Scalar particles and conditionals: Czech Conditional Inference Trees

FDSL 18

Mojmír Dočekal & Jarmila Černá, Masaryk university

2025-09-22

Intro

- existing theories of Negative Polarity Items (NPIs) predict that weak NPIs are licensed in the antecedents of conditionals (indicative or counterfactual) since:
  1. the antecedents are either treated as (Strawson) Downward Entailing (SDE): Heim (1984), Kadmon and Landman (1993), Von Fintel (1999);
  2. or as non-monotonic: Barker (2018);
- we adopt the second strategy, Barker (2018), as it is more compatible with the modern semantic and philosophical literature on conditionals
  - which, at least since Stalnaker (1976), treats conditionals as non-monotonic;

- however, it has been observed, Schwab and Liu (2022a), that some kinds of NPIs (attenuating NPIs like English *all that*) are preferred in counterfactual antecedents of conditionals (experimental evidence: Schwab and Liu (2022b));
- similar observations can be made for Czech NPIs, but with different kinds of NPIs and patterns
- we focus on two kinds of Czech scalar particles:

1. NPIs of the type *sebe*-comparative denoting a low degree of a property (e.g., *sebemenší* 'the slightest' in (1) and (2), see also Dočekal and Juřen (2023)), referred to as *sebe-low*;

(1) Budeme vděčni, pokud přispějete
be.FUT.1PL grateful.M.PL if contribute.FUT.2PL
sebemenším finančním darem.
sebe-low.INS financial.INS gift.INS
'We will be grateful if you contribute the slightest financial gift.' indicative

(2) Nacisté by s ním rychle
Nazis.NOM would.SUBJ with him quickly
zatočili, kdy-by pojali
deal.PFV.3PL when-SUBJ.1PL take.PFV.3PL
sebemenší podezření o jeho
sebe-low.ACC suspicion.ACC about his.LOC
stycích s námi.
contacts.LOC with us.INSTR
'If the Nazis had the slightest suspicion about his contacts with us, they would quickly deal with him.'

2. strong scalar particles of the type *sebe*-comparative denoting a high degree of a property (e.g., *sebevětší* 'the highest' in (3)), referred to as *sebe-top*;

(3) I kdy-by měli sebevětší
even when-SUBJ.3PL have.PST.3PL sebe-top.ACC
příjem, bude jim to málo.
income.ACC be.3SG.FUT them.DAT it.NOM little
'Even if they had the highest income, it would be too little for them.' even if

- *Sebe-low* in Czech are NPIs with a likelihood profile corresponding to the English NPI *even one* (see Crnič (2011))
- intuitively, they seem to be preferred in the antecedents of indicative and subjunctive conditionals but not in the antecedents of *even if* conditionals;
- such preferences are hard to capture in current theories of NPIs;
    - *sebe-low* are minimizers/NPIs, so they should be licensed in all kinds of conditionals (indicative, subjunctive, *even if*);

- our research goal is to map Czech NPIs' preferences in conditionals and provide a theoretical explanation of the observed patterns;
- to achieve this goal, we conducted a corpus study (examples (1), (2), and (3) are taken from SketchEngine: Kilgarriff et al. (2014))
- the results were interpreted via non-parametric statistical recursive partitioning methods;
- finally, the consequences for the theories of conditionals are discussed.

- joint work with Jarmila Černá (Masaryk University)

Corpus study

- from SketchEngine corpus csTenTen23, we extracted conditionals with antecedents containing the NPI *sebe*-low/top
  - (CQL query: [lemma="jestli" | lemma="kdyby" | lemma="pokud"] []{1,5} [word="sebevětší" | lemma="sebemenší"], resulting in 1040 hits)
- we then obtained a random sample of 500 such sentences;
- the sentences were manually annotated for the type of conditional (indicative, counterfactual, *even if*), the type of NPI (*sebe*-low/*sebe*-top), and the presence of local negation in the antecedent

- with all variables being of the boolean type;
- the descriptive results are shown in Figure 1.

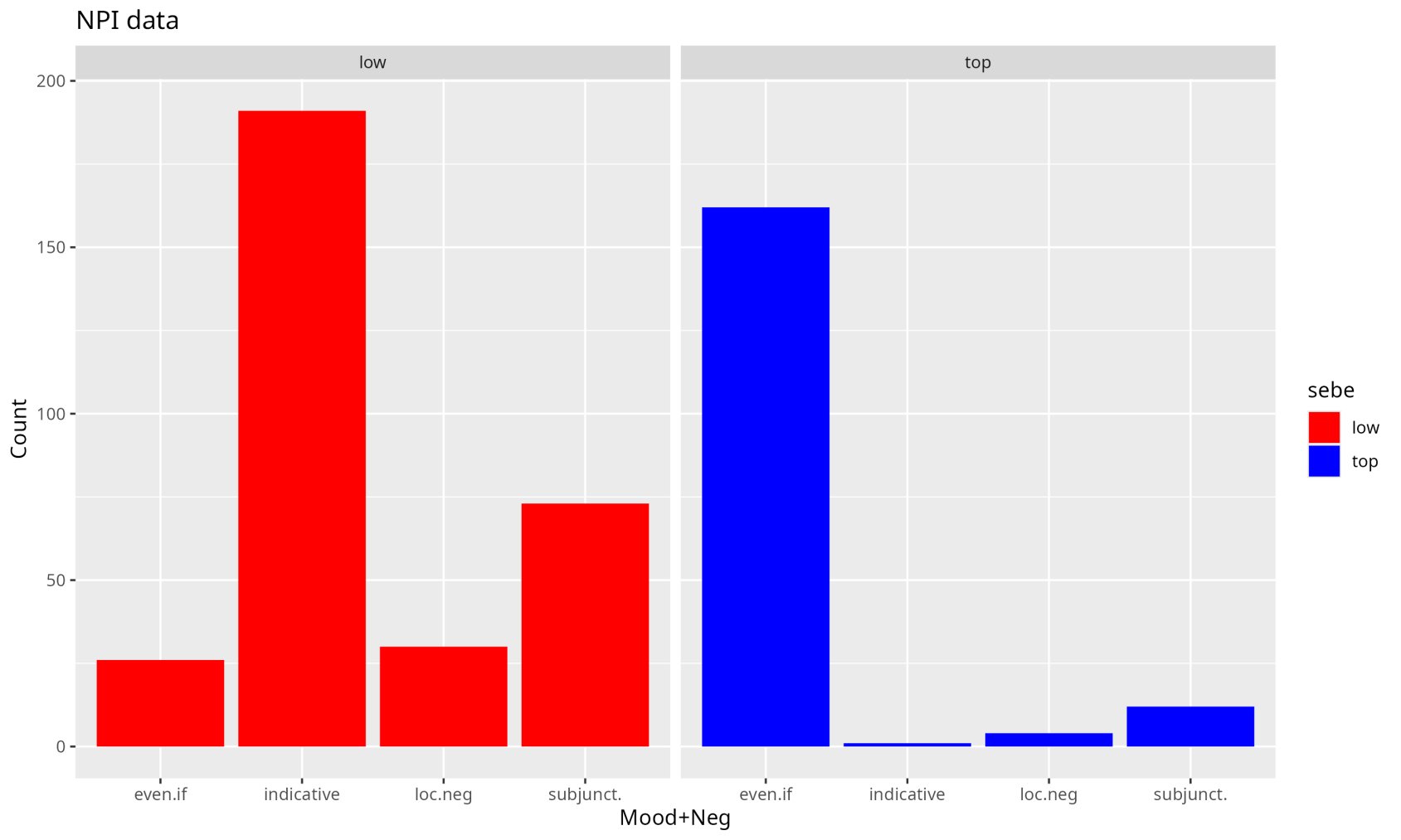

Figure 1: Figure 1: Descriptive results

- the inferential statistics were conducted using conditional inference trees (CIT, see Levshina (2020) a.o.), a non-parametric statistical recursive partitioning method;
- the analysis used the R (R Core Team (2024)) package *party* (Strobl et al. (2008));
- Figure 2 shows the individual tree fitted to the data:
- the dependent variable is the type of scalar particle (*sebe*-low/*sebe*-top)
- the independent variables are the type of conditional (indicative, counterfactual, *even if*) and the presence of local negation in the antecedent.

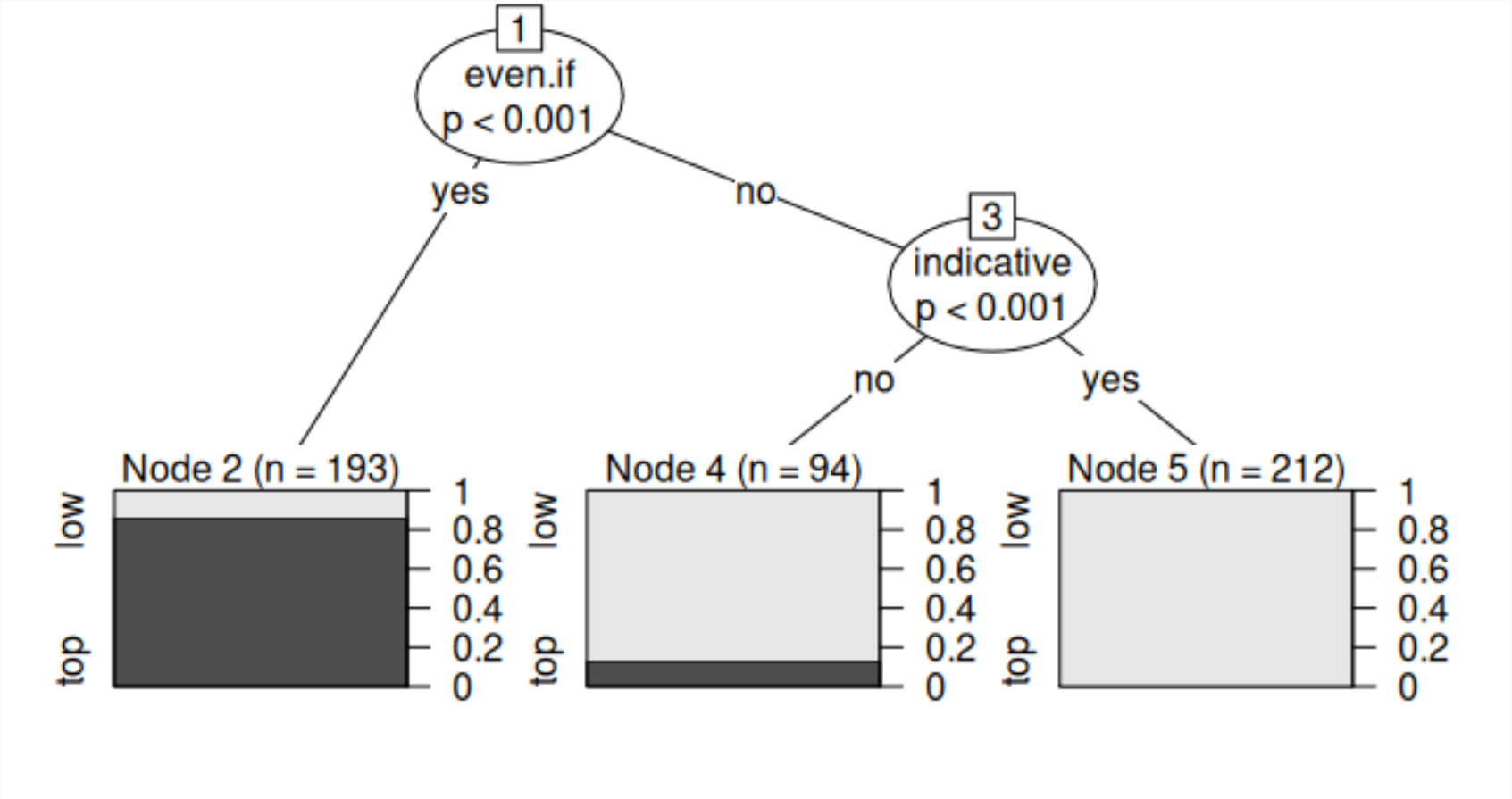

Figure 2: Figure 2: Conditional inference tree

- the top split (Node 1) is based on the variable *even if*: the predictor (*even if* type of conditional vs. indicative+subjunctive) is the most important and significant ($p<0.001$) predictor of the type of scalar particles;
- the data are then split into two subsets:
- the left subset (Node 2) contains sentences with conditionals of the type *even if*;
- the terminal Node 2 contains 193 sentences, with the majority (0.86) being *sebe*-top;

- the next split is based on the variable *indicative* (Node 3 – also significant): conditionals in the subjunctive mood (Node 4) contain 94 conditionals, with the majority (0.87) being *sebe*-low;
- the right subset (Node 5) contains 212 indicative conditionals, nearly all of which are *sebe*-low (0.98).

- the predictive power of the models is very satisfactory:
  - the CIT has a classification accuracy of 0.92 (with a baseline value of 0.51),
  - and the concordance index C is 0.94
- the results of the CIT were corroborated by a logistic regression model (not shown here – APPENDIX?) with the same dependent and independent variables, yielding very similar results.

Interpretation

- we adopt Stalnaker's unified theory of conditionals, which uniformly treats indicative and subjunctive conditionals as non-monotonic:
  - the conditional $\phi > \psi$ is true at w iff the closest (to w) most normal ϕ-world is a ψ-world)
- indicative conditionals pragmatically signal that the ϕ-world is in the context set (common ground), while the speaker using subjunctive conditionals does not guarantee that the ϕ-world is in the context set.
  - Stalnaker's approach selected for simplicity, in Lewis' approach it would work the same;

- the formalization of (1) would be:

(4) $f(w, [\![\text{you contribute } d\text{—donation in } w]\!]) \in \{w \in W : \text{we are happy in } w\}$

- and analogously for (2) and (3).

- the NPI *sebe*-low is licensed in the antecedent of indicative/subjunctive conditionals after Barker (2018)
  - since the wide scope interpretation of the NPI fails to entail the narrow scope of the NPI w.r.t. other operators – true for (1)/(2)/(3) where the PI occurs in the antecedent of a non-monotonic conditional;
  - necessary condition

- but the particles also have a pragmatic component
- we follow Crnič (2011) (general framework) and Dočekal and Juřen (2023) (for Czech) in explaining the NPI *sebe*-low as a scalar particle with unlikelihood presupposition having wide scope:

(5) $[\text{even}](\phi > \psi)$ ($p \lhd_C q$: the prejacent p is at most as likely as focus induced alternatives q in C.

- the prejacent is either as likely (non-monotonic analysis) or less likely than the alternatives

- after Crnič (2011) (and many before): likelihood in monotonic contexts can be translated to entailment
- in (S)DE lower values of scale entail (and are less likely) than higher values since the likelihood follows entailment: logically stronger is less likely:

(6) a. $(p \wedge q) \models p / \models q$

b. $P(p \cap q) \lhd P(p) / \lhd P(q)$

- For (1), (2), and (3), the presupposition is satisfied since in SDE $give(d) \rightarrow give(d')$ (d is a minimal amount, d' alternative degrees) and therefore the prejaccent is logically stronger/less likely than alternatives $give(d')$ (for all $d' > d$).

- core formalization ($>$ is Stalnaker's conditional)

(7) $[\text{even}](\text{you give } d > \psi)$
(7) $\rightarrow (\text{you give } d' > \psi)$ (for all $d' > d$)

- intuitively, the scalar particle measures the likelihood/world-distance (minimal) from the actual world (indicative: still the actual world, subjunctive some possible accessible world)

- we treat the NPI *sebe*-top as a scalar particle with the same unlikelihood presupposition (the most unlikely, entailing all focus alternatives);
- since it can appear with top-of-the-scale elements:

(8) Už od roku 1992 vyrábíme kvalitní
since from year 1992 produce.PFV.1PL quality.M.PL
české plastové kompostéry, které
Czech plastic.ACC composter.ACC which
vydrží sebevětší horko i mráz.
withstand.PFV.3PL self-great heat.ACC and frost.ACC
'We have been producing high-quality Czech plastic composters that can withstand even the greatest heat and frost since 1992.'

- but the prediction then is that we should not find *sebe*-top in the antecedent of conditionals since only the bottom of the scale is logically stronger there
  - fits for indicative, and partially for subjunctive conditionals
- the puzzle:

(9) What explains the *sebe*-top in the antecedent of *even-if* conditionals?

- working idea: pragmatics of *even if*

Side note on the semantics of *even if*

- very short since not the main topic here: we follow Gärdenfors (1988) and Crupi and Iacona (2022) in treating the *even if* concessive conditionals as a conjunction of two conditionals:

(10) A concessive is a sentence of the form $(p \rightarrow q) \wedge (\neg p \rightarrow q)$. Crupi and Iacona (2022, def. 1, simplified)

- truth-table:

p	q	$(p \rightarrow q) \wedge (\neg p \rightarrow q)$
F	F	F
F	T	T
T	F	F
T	T	T

- the formula is in fact equivalent to just q but there is a pragmatic component, though we take this as the correct core semantics of *even if* conditionals (concessive conditionals)
  - the speaker communicates that p would normally imply $\neg q$

- unlike König (1988), König (2006), Mizuno (2008) and Milosavljević and Milosavljević (2024) we don't treat rectifying, and rhetorical "concessives", (11-a), (11-b) from Milosavljević and Milosavljević (2024) as real concessives
  - they don't have the core semantics of concessives but are simply conjunction $p \wedge q$ with some pragmatic component

(11) My family was unusual in that they only spoke Wolof, although my 19-year-old "sister" remembered a smattering of French from school.

(11) Even though this solution would be harmful to our enemies, the damage done to us would be even greater.

- but both at least introduced-if concessives (next slides) with *i kdyby* are standard concessives with the core semantics of conjunction of two conditionals

- the intuitive idea (can be formalized in different ways): the speaker signals that p would normally imply $\neg q$ but q is true regardless of whether p is true or false
- for such content, the sebe-top is appropriate: the speaker signals that p is very unlikely (the most unlikely on the scale of likelihood) but q is true regardless of whether p is true or false
- go to slide 40

Different conjunctions for each meaning?

- real concessive meaning (and the two others non-concessive meanings) too – default strategy for all *i kdyby*:

(12)

I	kdyby	pršelo,	budeme	grilovat.
even if		rained.SUBJ.3SG	be.FUT.1PL	grill.INF

'Even if it rains, we will grill.'

- the first conjunction meaning (“rectifying”): *ačkoliv* ‘although’:

(13)

Byl	to	úspěšný	den,	ačkoliv
was	it.NOM	successful.M.SG	day.NOM	although
jsme		prohráli.		
be.PST.1PL		lost.PFV.PST.1PL		

‘It was a successful day, although we lost.’

- the second conjunction meaning ("rhetorical"): *přestože* 'despite the fact that':

(14) Přestože ten nápad zní na papíře dobře, v
despite the fact sound.3SG on paper well in
praxi by byl naprosto neproveditelný.
practice would be.PST.3SG absolutely unfeasible
'Despite the fact that the idea sounds good on paper, in practice it would be absolutely unfeasible.'

- prediction: only the concessive meaning should allow sebe-low (especially if interpreted as standing-if concessive)

- crucially, the semantics for conditionals is very similar also for *even if* conditionals (3) – conjunction of two implications (modalized), just both the *even if* and the scalar particle *sebe-top* signal that there are bigger discrepancies between the antecedent (ϕ-world) and the context set – possible explanation for the different distribution of the two scalar particles: pragmatics of *even-if*;

Possible explanations

- *even-if* conditionals are known to come in two flavors (Pollock (1976), Bennett (1982), Lycan (1991)):

(15) a. Even if they were to raise my salary double, I would quit the job. introduced-if

b. Even if John said one word, I would leave. standing-if

- focusing on introduced-if: it implicates that the speaker will quit the job regardless of the circumstances – the speaker is committed to the truth of the consequent – Guerzoni and Lim (2007) calls this inference "factivity";

(16) a. Even if they were to raise my salary double, I would quit the job. introduced-if

b. Even if John said one word, I would leave. standing-if

- standing-if conditionals are not 'factive': can be true (like material implication) if the consequent is false (forcing falsity of the antecedent)

- Guerzoni and Lim (2007) argue that the "factivity" of introduced-if conditionals follows from the existential/additive presupposition of *even* (*John invited even Mary* $\rightarrow$ John invited at least one $x \neq Mary$ too)
- and that *even* in introduced-if conditionals is associated with the polarity head:

(17) Even if they were to raise my salary double, I would quit the job.

a. If they raise my salary double, I would quit. (assertion)
b. If they don't raise my salary double, I would quit. (existential presupp.)
c. $\models$ I would quit.

Application to the Czech

(18) Even if they had the highest income (sebe-top), it would be too little for them.

- following Guerzoni and Lim (2007): but here let's assume that *even* associates with the top of the scale *highest income*: existential presupposition kicks in (plus downward monotonicity):

(19) a. Even if they had the highest income, it would be too little for them.
b. If they had the lowest income, it would be too little for them.
c. $\nvDash$ It would be too little for them.

- the sebe-top (like affirmative and negative polarity head) extinguishes the whole space of alternatives
- in introduced-if conditionals, the speaker is committed to the truth of the consequent, which can be communicated via association with the strong scalar particle
- association with a weak scalar particle (sebe-low) would go against the speaker's commitment to the truth of the consequent
- that explains why *sebe-low* occurs very rarely in *even if* Czech conditionals

Last data wrinkle

- cases with sebe-top in subjunctive conditionals (around 10) look like standing-if conditionals with covert *even*:

(20) Kdyby hrála v sebevětší sračce, bude
even play.3SGFem in self-biggest crap, will.3SGFem
vynikající.
excellent
'Even if she played in the biggest crap, she would be excellent.'

- prediction: adding sebe-low should yield standing-if interpretation of even if (then sebe-low has communicative value again)
- seems to be true (other corpus examples look similar):

(21) I kdyby měl sebemenší
even if had.3SG.MASC sebe-low.ACC
výkon, je to plýtvání energií.
performance.ACC is it.NOM waste.NOM energy.GEN
'Even if it had the slightest performance, it would be a waste of energy.'

Conclusion

- the two types of Czech scalar particles *sebe*-low and *sebe*-top show different preferences in the antecedents of different types of conditionals:
  - *sebe*-low is preferred in indicative and subjunctive conditionals, while *sebe*-top is preferred in *even if* conditionals;
- this can be linked to the pragmatics of the *even if* concessive conditionals, which signal that the speaker is committed to the truth of the consequent (introduced-if conditionals);

Thank you for your attention!

Co-funded by the European Union

This work was supported by the European Regional Development Fund project " A lifetime with language: the nature and ontogeny of linguistic communication (LangInLife) " (reg. no.: CZ.02.01.01/00/23_025/0008726).

References i

Barker, Chris. 2018. "Negative Polarity as Scope Marking." *Linguistics and Philosophy* 41 (5): 483–510.

Bennett, Jonathan. 1982. "Even If." *Linguistics and Philosophy*, 403–18.

Crnič, Luka. 2011. "Getting Even." PhD thesis, Cambridge, MA: Massachusetts Institute of Technology. https://hdl.handle.net/1721.1/68912.

Crupi, Vincenzo, and Andrea Iacona. 2022. "On the Logical Form of Concessive Conditionals." *Journal of Philosophical Logic* 51 (3): 633–51.

References ii

Dočekal, Mojmír, and Martin Juřen. 2023. "Part-Whole Structure and NPIs Licensing: Experimental Evidence." https://olinco.upol.cz/wp-content/uploads/2023/05/Olinco2023_book_of_abstracts.pdf.

Gärdenfors, Peter. 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States.* The MIT press.

Guerzoni, Elena, and Dongsik Lim. 2007. "Even If, Factivity and Focus." In *Proceedings of Sinn Und Bedeutung*, 11:276–90.

Heim, Irene. 1984. "A Note on Negative Polarity and Downward Entailingness." In *Proceedings of NELS*, 14:98–107. 1983.

Kadmon, Nirit, and Fred Landman. 1993. "Any." *Linguistics and Philosophy*, 353–422.

References iii

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. "The Sketch Engine: Ten Years On." *Lexicography* 1 (1): 7–36. https://doi.org/10.1007/s40607-014-0009-9.

König, Ekkehard. 1988. "Concessive Connectives and Concessive Sentences: Cross-Linguistic Regularities and Pragmatic Principles." *Explaining Language Universals* 145166.

———. 2006. "Concessive Clauses."

Levshina, Natalia. 2020. "Conditional Inference Trees and Random Forests." In *A Practical Handbook of Corpus Linguistics*, edited by Magali Paquot and Stefan Th. Gries, 611–43. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-46216-1_25.

References iv

Lycan, William G. 1991. “Even” and” Even If.” *Linguistics and Philosophy*, 115–50.

Milosavljević, Aleksandra, and Stefan Milosavljević. 2024. “Standard Concessives Are Inherently Focused: Evidence from Serbian.” *Glossa: A Journal of General Linguistics* 9 (1).

Mizuno, Yuko. 2008. “Although Clauses in English Discourse: A Functional Analysis.” PhD thesis, .

Pollock, John L. 1976. *Subjunctive Reasoning*. Springer Netherlands. https://doi.org/10.1007/978-94-010-1500-4.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

References v

Schwab, Juliane, and Mingya Liu. 2022a. "Attenuating NPIs in Indicative and Counterfactual Conditionals." In *Proceedings of Sinn Und Bedeutung*, 26:772–89.

———. 2022b. "Processing Attenuating NPIs in Indicative and Counterfactual Conditionals." *Frontiers in Psychology* 13: 894396.

Stalnaker, Robert C. 1976. "Indicative Conditionals." In *IFS: Conditionals, Belief, Decision, Chance and Time*, 193–210. Springer.

Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin, and Achim Zeileis. 2008. "Conditional Variable Importance for Random Forests." *BMC Bioinformatics* 9 (307). https://doi.org/10.1186/1471-2105-9-307.

References vi

Von Fintel, Kai. 1999. "NPI Licensing, Strawson Entailment, and Context Dependency." *Journal of Semantics* 16 (2): 97–148.