone A’-movement out of the *wh*-clause, abbreviated as C, holds in Czech. The pattern supports our claim

² Hirsch proposal concerns English ever free relatives expressing ignorance. Šimík (2018) extends Hirsch’s analysis to other languages and other semantic types of ever free relatives.

that C = CP in (5a) and C = NP in (5b).³ Further independent evidence will be used to support the assumption that NP/DP-hood is the factor behind the island-status of wh-clauses.

- (5) a. [XP₁ [C ... t₁ ...] main clause]
 b. * [XP₁ main clause [C ... t₁ ...]]

Consider examples (6)–(8), all of which exemplify the pattern in (5).⁴ The extracted element is a relative pronoun, which takes part in the derivation of a headed relative clause. (The intended interpretations of the ungrammatical (b)-examples are the same as of the corresponding (a)-examples.)

(6) *Extraction from (a) correlative vs. (b) free relative*

- a. To je ten chlap, [kterému₁ [C co(koliv) dáš t₁], to ztratí].
 it is the man.NOM which.DAT what(ever) give.2SG that lose
 ‘This is the man such that he will lose whatever you give him.’
 b. * To je ten chlap, [kterému₁ ztratí [C co(koliv) dáš t₁]].
 it is the man.NOM which.DAT lose what(ever) give.2SG

(7) *Extraction from (a) left-adjoined vs. (b) integrated conditional*

- a. Potkal člověka, [kterého₁ [C když poprosíš t₁], tak to zařídí].
 met man.ACC which.ACC when ask.2SG so it arrange
 ‘He met a man such that he/she/they will arrange it if you ask him.’
 b. * Potkal člověka, [kterého₁ to zařídí [C když poprosíš t₁]].
 met man.ACC which.ACC it arrange when ask.2SG

(8) *Extraction from (a) left-adjoined vs. (b) integrated purpose clause*

- a. To je řečník, [kterého₁ [C aby nalákali t₁], museli by mít peníze].
 it is speaker which.ACC so.that attract must SBJ have money
 ‘This is a speaker such that they need more money in order to attract him.’
 b. * To je řečník, [kterého₁ by museli mít peníze [C aby nalákali t₁]].
 it is speaker which.ACC SBJ must have money so.that attract

We proceed as follows. We demonstrate (i) that the extraction site is within C rather than in the main clause, (ii) that the landing site is outside of C (not just at its edge) and, more particularly, in the extended projection of the main clause, and finally (iii) that what blocks the extraction in cases like (b) above is the presence of a nominal layer on top of C.

³ The pattern is surprising because it goes in the opposite direction to what many influential approaches to locality would predict, including theories based on Condition on Extraction Domains (Huang 1982, Chomsky 1986, Manzini 1992), which prohibit extraction from adjuncts and/or derived positions, or theories based on information structure (Erteschik-Shir 1973, Erteschik-Shir & Lappin 1979), which prohibit extraction from “non-dominant” (topicalized, presuppositional) constituents.

⁴ This pattern was first discussed by Lešnerová & Oliva (2003), who provide many corpus examples, as well as an analysis within a dependency approach to syntax.

2.1. Extraction site is in C

Let us consider example (6) and assume, for the sake of the argument, that the relative pronoun has been extracted from the main clause subject position. The subject is indeed covert and coreferential with the relative pronoun, so it could well be a trace, while the gap internal to C (the indirect object of ‘give’) could be parasitic. Such an analysis would indeed conform to much of what we know or believe about how natural language syntax works. Yet, it is demonstrably wrong. First, the relative pronoun carries dative – assigned within C – and not nominative, as expected under our strawman analysis. A nominative relative pronoun is in fact ungrammatical; see (9a). Second, if the main clause gap corresponds to an oblique-case-marked argument, it is ungrammatical, as shown in (9b), where the argument position must be filled by an overt (clitic) dative pronoun.⁵ Third, the main clause need not contain an occurrence of a gap or pronoun coreferential with the extracted element. This is evident from (8) and reiterated in (9c). We conclude that if there is a subject gap coreferential with the extracted element, as in (6a)/(9a) or (7), then it is a plain vanilla *pro*, not a trace.

- (9) a. To je ten chlap, {kterému / *který} [C co(koliv) dáš], to ztratí.
 it is the man which.DAT ~.NOM what(ever) give.2SG that lose
 ‘This is the man such that he will lose whatever you give him.’
- b. To je ten chlap, kterému [C když zavoláš], pomůžeš { mu / *Ø}.
 it is the man which.DAT when call.2SG help.2SG him.DAT
 ‘This is the man such that if you call him, you’ll help him.’
- c. To je ten chlap, kterému [C když zavoláš], (tak) odejdu.
 it is the man which.DAT when call.2SG so leave.1SG
 ‘This is the man such that if you call him, I’ll leave.’

Example (10) shows that the extracted element reconstructs into C for purposes of reflexive anaphor binding (*svou* ‘his.REFL’ is bound by *Karel*), which is known to be subject to stringent locality conditions in Czech (Dočekal 2005). Notice that while the main clause has a *pro* coreferential with *Karel*, there is no position for the extracted phrase to reconstruct to and hence no way that *svou* ‘his.REFL’ could be bound by this *pro*. (The relevance of the two interpretations will become evident shortly.)

- (10) [Kterou svou_i cennost]_i říkali, že [C když si Karel_i uschová t₁],
 which his.REFL valuable said.PL that when REFL Karel deposits
 tak udělá nejlíp?
 so does best
 a. *Available interpretation*: ‘Which one of Karel’s valuables is such that they said that it would be wise of Karel if he deposits it?’

⁵ Main clause gaps are sometimes licensed (and attested in corpus) if they correspond to accusative-marked pronouns (Lešnerová & Oliva 2003: 249). This is arguably related to the fact that accusative-marked resumptive pronouns sometimes alternate with gaps in Czech resumptive relatives (see e.g. Toman 1998).

- b. *Unavailable interpretation*: ‘Which one of Karel’s valuables is such that it would be wise of Karel if they said that he deposits it.’

2.2. Extraction targets a position external to C

Believers in the ban on extraction out of (left) adjuncts might have thought – at least until encountering our last example – that the extracted element has not left the adjunct in the first place – perhaps it just moved to the edge of the adjunct. This assumption has in fact been adopted by Heck (2008) or Grewendorf (2015), in order to make sense of comparable constructions in (Bavarian) German, exemplified in (11).

- (11) De Mass wenn i no drink, bin i bsuffa. (Grewendorf 2015: 232–233)
 the liter if I still drink am I drunk
 ‘If I still drink this Mass, I’ll be drunk.’

Example (10) clearly shows that the extracted phrase can escape C in Czech. Importantly, as the unavailability of interpretation (10b) explicates, the verb *řikali* ‘said’ embeds the main clause (modified by C), not just C. Similar kind of evidence is provided by the placement of epistemic and evidential particles, such as *prý* ‘allegedly’, exemplified in (12). As the availability of the interpretations in (a)/(b) suggests, *prý* is attached to (semantically modifies) the main clause (MC), not the conditional. In fact, in Czech, as in English, epistemic and evidential adverbs cannot even modify conditional antecedents (for English, see McDowell 1987, Drubig 2001), which is also the reason why we have replaced *allegedly* with *they say that* in (12b). We conclude that the landing site of elements extracted out of C is, or at least can be, external to C.

- (12) To je ten člověk, který [MC *prý* [C *když* promluví], všichni
 it is the man who.NOM allegedly when speaks everybody
 ztichnou].
 falls.silent
 a. *Available interpretation*: ‘This is the man such that allegedly everybody falls silent when he speaks.’
 b. *Unavailable interpretation*: ‘This is the man such that everybody falls silent if they say that he speaks.’

2.3. What blocks extraction is a nominal layer

In this section, we show that it is the nominal element that blocks extraction out of C. Compare the grammatical (12) with the ungrammatical (13), in which C is embedded in the *v* ‘in’ PP.

- (13) * To je ten člověk, kterého₁ [*prý* [PP *v* ten moment [C *když* uvidíš t₁]],
 it is the man who.ACC allegedly in the moment when see.2SG
 budeš žasnout].
 will.2SG be.amazed

Example (14) shows that ungrammaticality also arises when the dominating NP is not embedded in a PP. Compare the grammatical extraction in (6a) with the ungrammatical one in (14), with the pronoun *to* ‘it’.

- (14) * To je ten chlap, [kterému₁ [NP to [C co dáš t₁]], to ztratí].
 it is the man.NOM which.DAT it what give.2SG that lose.3SG

The following data show that the nominal layer blocks extraction irrespective of whether extraction happens from a left-adjoined wh-clause, as in (13) and (14), or from an integrated wh-clause, as in (15) and (16).

- (15) * To je ten člověk, kterého [prý budeš žasnout [PP v ten moment [C když uvidíš t₁]].
 it is the man who.ACC allegedly will.2SG be.amazed in the moment when see.2SG

- (16) * To je ten chlap, [kterému₁ ztratí [NP to [C co dáš t₁]].
 it is the man.NOM which.DAT lose.3SG it what give.2SG

The proposal is also supported by the fact that adjuncts introduced by subordinators containing a nominal element like *protože* ‘because’ (lit. for.it.DECL-COMP) *poté co* ‘after’ (lit. after.it REL-COMP, cf. Biskup 2019) also do not allow extraction, as demonstrated below.

- (17) * To je ta knížka, [kterou₁ [{ protože / poté co } koupil t₁] už nemá peníze].
 it is the book which.ACC because after bought already NEG.have.3SG money

3. A syntactic analysis of the extraction pattern

In this section, we provide a particular syntactic analysis of the extraction pattern observed above, building on the idea that left-peripheral clausal adjuncts are CPs, but their integrated variants are NPs, proposed in section 1 and empirically supported in section 2.

3.1 Extraction from the integrated/NP-headed wh-clause

Following Abels (2003), Grohmann (2003), Ticio (2003) and Bošković (2005), we assume that phrasal movement observes antilocality, specifically, that phrasal movement must cross at least one full phrase. In addition, we assume that Czech, as an article-less language, has the following phases: CP, vP and NP (cf. Chomsky 2007, 2008; Dyakonova 2009; Bošković 2014). In earlier generative frameworks, the impossibility of extraction from a wh-clause embedded in an NP would be explained in terms of Subjacency (Chomsky 1973) or Barrierhood (Chomsky 1986), with CP and NP being bounding nodes/barriers. The phase model can mimic it with two adjacent phases (CP

and NP), where movement from the lower phase edge to the higher phase edge violates antilocality (see Bošković 2017). This means that the illicit movement from the integrated wh-clause looks like (18).

- (18) [... [NP **whP₁** [NP (demonstrative) [NP N [CP **t₁** C ... **t₁**...]]]]]
-
- The diagram shows a horizontal line representing a phase boundary. Above the line, on the left, is an upward-pointing arrow labeled 'whP₁'. Below the line, on the right, is a downward-pointing arrow labeled 't₁'. The two arrows cross each other, indicating movement from the lower phase to the higher phase across the boundary.

The moving element (**whP₁** in (18)) must move out of the integrated clause via the CP-edge. Since the next higher phase edge is the edge of NP and there is no full phrase on the way (which can be crossed), the CP-to-NP movement violates antilocality. How this proposal applies to the data under discussion is shown below. As an illustration consider the left-peripheral NP-embedded conditional in (19), repeated from (13), and the integrated conditional in (20), repeated from (15). In both cases, the problematic traces (the illicit movement step) are in italics. As we proposed in section 1, the integrated case – here (20) – is always represented as an NP (and interpreted as a donkey-anaphoric entity), even if an overt nominal (*ten moment* ‘the moment’) is missing.

- (19) * To je ten člověk, **kterého₁** [prý [PP v [NP **t₁** ten moment [CP **t₁** když uvidíš **t₁**]], budeš žasnout].
 it is the man who.ACC allegedly in the moment when
 see.2SG will.2SG be.amazed

- (20) * To je ten člověk, **kterého₁** [prý budeš žasnout [PP v [NP **t₁** ten moment [CP **t₁** když uvidíš **t₁**]]].
 it is the man who.ACC allegedly will.2SG be.amazed in the
 moment [CP **t₁** když uvidíš **t₁**]].
 moment when see.2SG

3.2 Extraction from the left-adjoined wh-clause

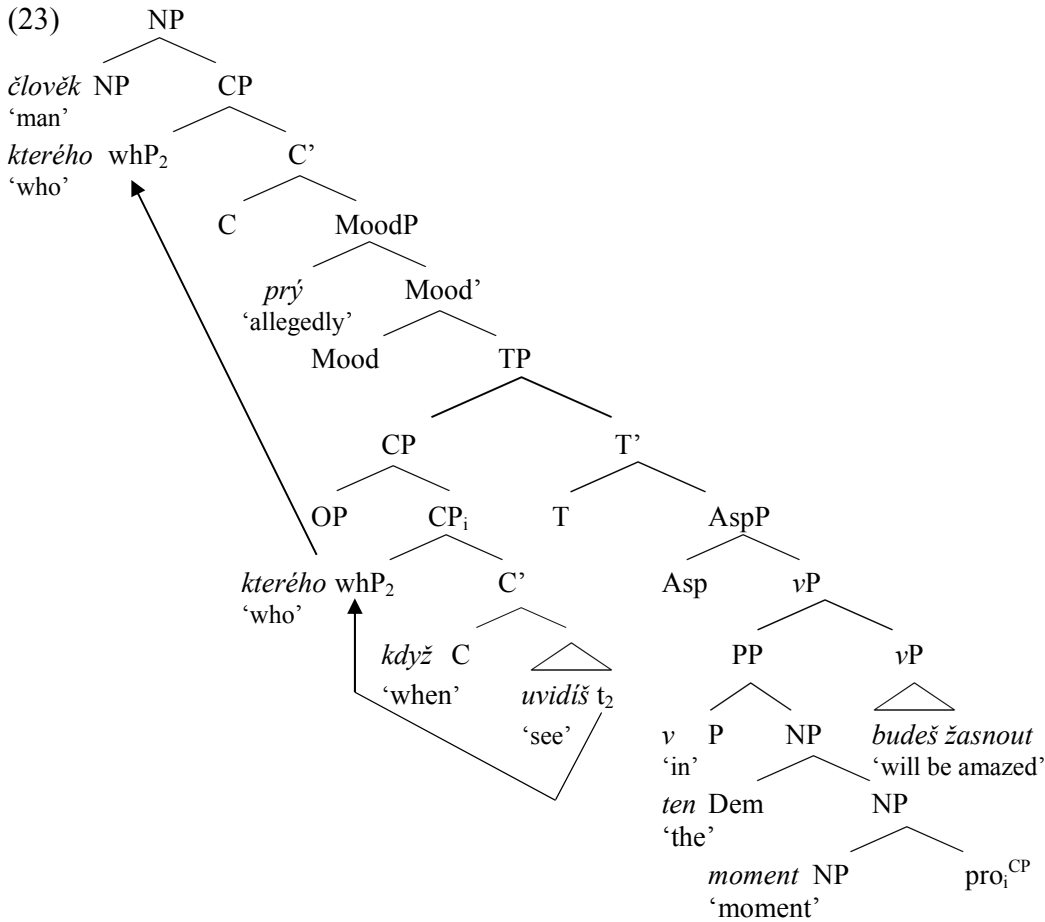
Recall that left-adjoined wh-clauses (without an overt nominal head) have the syntactic structure in (21) and that they differ from cases like (19), in which the wh-clause is embedded in NP.

- (21) [CP ... [OP [CP wh...]] ... [VP]]

Thus, given the proposal in the preceding section, these wh-clauses are not expected to block extraction, which is in accordance with the data presented in section 2. For instance, example (12), slightly modified as (22), will be derived as shown in (23).

- (22) To je ten člověk, kterého prý když uvidíš, (v ten moment)
 it is the man who.ACC allegedly when see.2SG in the moment
 budeš žasnout.
 will.2SG be.amazed
 ‘This is the man such that allegedly you will be amazed (in the moment) when you spot him.’

The whP *kterého* moves to the edge of CP₁ and the adjunct merges with the conditional operator and functions as its restrictor.⁶ In the next step in (23), CP₁ adjoins to the matrix TP (the nucleus). After that, *kterého* is going to move to CP of the matrix clause. As discussed in the preceding section, this step must cross at least one full phrase in order not to violate antilocality. Given that phrasal elements like the adverbial *prý* can intervene between the moved *který* and the adjunct clause, this seems to be the case (we assume a Mood projection for such adverbials). And since the wh-clause in (22) is not embedded in an NP, the movement step is licit.



4. Conclusion

We showed that A'-extraction in Czech is possible out of left-peripheral clausal adjuncts (conditionals, correlatives, and the like), but impossible out of string-identical, but syntactically integrated clauses. We argued that this surprising extraction pattern – previously observed in Lešnerová & Oliva 2003, but novel to the generative discourse – follows from the idea that left-peripheral clausal adjuncts are CPs (semantically propositions), while their integrated versions are NPs (semantically entities). We

⁶ For the sake of explicitness, we assume that in the optional NP *moment*, the adjunct clause is represented as a clausal *pro*.

proposed an analysis couched in a phase- and antilocality-based approach to syntactic islands.

References

- Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. University of Connecticut PhD dissertation.
- Arsenijević, Boban. 2009. {Relative {conditional {correlative clauses}}}. In Anikó Lipták (ed.), *Correlatives crosslinguistically*, 131–156. Amsterdam: John Benjamins.
- Biskup, Petr. 2019. *Prepositions, case and verbal prefixes. The case of Slavic*. Amsterdam: John Benjamins.
- Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59: 1-45.
- Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45(1): 27–89.
- Bošković, Željko. 2017. Extraction from complex NPs and detachment. In Martin Everaet and Henk van Riemsdijk (eds.), *The Wiley-Blackwell companion to syntax*, 2nd edition, Chapter 42. Oxford: Wiley-Blackwell.
- Brasoveanu, Adrian. 2008. Uniqueness effects in correlatives. In Atle Grønn (ed.), *Proceedings of Sinn und Bedeutung 12*, 47–65. Oslo: ILOS.
- Chomsky, Noam. 1973. Conditions on Transformations. In Stephen Anderson and Paul Kiparsky (eds.), *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston, 232–286.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland and Hans Martin Gärtner (eds), *Interfaces + Recursion = Language?*, 1–29. Berlin: Mouton de Gruyter.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria L. Zubizarreta (eds), *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, 133–166. Cambridge MA: MIT Press.
- Dočekal, Mojmír. 2005. *Vázáné proměnné v češtině*. Brno: Masaryk University dissertation.
- Drubig, Hans Bernhard. 2001. On the syntactic form of epistemic modality. Manuscript, University of Tübingen.
- Dyakonova, Marina. 2009. A phase-based approach to Russian free word order. Utrecht: LOT.
- Erteschik, Nomi. 1973. On the nature of island constraints. Doctoral dissertation, MIT, Cambridge, MA.
- Erteschik, Nomi and Shalom Lappin. 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6(1–3): 41–85.
- Grewendorf, Günter. 2015. Double fronting in Bavarian left periphery. In Ur Shlonsky (ed.), *The cartography of syntactic structures 10: Beyond functional sequence*, 232–252. Oxford: Oxford University Press.

Biskup and Šimík

- Grohmann, Kleanthes. 2003. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins.
- Heck, Fabian. 2008. *On pied-piping: Wh-movement and beyond*. Berlin: de Gruyter.
- Hirsch, Aron. 2016. A compositionl semantics for *wh-ever* free relatives. In Nadine Bade, Polina Berezovskaya, and Anthea Schöller (eds.), *Proceedings of Sinn und Bedeutung 20*, 341–358.
- Huang, James C.-T. 1982. Logical relations in Chinese and the theory of grammars. Cambridge, MA: MIT dissertation.
- Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara Partee (eds.), *Quantification in natural languages*, vol. II, 451–486. Dordrecht: Kluwer.
- Kratzer, Angelika. 1979. Conditional necessity and possibility. In Rainer Bäuerle (ed.), *Semantics from different points of view*, 117–147. Berlin: Springer.
- Lešnerová, Šárka and Karel Oliva. 2003. Česká vztážná souvětí s nestandardní strukturou. *Slovo a slovesnost* 64(4): 241–252.
- Manzini, Maria Rita. 1992. *Locality: A theory and some of its empirical consequences*. Cambridge, MA: MIT Press.
- McDowell, Joyce P. 1987. Assertion and modality. Los Angeles: University of Southern California dissertation.
- Rawlins, Kyle. 2013. (Un)conditionals. *Natural Language Semantics* 40(2): 111–178.
- Šimík, Radek. 2018. Ever free relatives crosslinguistically. In Uli Sauerland and Stephanie Solt (eds.), *Proceedings of Sinn und Bedeutung 22*, vol. 2, 375–392. Berlin: Leibniz-ZAS.
- Šimík, Radek. To appear. Free relatives. In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, and Thomas Ede Zimmermann (eds.), *Semantics companion*. Wiley-Blackwell.
- Ticio, Maria Emma. 2003. On the structure of DPs. Storrs, University of Connecticut PhD dissertation.
- Toman, Jindřich. 1998. A discussion of resumptives in Colloquial Czech. In Željko Bošković, Steven Franks, and William Snyder (eds.), *Formal Approaches to Slavic Linguistics 6: The Connecticut Meeting 1997*, 303–318. Ann Arbor, MI: Michigan Slavic Publications.

Petr Biskup, Radek Šimík
petr.biskup@hu-berlin.de, radek.simik@hu-berlin.de