Extraction from clausal adjuncts in Czech: A rating experiment*

Radek Šimík[1], Petr Biskup[2], Kateřina Bartasová[1], Markéta Dančová[3], Eliška Dostálková[1], Kateřina Hrdinková[1], Gabriela Kosková[1], Jaromír Kozák[1], Klára Lupoměská[1], Albert Maršík[1], Edita Schejbalová[1] & Illia Yekimov[1]

[1]Charles University, [2]University of Leipzig, [3]Palacký University

1. Introduction

This paper focuses on the issue of wh-extraction from clausal (tensed) adjuncts in Czech. Building on the observation that Czech clausal adjuncts are in some cases transparent for extraction (Lešnerová and Oliva 2003, Biskup and Šimík 2019), we have designed a naturalness-rating experiment with the aim to evaluate the adequacy of two competing approaches: the syntactic/semantic approach of Biskup and Šimík (2019) and the discourse-based approach of Abeillé et al. (2020).

According to Biskup and Šimík (2019), adjuncts in Czech are transparent for extraction if they are proposition-denoting clauses (CPs), syntactically and semantically akin to (embedded) questions. They are islands if they are entity-denoting complex nominals (often prepositional), akin to (free) relative clauses. The former scenario obtains if the adjunct is left-peripheral, the latter if it is right-peripheral or if the nominal structure is overtly represented. According to the discourse-based approach, left-peripheral clauses are informationally backgrounded, which makes them less transparent than the corresponding right-peripheral ones, which are focused. Nevertheless, as recently argued by Abeillé et al. (2020), even backgrounded constituents can be transparent as long as the extracted element is also backgrounded. We test this prediction by looking at the difference between the extraction of relative pronouns (backgrounded) vs. interrogative phrases (focused).

Our experimental results largely support the syntactic/semantic analysis of Biskup and Šimík (2019) and cannot be easily reconciled with the discourse-based approach of Abeillé et al. (2020). In addition, we discover effects not predicted by neither account.

*Besides NELS 52, this work was presented in the ERCEL research group (June 2021; Charles University) and at the workshop *Adverbial clauses: between subordination and coordination* (May 2022; University of Cologne). We received valuable feedback from James Brand, Sam Featherston, Jan Chromý, Radim Lacina, and Mikuláš Preininger.

The paper is structured as follows. In section 2 we provide some background and the relevant empirical pattern as reported in the literature. In 3 we discuss theoretical approaches applicable to the pattern, with special attention paid to Biskup and Šimík (2019) and Abeillé et al. (2020). Section 4 reports on our naturalness rating experiment. Finally, section 5 discusses the results and concludes.

2. Background

Clausal adjuncts have traditionally been considered islands for extraction because of the ungrammaticality of sentences like (1) (e.g. Huang 1982, Cinque 1990).

(1) *Who$_1$ did Mary cry [after John hit t$_1$]?

Yet over the years exceptions have frequently been reported (Grosu 1981, Deane 1991, Kluender 1998); for more examples and references see Chaves's (2021) survey, which is also the source of the examples below (p. 691).

(2) a. These are the pills OP$_1$ that Mary died [before she could take t$_1$].
b. The person who$_1$ I would kill myself [if I couldn't marry t$_1$] is Jane.
c. Which book$_1$ will Kim understand linguistics better [if she reads t$_1$]?

There is also a growing body of experimental evidence supporting the stance that at least some clausal adjuncts in some languages are indeed transparent for extraction (Dal Farra 2019, Kush et al. 2019, Bondevik et al. 2021). Below are two examples from Bondevik et al.'s (2021) experiment, involving the extraction of a contrastive topic (CT; manipulated by a preamble) out of a clausal adjunct. Example (3a) (if-clause) was rated as rather "good" while (3b) (because-clause) as "bad" (here *), which demonstrates that extraction from clausal adjuncts is in principle possible, but also that not all adjuncts are alike.

(3) Norwegian (Bondevik et al. 2021:229)

a. [...] men [$_{CT}$ takkekortene$_1$] blir hun skuffet [om de
but thank.you.cards.DEF becomes she disappointed if they
glemmer å sende ut t$_1$ med en gang].
forget to send out with one time
Literally: '[...] but the thank-you cards she will be disappointed if they forget to send out right away.'

b. *[...] men [$_{CT}$ vintertemperaturene$_1$] blir hun boende [fordi hun
but winter.temperatures.DEF becomes she living because she
liker t$_1$].
likes
Literally: '[...] but the winter temperatures she stays there because she likes.'

Extraction from clausal adjuncts is also attested in Czech, as first noticed and discussed – based on corpus evidence – by Lešnerová and Oliva (2003).

(4) Lešnerová and Oliva (2003:241)
Na každé zakoupené plyšové hračce je nálepka, kterou$_1$ [když dítě odevzdá
on every purchased plush toy is sticker which when child hands.in.3SG
t_1 v ZOO], obdrží navíc drobný dárek.
in ZOO receive.3SG in.addition small gift
'On every purchased plush toy there's a sticker such that when a child hands in the sticker in the ZOO, the child will receive a small gift in addition.'

Biskup and Šimík (2019), who offered a formal analysis of this type of extraction (see below), noticed that extraction is possible not just from conditional adjuncts, but also purpose clauses or correlatives. Crucially, however, extraction is only claimed to be grammatical if the adjunct appears to the left of the clause it modifies; compare the grammatical (a) examples with the ungrammatical (b) examples below.

(5) a. To je ten chlap, kterému$_1$ [co(koliv)$_2$ dáš t_2 t_1], to ztratí.
it is that man which.DAT what(ever).ACC give.2SG it lose.3
b. *To je ten chlap, kterému$_1$ ztratí [co(koliv)$_2$ dáš t_2 t_1].
it is that man which.DAT lose.3 what(ever).ACC give.2SG
(Intended:) 'This is the man such that he will lose whatever you give him.'

(6) a. To je ten řečník, kterého$_1$ [aby nalákali t_1], museli by
it is the speaker which.ACC in.order.SBJV.3 attract.PL must.PL SBJV.3
mít peníze.
have money.
b. *To je ten řečník, kterého$_1$ museli by mít peníze [aby
it is the speaker which.ACC must.PL SBJV.3 have money in.order.SBJV.3
nalákali t_1].
attract.PL
(Intended:) 'This is a speaker such that they need money in order to attract him.'

Biskup and Šimík (2019) further observed that when the adjunct clause has the form of a complex nominal, extraction is blocked. This is the case for the adjunct types mentioned above, as illustrated in (7) for the purpose adjunct, as well as for adjuncts that are introduced by complementizers whose internal structure involves a nominal layer. This is illustrated in (8) for a causal adjunct, whereby the Czech causal complementizer *protože* 'because' is composed of *pro* 'for', *to* 'that' (demonstrative), and *že* 'that' (declarative complementizer), suggesting that *protože*-clauses in Czech are in fact complex nominals embedded in a PP.

(7) a. *To je ten řečník, kterého$_1$ [na to, aby nalákali t_1],
it is the speaker which.ACC for that.DEM in.order.SBJV.3 attract.PL
museli by mít peníze.
must.PL SBJV.3 have money.
Intended: 'This is a speaker such that they need money in order to attract him.'

b. *To je ta knížka, kterou$_1$ [protože koupil t_1], už
it is that book which.ACC because bought..SG.M anymore
nemá peníze.
NEG.has.3SG money
Intended: 'This is a book such that he doesn't have any more money because he bought that book.'

3. Competing theories

The reported empirical pattern, or more precisely the correlation between adjunct position and its transparency, might strike an expert in island syntax as puzzling. If there is a left–right asymmetry, then it should go in the opposite direction: one would expect a right-peripheral structure to be more transparent for extraction than the corresponding left-peripheral structure, not conversely.

The asymmetry has received significant attention in the literature. For instance, Haegeman (2003, 2010) has accummulated plenty of evidence that left (called "peripheral") adjuncts are syntactically more complex than right ("central") adjuncts. If syntactic complexity positively correlates with opacity for extraction (the more complex a structure, the more opaque for extraction; e.g. Starke 2001, Abels 2012), the reported pattern is unexpected. The same holds for analyses based on movement: while in-situ structures may be transparent for extraction, they become islands once they move. This type of analysis, originally devised to account for the subject–object island asymmetry or in-situ vs. scrambled object asymmetry (so-called condition on extraction domains/CED; Huang 1982, Müller 2010), could in principle be applied in our case, too, but – alas – with the wrong result. Finally, it has been argued that informationally backgrounded (given, presupposed, topicalized) constituents are less transparent for extraction than focused constituents (Erteschik-Shir 1973, 2007, Goldberg 2006, Abeillé et al. 2020). The backgrounded vs. focused nature of left vs. right adjunct clauses makes sense for Czech, a language that heavily relies on word order in the expression of information structure. Basically all the existing theories of Czech information structure (see, e.g., Mathesius 1941, 1947, Daneš 1964, Sgall et al. 1986, Firbas 1992, Kučerová 2007, Biskup 2011) rely on the idea the backgrounded or given is placed before focused or new.[1] This is considered to hold for non-clausal and clausal constituents alike (Mathesius 1947, Chudobová 2011).

[1] See Šimík and Wierzba (2015, 2017) for a dissenting view, although one that does not make much of a difference for the case at hand.

3.1 Biskup and Šimík (2019)

Biskup and Šimík (2019) propose to account for the left-right asymmtry in the following way. There is no adjunct island condition per se (in Czech). What seems like adjunct islands are in fact complex nominal islands. If adjuncts are placed right-peripherally (or "centrally" in terms of Haegeman 2010), they are always nominals. More particularly, they are like free relatives in that they have an external nominal syntax as well as semantics – they denote entities of the relevant sort, e.g. a particular time interval in the case of temporal adjuncts (Hall and Caponigro 2011) or a particular (maximal) situation/world in the case of conditionals (Schlenker 2004). The referential semantics facilitates their integration into the event structure of the host clause, but the side effect is that their effectively nominal nature renders the adjunct clause a (complex NP) island. Biskup and Šimík (2019) propose a phase-based syntactic analysis of the complex nominal island fine-tuned for the purpose of NP languages, but, in principle, any complex nominal island analysis will do the job of deriving the observed empirical pattern.

Left adjuncts (called "peripheral" by Haegeman), on the other hand, are syntactically clausal/CPs. Semantically, they are not entities, but rather propositions, as in the standard Kratzerian account of conditionals (Kratzer 2012), in proposition-based analyses of correlatives (Bittner 2001, Brasoveanu 2008), or unconditionals (Rawlins 2013). The proposition restricts a potential modal or quantificational operator in the host clause. At the same time, it introduces the relevant kind of variable (e.g. time interval in temporal clauses) which is then referentially related to the corresponding participant in the event structure of the host clause (the event time).

If a clausal adjunct is headed by a nominal – possibly integrated in the complex adverbial complementizer (see (7)), it is predicted to be an island independently of its position.

3.2 Abeillé et al. (2020)

Abeillé et al. (2020) recently proposed a refinement of the discourse-based theory (going back to Erteschik-Shir 1973). They postulate the focus-background conflict (FBC) constraint, which states that "a focused element should not be part of a backgrounded constituent" (p. 3). In their experiment, they compare the extraction from subjects and objects on the one hand and the extraction of relative vs. interrogative pronouns on the other. They find that it is more acceptable to extract relative than interrogative pronouns from subjects. This is claimed to be predicted by the FBC constraint. Since relative pronouns, as opposed to interrogative ones, are backgrounded, they do not give rise to a violation of the FBC constraint.[2]

[2]There are some caveats in their theory. First, it is not clear why the grammatical categories of subject vs. object should correspond to the discourse categories backgrounded vs. focused (obviously, subjects can be focused and objects backgrounded). Second, it is unclear why the FBC constraint should be violated by an interrogative pronoun (focus) that is *extracted* from the subject (background). When it is extracted, it is not really a part of the extraction site, unless one assumes a copy or trace theory of filler-gap dependencies, which the authors do not (p. 1). We leave these caveats aside and concentrate on the predictions stated.

The FBC-based theory correctly predicts the acceptability of the extractions from the left adjuncts observed by Biskup and Šimík (2019). This is because all of them involve relative pronouns, whose backgrounded nature makes the extraction from a backgrounded clause possible. The theory further makes the prediction that extracting an interrogative pronoun/phrase should be less acceptable because it would consitute a violation of the FBC constraint. However, it also expects that the extraction from right (focused) adjuncts should be acceptable, irrespective of the relative vs. interrogative nature of the extracted element. This prediction is clearly problematic in the light of the data reported in Biskup and Šimík (2019). However, it is good to keep in mind that Biskup and Šimík's data rely on their own individual judgments, which, of course, might be wrong.

3.3 Filling in the gaps

We have so far identified three factors relevant for the extraction out of clausal islands in Czech: (i) the adjunct position, (ii) the presence of a nominal layer, and (ii) the nature of the extracted element. So far, the syntactic/semantic theory makes predictions about (i) and (ii) (but not (iii)) and the discourse-based theory about (i) and (iii) (but not (ii)). Is there a natural way to extend the theories and fill in the missing predictions?

Consider first the syntactic/semantic theory and factor (iii), i.e., the relative vs. interrogative contrast. Following Starke (2001), Abels (2012) one could work on the assumption that the larger (featurally more specific) a structure is, the more mobile it is. There are good reasons to assume that relative pronouns are structurally more specific than interrogative pronouns. It holds that a relative pronoun can be a structural superset of the interrogative, but not conversely. A paradigmatic example is the Bulgarian relative pronoun *kojto* 'which.REL', which is based on the interrogative *koj* 'which.INTER' (Rudin 2009; see also Mitrović 2016 for Slovenian or Daskalaki 2020 for Greek). To the best of our knowledge, the opposite pattern is not documented. A syntactic (structure-based) analysis would therefore expect relative pronouns to be more mobile (to be able to escape wh-/adjunct-islands) than interrogative ones.

Let us now turn to the discourse-based theory and factor (ii), i.e. the presence/absence of a nominal layer on the adjunct. It seems natural to assume that if an adjunct is further embedded in a nominal phrase (e.g. *when…* vs. *the time when…*), it becomes more backgrounded and therefore more difficult to extract from. This aligns with Erteschik-Shir and Lappin's (1979) assumptions, who consider headed relative clauses islands for extraction, unless they occur in specific syntactic environments, as in pivots of existential constructions.

4. Experiment

4.1 Questions and hypotheses

In our experiment, we investigated the naturalness of A′-extraction from clausal adjuncts in Czech. Below we list the main research questions (Q) and the corresponding hypotheses associated with the two types of theories under discussion: syntactic/semantic (Biskup and

Šimík 2019, complemented by Starke 2001/Abels 2012-style approach to locality) and discourse-based (Abeillé et al. 2020).

Q1 What is the difference between the extraction from left (peripheral) vs. right (central) adjuncts?

Q2 What is the difference between the extraction of relative vs. interrogative pronouns or phrases?

Q3 What is the difference between the extraction from clausal vs. complex nominal adjuncts?

The syntactic/semantic theory predicts that the only completely natural (grammatical) extraction should be the one of relative pronouns from left clausal adjuncts without a nominal layer. Extraction from right and/or nominally headed adjuncts is predicted to be unnatural. Extraction of interrogative phrases could be marginally acceptable from left adjuncts without a nominal layer, corresponding to a weak island violation.

The discourse-based theory predicts that the extraction of relative pronouns and interrogative phrases from right adjuncts (without a nominal layer) should be at least as acceptable as the extraction of relative pronouns from left adjuncts (without a nominal layer). Extraction of interrogative phrases from left adjuncts is expected to be less natural, just as extraction from nominally headed adjuncts.

4.2 Design and materials

In line with the above questions and hypotheses, we manipulated three variables in a fully crossed $2 \times 2 \times 2$ design: the POSITION of the adjunct (left vs. right), the EXTRACTED ELEMENT (relative vs. interrogative), and the NOMINAL layer on top of the adjunct (absent vs. present). Table 1 provides schematic stimuli in all the conditions. Example (8) exemplifies the actual stimuli (item 11), where (a)/(b) formed a sentence together with (i)/(ii).[3]

(8) a. rel
Znám písničku, kterou…
know.1SG song which.ACC
'I know a song such that…'

b. inter
Nevím, kterou písničku
NEG.know.1SG which song.ACC
'I don't know which song is such that…'

(i) left+abs/pres
[{když / ve chvíli, kdy} poslouchášt], lépe se soustředíš.
when at time when listen.2SG better REFL concentrate.2SG

[3]All the materials and a preregistration can be accessed here: `https://osf.io/zu9d4/?view_only=e813eb862f4b4052ad55c2f91659048d`.

(ii) right+abs/pres
se lépe soustředíš [{když / ve chvíli, kdy} poslouchášt].
REFL better concentrate.2SG when at time when listen.2SG
'you can concentrate better when you listen to it.'

stimulus (schematic)	POSITION	EXTRACTED	NOMINAL
song which [when hear t] relax	left	rel	abs
song which [at time when hear t] relax	left	rel	pres
wonder which song [when hear t] relax	left	inter	abs
wonder which song [at time when hear t] relax	left	inter	pres
song which relax [when hear t]	right	rel	abs
song which relax [at time when hear t]	right	rel	pres
wonder which song relax [when hear t]	right	inter	abs
wonder which song relax [at time when hear t]	right	inter	pres

Table 1: Set of schematic stimuli for a single item

We created 48 experimental items and complemented them with 64 fillers (which contained a number of smaller experiments manipulating factors that could potentially be relevant or confounding for the critical manipulations; see discussion). Of the 48 items, all involved the wh-word *který* 'which', which is very common in both relatives and interrogatives. In interrogatives, *který* was accompanied by a nominal which corresponded to the relative clause nominal head. Given the growing and converging evidence for the so-called matching analysis of relative clauses, whereby the relative wh-word is followed by an elided nominal under identity with the head, the interrogative and relative wh-clauses could even be considered string-identical, differing only in the critical manipulation. Finally, we used embedded interrogatives, to minimize the difference between the relative vs. interrogative condition – both in terms of embedding and the overall illocutionary force (all stimuli were statements). Many other properties not central to our research questions varied across the items, to cover more empirical ground and to make the stimuli less tiring for the participant. The extracted element was mostly an object (33 items), sometimes an adverbial (10 items) or a subject (5 items). It was mostly in singular (41 items). The adjunct was mostly a conditional or temporal clause, using a variety of complementizers, in 9 cases it was a purpose clause. The nominal used in the pres condition of the NOMINAL factor was mostly a suitable noun ('moment', 'case', etc.) and sometimes a demonstrative ('this/that'). In the majority of cases (43 items) the clause modified by the adjunct clause contained no pronominal element (or a subject *pro*) bound by the extracted wh-expression, mainly to avoid a potential weak crossover effects in the right adjunct condition.

4.3 Procedure and participants

The experiment was coded, pseudo-randomized, distributed on lists using the Latin Square design, and eventually administered with the help of the L-Rex software (Starschenko and

Wierzba 2021). The participants were instructed to rate the naturalness of the stimuli (always a single complex sentence) on the scale 1 (completely unnatural) to 7 (completely natural). Only the extremes of the scale were labeled. The procedure was exemplified by the means of two examples – one natural, the other unnatural – where the naturalness contrast was qualitatively unrelated to the main manipulations. Each participant rated 112 stimuli (48 experimental and 64 filler items). The pace was individual and most participants completed the experiments in 20 to 30 minutes. 96 participants (non-experts) completed the experiment, amounting to 576 ratings per each unique condition. They were recruited on an individual basis by the coauthors of this paper and completed the experiment in their homes at their own computers.

4.4 Results

The results are visualized by the bar plots in Figure 1. Each bar plot indicates the proportional distribution of the individual ratings for one of the eight conditions (see section 4.2). The darker the plot, the more natural the condition. The 50% line cuts through the median rating and the tallest shade in each bar indicates the modus (the most frequent rating in the condition). For instance, for the extraction of a relative pronoun from a left clausal adjunct (nominal absent) – the top left bar – the median rating was 6 and modus 7. That means that it was rated as very natural. Compare this with the corresponding extraction from right adjunct (the bar next to it), where the median was 2 and modus 1. That condition was rated as very unnatural.

We fitted a cumulative link mixed model with random intercepts for participants and items, using the clmm function of the ordinal package (Christensen 2013) of R (R Core Team 2021), to estimate the effects of the manipulated variables and their interactions on the dependent variable (rating). All three variables were sum coded, allowing us to estimate their main effects.

The main effects of all the variables as well as all their mutual interactions (two-way and three-way) turned out significant. Extracting a relative pronoun is more natural than extracting an interrogative phrase ($z = 23.125, p < .0001$); extraction from a left adjunct is more natural than from a right adjunct ($z = 38.204, p < .0001$); finally, extracting from a clause (abs) is more natural than from a complex nominal (pres) ($z = 12.874, p < .0001$). The three-way interaction ($z = 3.936, p < .0001$), which in a way subsumes all the two-way interactions, can be formulated as follows. Extraction from right adjuncts is very clearly the least natural condition in all four subcases, with the modus or even median rating at 1. How much better the extraction from the left adjunct is depends on the other two factors in the following way: the naturalness increase is more pronounced for relative than for interrogative elements and it is more pronounced if the nominal layer is absent than when it is present.

5. Discussion and conclusion

Overall, the syntactic/semantic approach provides a better fit of the data than the discourse-based approach. It correctly predicts that the extraction of relative pronouns from the left

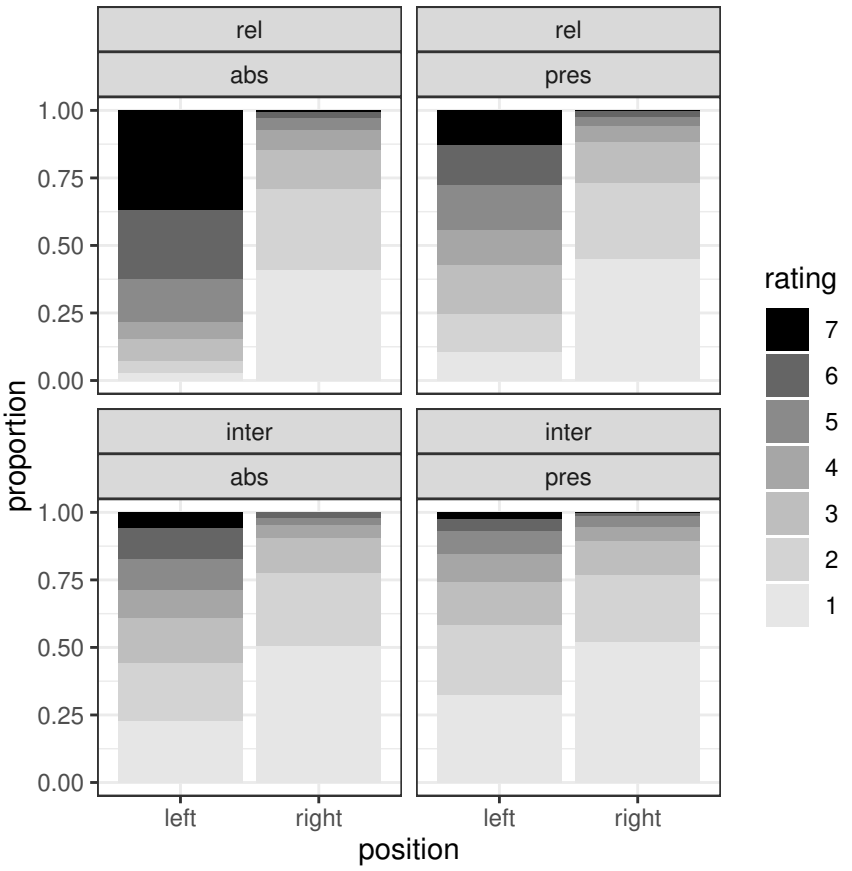

Figure 1: Proportions of naturalness ratings on the scale 1 = unnatural to 7 = natural

adjuncts without a nominal layer is the only completely natural condition (cf. the aforementioend three-way interaction in the expected form). With the median of 6 (modus 7), it is comparable to extraction from complement clauses (filler exp.; median 5, modus 6). The prediction of the discourse-based approach, on the other hand, is not borne out. Extraction from right adjuncts is generally rated as very unnatural (median 1 or 2, modus 1) and directly comparable to the extraction from another type of strong island (filler exp.; median 2, modus 1).[4] This is at odds with the prediction that the extraction from right adjuncts should be at least as natural as the extraction of relative pronouns from left adjuncts.

It is good to notice that the effect of POSITION is very close to being categorical: the rating of the right condition is at or very close to floor. This interpretation is supported by the results of our filler experiments, where (i) extractions from another type of strong island received comparable ratings and (ii) linearly long extractions from non-islands imposed only a very minor penalty. The categorical effect can be considered to reflect the violation of a grammatical constraint, in line with the syntactic/semantic analysis, and not just a processing issue.

[4]The strong island used in the fillers was a complement clause headed by a demonstrative (e.g. 'I'd like to know which room$_1$ she claimed (*that.DEM) that.COMP her boss redecorated t$_1$.').

Both approaches correctly capture the decreased naturalness of the extraction of interrogative phrases (effect of EXTRACTED element). The fact that this effect is less pronounced than the effect of POSITION follows naturally from both approaches. The syntactic one can resort to the notion of a weak island; the discourse-based one to processing issues.

Both approaches also correctly capture the decreased naturalness of extractions from complex nominals (effect of NOMINAL layer). What is unexpected for the syntactic approach is the relative naturalness of the extraction from overt complex nominals: extraction of relative pronouns from left-positioned complex nominal adjuncts reaches the median of 4 (modus 3), which is clearly a better rating than the observed extraction from right adjuncts (median 2, modus 1), which Biskup and Šimík (2019) also consider to be complex nominals. This is even more pressing in the light of a filler experiment, which showed that left-positioned complex nominal adjuncts are less natural (than clausal ones) independently of extraction. The syntactic/semantic analysis could in principle be upheld if what mattered was not just the nominal nature, but the particular *kind* of nominal: while right adjuncts are, by hypothesis, donkey-anaphoric definites (Hirsch 2016, Šimík 2018), many of the overt nominals we used in our items might well have been interpreted as indefinite. If definiteness matters for extraction (see e.g. Kluender 1998), we might have a handle on the unexpected difference between extraction from left complex nominals and right clausal adjuncts.

To conclude, we have demonstrated that wh-extraction from clausal adjunct is natural in Czech, adding to the recent literature (e.g. Bondevik et al. 2021), but only in very specific conditions, namely when a relative pronoun is extracted from a left adjunct that lacks an overt nominal layer, i.e., when it is (by assumption) not a complex nominal. This overall naturalness profile is predicted by the syntactic/semantic approach of Biskup and Šimík (2019), but not by the discourse-based approach of Abeillé et al. (2020). Our experiment further revealed significant distinctions among the other conditions, which could all be considered unnatural. The gradient character of these distinctions points to processing rather than grammatical issues and is thus more in line with the discourse-based account. More research is needed to disentangle the sources of the observed naturalness distinctions.

References

Abeillé, Anne, Barbara Hemforth, Elodie Winckel, and Edward Gibson. 2020. Extraction from subjects: Differences in acceptability depend on the discourse function of the construction. *Cognition* 204:104293. URL `https://doi.org/10.1016/j.cognition.2020.104293`.

Abels, Klaus. 2012. The Italian left periphery: A view from locality. *Linguistic Inquiry* 43:229–254. URL `https://doi.org/10.1162/LING_a_00084`.

Biskup, Petr. 2011. *Adverbials and the phase model*. Amsterdam: John Benjamins. URL `https://doi.org/10.1075/la.177`.

Biskup, Petr, and Radek Šimík. 2019. Structure of conditional and (cor)relative clauses: New evidence from locality. In *NELS 49: Proceedings of the 49th Annual Meeting of*

the North East Linguistic Society, Volume 1, ed. by Maggie Baird and Jonathan Pesetsky, 135–144. Amherst, MA: GLSA Publications.

Bittner, Maria. 2001. Topical referents for individuals and possibilities. In *SALT 11: Proceedings from the 11th Conference on Semantics and Linguistic Theory*, ed. by Rachel Hastings, Brandon Jackson, and Zsofia Zvolenszky, 36–55. Ithaca, NY: Cornell University. URL https://doi.org/10.3765/salt.v11i0.2854.

Bondevik, Ingrid, Dave Kush, and Terje Lohndal. 2021. Variation in adjunct islands: The case of Norwegian. *Nordic Journal of Linguistics* 44:1–32. URL https://doi.org/10.1017/S0332586520000207.

Brasoveanu, Adrian. 2008. Uniqueness effects in correlatives. In *Proceedings of Sinn und Bedeutung 12*, ed. by Atle Grønn, 47–65. Oslo: ILOS.

Chaves, Rui P. 2021. Island phenomena and related matters. In *Head-Driven Phrase Structure Grammar: The handbook*, ed. by Stefan Müller, Anne Abeillé, Robert D. Borsley, and Jean-Pierre Koenig. Berlin: Language Science Press. URL https://doi.org/10.5281/zenodo.5599846.

Christensen, Rune Haubo Bojesen. 2013. Ordinal: Regression models for ordinal data. URL http://www.cran.r-project.org/package=ordinal/, r package version 2013.9-30.

Chudobová, Martina. 2011. Větosled v současné češtině [Clause order in contemporary Czech]. Master's thesis, Charles University, Praha. URL http://hdl.handle.net/20.500.11956/33727.

Cinque, Guglielmo. 1990. *Types of A′-dependencies*. Cambridge, MA: MIT Press.

Dal Farra, Chiara. 2019. Variability in adjunct islands: An experimental study in Italian. In *WCCFL 36: Proceedings of the 36th West Coast Conference on Formal Linguistics*, ed. by Richard Stockwell, Maura O'Leary, Zhongshi Xu, and Z. L. Zhou, 115–120. Somerville, MA: Cascadilla Proceedings Project. URL http://www.lingref.com/cpp/wccfl/36/paper3453.pdf.

Daneš, František. 1964. A three-level approach to syntax. In *Travaux Linguistiques de Prague 1: L'École de Prague d'aujourd'hui*, ed. by Josef Vachek, 225–240. Praha: Éditions de l'Académie Tchécoslovaque des Sciences.

Daskalaki, Evangelia. 2020. Types of relative pronouns. In *Syntactic architecture and its consequences II: Between syntax and morphology*, ed. by András Bárány, Theresa Biberauer, Jamie Douglas, and Sten Vikner, 277–296. Berlin: Language Science Press. URL https://doi.org/10.5281/zenodo.4280653.

Deane, Paul D. 1991. Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2:1–63. URL https://doi.org/10.1515/cogl.1991.2.1.1.

Erteschik-Shir, Nomi. 1973. On the nature of island constraints. Doctoral dissertation, MIT, Cambridge, MA. URL http://dspace.mit.edu/handle/1721.1/7582.

Erteschik-Shir, Nomi. 2007. *Information structure: The syntax-discourse interface*. Oxford: Oxford University Press.

Erteschik-Shir, Nomi, and Shalom Lappin. 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6:41–85. URL https://doi.org/10.1515/thli.1979.6.1-3.41.

Firbas, Jan. 1992. *Functional sentence perspective in written and spoken communication*. Cambridge: Cambridge University Press.

Goldberg, Adele. 2006. *Constructions at work*. Oxford: Oxford University Press.

Grosu, Alexander. 1981. *Approaches to island phenomena*. Amsterdam: North-Holland.

Haegeman, Liliane. 2003. Conditional clauses: External and internal syntax. *Mind & Language* 18:317–339. URL https://doi.org/10.1111/1468-0017.00230.

Haegeman, Liliane. 2010. The internal syntax of adverbial clauses. *Lingua* 120:628–648. URL https://doi.org/10.1016/j.lingua.2008.07.007.

Hall, David P., and Ivano Caponigro. 2011. On the semantics of temporal *when*-clauses. In *SALT 20: Proceedings from the 20th Conference on Semantics and Linguistic Theory*, ed. by Nan Li and David Lutz, 544–563. Ithaca, NY: CLC Publications. URL https://doi.org/10.3765/salt.v20i0.2566.

Hirsch, Aron. 2016. A compositional semantics for *wh-ever* free relatives. In *Proceedings of Sinn und Bedeutung 20*, ed. by Nadine Bade, Polina Berezovskaya, and Anthea Schöller, 341–358. URL http://semanticsarchive.net/sub2015/SeparateArticles/Hirsch-SuB20.pdf.

Huang, Cheng-Teh James. 1982. Logical relations in Chinese and the theory of grammars. Doctoral dissertation, MIT, Cambridge, MA. URL http://dspace.mit.edu/handle/1721.1/15215.

Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. In *Syntax and semantics 29: The limits of syntax*, ed. by Peter Culicover and Louise McNally, 241–279. San Diego, CA: Academic Press. URL https://doi.org/10.1163/9789004373167_010.

Kratzer, Angelika. 2012. *Modals and conditionals: New and revised perspectives*. Oxford: Oxford University Press.

Kush, Dave, Terje Lohndal, and Jon Sprouse. 2019. On the island sensitivity of topicalization in Norwegian: An experimental investigation. *Language* 95:393–420. URL https://doi.org/10.1353/lan.2019.0051.

Kučerová, Ivona. 2007. The syntax of givenness. Doctoral dissertation, MIT, Cambridge, MA. URL http://hdl.handle.net/1721.1/39580.

Lešnerová, Šárka, and Karel Oliva. 2003. Česká vztažná souvětí s nestandardní strukturou. *Slovo a slovesnost* 64:241–252. URL http://sas.ujc.cas.cz/archiv.php?art=4129.

Mathesius, Vilém. 1941. Základní funkce pořádku slov v češtině. *Slovo a slovesnost* 7:169–180. URL http://sas.ujc.cas.cz/archiv.php?lang=en&art=413.

Mathesius, Vilém. 1947. *Čeština a obecný jazykozpyt*. Praha: Melantrich.

Mitrović, Moreno. 2016. A relative syntax and semantics for Slovenian. In *Formal studies in Slovenian syntax: In honor of Janez Orešnik*, ed. by Franc Marušič and Rok Žaucer, 221–252. Amsterdam: John Benjamins. URL https://doi.org/10.1075/la.236.10mit.

Müller, Gereon. 2010. On deriving the CED from the PIC. *Linguistic Inquiry* 41:35–82. URL http://www.jstor.org/stable/40606828.

R Core Team. 2021. R: A language and environment for statistical computing. URL https://www.r-project.org, vienna: R Foundation for Statistical Computing.

Rawlins, Kyle. 2013. (Un)conditionals. *Natural Language Semantics* 40:111–178. URL https://doi.org/10.1007/s11050-012-9087-0.

Rudin, Catherine. 2009. The Bulgarian relative marker *-to*. In *A linguist's linguist: Studies in South Slavic linguistics in honor of E. Wayles Browne*, ed. by Steven Franks, Vrinda Chidambaram, and Brian Joseph, 403–422. Bloomington, IN: Slavica Publishers.

Schlenker, Philippe. 2004. Conditionals as definite descriptions (A referential analysis). *Research on Language and Computation* 2:417–462. URL https://doi.org/10.1007/s11168-004-0908-2.

Sgall, Petr, Eva Hajičová, and Jarmila Panevová. 1986. *The meaning of the sentence and its semantic and pragmatic aspects*. Dordrecht: Reidel.

Šimík, Radek. 2018. Ever free relatives crosslinguistically. In *Proceedings of Sinn und Bedeutung 22*, ed. by Uli Sauerland and Stephanie Solt, volume 2, 375–392. Berlin: Leibniz-Center General Linguistics. URL https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/112.

Šimík, Radek, and Marta Wierzba. 2015. The role of givenness, presupposition, and prosody in Czech word order: An experimental study. *Semantics & Pragmatics* 8:1–103. URL https://doi.org/10.3765/sp.8.3.

Šimík, Radek, and Marta Wierzba. 2017. Expression of information structure in West Slavic: Modeling the impact of prosodic and word-order factors. *Language* 93:671–709. URL https://doi.org/10.1353/lan.2017.0040.

Starke, Michal. 2001. Move dissolves into merge: A theory of locality. Doctoral dissertation, University of Geneva, Geneva. URL http://ling.auf.net/lingbuzz/000002.

Starschenko, Alexej, and Marta Wierzba. 2021. L-Rex (beta). URL https://github.com/2e2a/l-rex/, computer software.

Radek Šimík, Petr Biskup, Kateřina Bartasová, Markéta Dančová, Eliška Dostálková, Kateřina Hrdinková, Gabriela Kosková, Jaromír Kozák, Klára Lupoměská, Albert Maršík, Edita Schejbalová & Illia Yekimov
radek.simik@ff.cuni.cz, biskup@rz.uni-leipzig.de