Genericity licenses Negative Polarity Items: Experimental evidence from Czech

The aspectual architecture of the Slavic verb

Mojmír Dočekal, Masaryk University

2025-10-23

Intro

- Current theories of Negative Polarity Item (NPI, or more generally PI) licensing (Gajewski (2011), Chierchia (2013)) consider them unlicensed in affirmative episodic statements
  - if no traditionally recognized licenser is present;
- Experiments (for English: Gajewski (2016) – definite DPs; for Czech: Dočekal and Juřen (2023)) show that the number of the head NP may affect NPI licensing:

(1) a. The students who have **any** books on NPIs are selling them.
b. *The student who has **any** books on NPIs is selling them. (Guerzoni & Sharvit, 2007, p. 372, (29))

- Another theoretically expected factor in acceptability is genericity
  - Kadmon and Landman (1993) point out the relation between genericity and free-choice *any*;
- However, the effect on potential indefinite licensors for PIs has not been tested;

- We constructed an experiment testing the number of the head NP and the genericity/episodicity of the verb as factors in NPI licensing in Czech;
- We asked whether speakers recognize genericity as a boost for the acceptability of sentences with *sebemenší* (NPI);

- Generic sentences are non-monotonic (Nishiguchi (2003), Kirkpatrick (2019), among others), which poses a problem for the most widely used theories of NPIs, which derive their distribution from downward entailingness (Ladusaw (1979) and subsequent work);

(2) a. Birds lay eggs.
b. ↛ Male birds lay eggs.

(3) Our research question:

a. Is genericity a factor in NPI licensing in Czech?

- Especially considering that the previous experiment on English definite descriptions did not show this effect
- But considering that Czech is one of the languages with dedicated morphological marking of genericity

- Joint work with Rhiana Horovská
- slides:

- We used the scope theory of NPI licensing by Barker (2018), which can account for NPI licensing in non-monotonic contexts;
- NPIs are licensed if their wide scope does not entail their narrow scope:
  - licensing by negation
  - The NPI signals the narrow scope of the indefinite with respect to negation

(4) John didn't see *any* students.

a. $\exists \mathrm{x}[\mathrm{Student}'(x) \wedge \neg \mathrm{See}'(\mathrm{J}, x)]$

b. $\nvDash \neg \exists \mathrm{x}[\mathrm{Student}'(x) \wedge \mathrm{See}'(\mathrm{J}, x)]$

- In non-monotonic contexts, the entailment does not hold, so NPIs can be licensed as well

Experiment

- Series of experiments: reporting the third one here
- We conducted an experiment in PCIbex with native speakers of Czech
- acceptability judgement task: 1 = unacceptable ... 5 = acceptable
- the fillers of corresponding complexity;
- 112 participants; 98 passed the fillers

- 2×2×2 factorial design controlling for:

1. number (sg/pl on the subject and predicate)
2. Genericity (generic: the adverb *většinou* – "mostly" + habitual *-av-* morpheme in the predicate × episodic: the adverb *právě* – "just" + perfective predicate)
3. PI presence (PI: *sebemenší* "the slightest"; adj: a non-polarity-sensitive adjective)

- The participants rated the acceptability of the stimuli on a scale from 1 (low) to 5 (high)
- We hypothesized:

a) genericity would be rated higher than episodicity

b) plural higher than singular

c) PI-sentences less accepted than non-PI sentences.

Example item (generic conditions): - preceded by a short context:

(5) Kontext: Jiřka se zajímá o
context Jiřka REFL take-interest about
jasnovidky, což jsou obvykle mystické
clairvoyant.F.PL.ACC which are usually mystical
ženy. Tvrdí o nich:
women claim.3SG about them:
Context: Jiřka takes interest in clairvoyants who are usually mystical women. She claims about them:

(6) **(pl)**:Jasnovidky/**(sg)**:Jasnovidka se [**(PI)**:*sebemenší*
clairvoyant.F.PL/SG with slightest.INS
známkou]/[**(adj)**:*vysokou* úrovní] nadání většinou
hint.INS/high.INS level.INS talent.GEN mostly
vídávají/vídává smutnou budoucnost.
see.HAB.3PL/3SG sad.ACC future.ACC
'Clairvoyants/A clairvoyant with the slightest hint/a high level of talent see/-s a sad future most of the times.'

(episodic conditions) - context first:

(7) Kontext: Jiřka vyzpovídala jasnovidky, z nichž
context Jiřka interviewed clairvoyants from whom
některé vyvěštily samé trápení. Tvrdí o
some prophesied all anguish claim.3SG about
nich:
them:
'Context: Jiřka has interviewed clairvoyants out of whom some prophesied nothing but anguish. She claims about them:'

(8) **(pl)**:Jasnovidky **(sg)**:Jasnovidka se [**(PI)**:*sebemenší*
clairvoyant.F.PL /SG with slightest.INS
známkou]/[**(adj)**:vysokou úrovní] nadání právě
hint.INS/high.INS level.INS talent.GEN just
uviděly/uviděla smutnou budoucnost.
see.PF.PST.3PL/3SG sad.ACC future.ACC
'Clairvoyants/A clairvoyant with the slightest hint/a high level of talent have/-s just seen a sad future.'

Descriptive Statistics

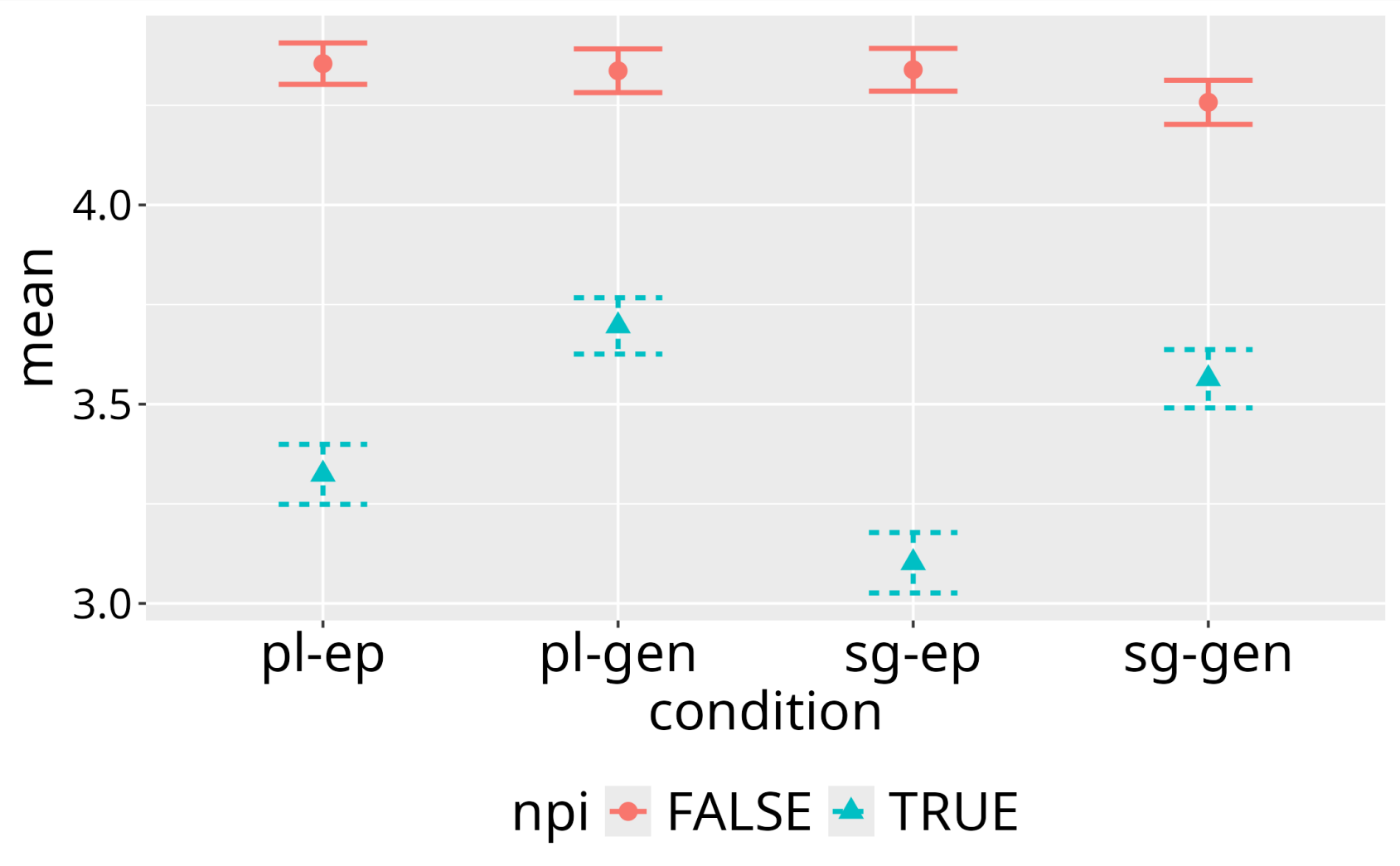

Figure 1: Responses: mean and standard errors

Inferential statistics

- The data were analyzed in a mixed-effects Bayesian linear regression model using the *rstanarm* package (Brilleman et al. (2018)) in *R* (R Core Team (2024));
- The model uses sum-coded contrasts in a 2×2×2 design with interactions (the conditions GEN, PLUR, NPI against EP, SG, ADJ);
- The dependent variable is the participant's response on the Likert scale;
- The model uses the full random effects structure and *rstanarm*'s default weakly informative priors.

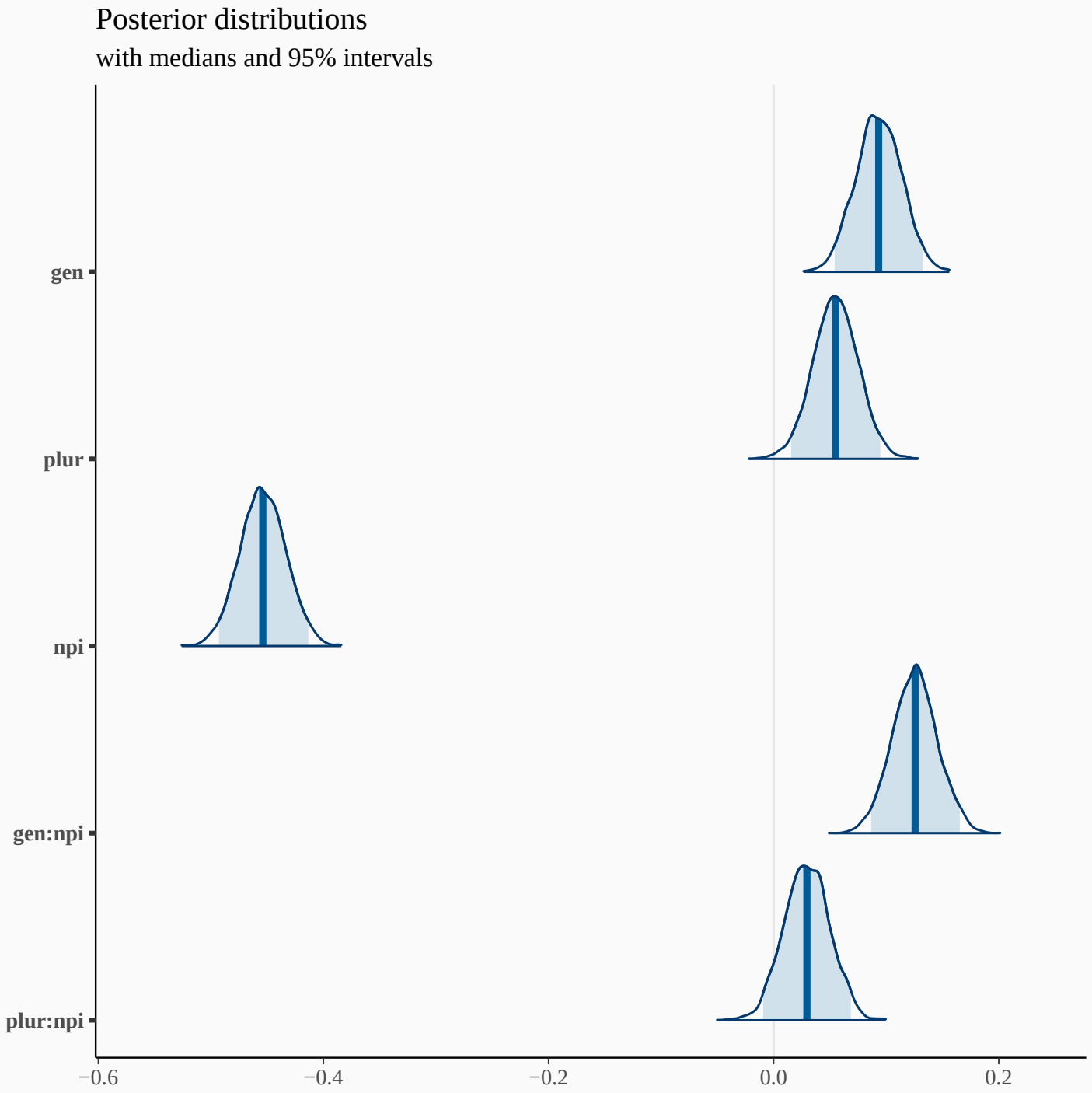

Figure 2: Main effects and interaction

- The model shows:

0. the effects (and interactions) are interpreted against the grand mean (Intercept): $\widehat{\beta} = 3.87$, CrI = [3.73, 4.02], BF 2.13e+93;
1. A credible main positive effect of genericity (GEN: $\widehat{\beta} = 0.09$, 95% Credibility Interval (CrI) = [0.05, 0.13], with Bayes Factor (BF) 31.8, indicating very strong evidence in favor of the effect)

2. The strongest main effect is negative for the PI (NPI: $\widehat{\beta} = -0.45$, CrI = [-0.49, -0.41], BF = 2.36e+23, indicating extreme evidence)
3. The third main effect is positive for the plural (PLUR: $\widehat{\beta} = 0.06$, CrI = [0.02, 0.10]), but there is moderate evidence against its existence due to a BF of 0.330;

5. The strongest interaction effect is between genericity and PI (GEN:NPI: $\widehat{\beta} = 0.13$, CrI = [0.09, 0.16], BF = 3.00e+04, indicating extreme evidence)

All other interactions (GEN:PLUR, PLUR:NPI, GEN:PLUR:NPI) receive BF 0.018, indicating very strong evidence against their existence (direct support for their H_0).

Linguistic Interpretation:

1. PIs are licensed by genericity (credible interaction GEN:NPI)
2. But not by plurality (non-credible PLUR:NPI).

- In the previous experiment (without context), the interaction effect was much less credible
  - Genericity interpretation: morphological marking plus context information
    - Even in Czech, the genericity marker can be used to derive non-generic readings (secondary imperfectives, etc.) – morphology is not a foolproof marker of genericity

Preliminary steps to formalization

- To formalize genericity, we follow the dynamic approaches to counterfactuals and generics (von Fintel 2001; Kirkpatrick (2019), among others)
- Which use the dynamic conditional '$>$' quantifying over most normal worlds

- The generic (1) is then formalized in (9) as non-monotonic modal universal quantification (UQ).

(9) a. $\forall x(\text{Clairvoyant}'(x) > \text{SeeSadFuture}'(x))$

$\forall x \in^* (w, \text{ClairvoyantWithTalent}'(x) \wedge \text{HasTalent}'(x) \text{ in } w) \subseteq$
$\{w \in W : \text{SeeSadFuture}'(x) \text{ in } w\}$

- Where $^*(w, [\![\phi]\!])$ is a function from the actual world (w) to the set of most normal worlds with respect to ϕ being true in w
  - The normality makes the quantification non-monotonic
  - After Stalnaker and Lewis; newer formalization: Asher and Morreau (1995), among others

- Assuming that *sebemenší* is an NPI, its occurrence in (1)/(2) is unexpected in standard DE theories of NPI licensing;
- Since it appears in the first argument of the non-monotonic UQ, we use a non-standard approach to NPI licensing (Barker (2018))
- Licensed if the wide scope interpretation of the NPI fails to entail the narrow scope of the NPI with respect to other operators – true for (9) where the PI occurs in the first argument ($ClairvoyantWithTalent(x)$).

- Preliminary truth conditions:

(10) $\forall x \in^{*} (w, Clairvoyant'(x) \wedge max\{d : HasTalent'(x, d)\} > min(S_{talent})$ in $w) \subseteq \{w \in W : SeeSadFuture'(x)$ in $w\}$

- The NPI contributes the $min(S_{talent})$ – for future work: decompose *sebemenší*
- The non-monotonic modal universal quantifier has wide scope, and the wide scope of the NPI would not entail the narrow scope

- In the episodic sentence (depending on formalization):

1. Either there is no operator over which the NPI can take wide scope – in which case inserting an expression signaling narrow scope (NPI) is uninformative;
2. There is existential closure (of the event variable) – the wide and narrow scope of the NPI are equivalent; thus, the NPI is not licensed.

Side remarks on genericity

- Future goal: integrate our results into a framework close to Kratzer (1995)
- Since it seems to fit the data well, here is a sketch of the idea:

1. individual-level predicates: *inteligentní* 'intelligent', *vysoký* 'tall', *pocházet z Německa* 'come from Germany', *umět německy* 'know German', ...
2. stage-level predicates: *nahatý* 'naked', *vzteklý* 'angry', *opilý* 'drunk', *mluvit* 'speak'

- Stage-level predicates have a "Davidsonian" argument for the spatiotemporal location of the event
- Individual-level predicates lack this argument and (ceteris paribus) cannot be used with the generic marker *-av-*:

(11) a. Marie *umívá německy/*pocházívá z
Marie knows.GEN German/comes.GEN from
Německa/*bývá vysoká.
Germany/is.GEN tall.F.SG
Intended: 'Marie knows German often/comes from Germany often/is often tall.'

b. Marie mluvívá německy/bývá vzteklá/bývá
Marie speaks.GEN German/is.GEN angry/is.GEN
opilá.
drunk.F.SG
'Marie speaks German often/is often angry/is often drunk.'

- But with bare plurals, even the generic version is acceptable with ILP:

(12) Češi bývají vysocí/bývají z
Czechs are.GEN tall.PL/are.GEN from
Prahy/bývají inteligentní.
Prague/are.GEN intelligent.PL
'Czechs are often tall/often come from Prague/are often intelligent.'

- This seems to show that Czech generic verbs behave like Q-adverbs (*usually*, *generally*) – see also Filip (2023);
- In Kratzer's analysis, the Q-adverb binds the free variables (we need not decide whether it is truly an unselective binder or not, but it seems to be less restricted than at least regular quantifiers like *every* or *most*):

(13) Marie mluvívá německy dobře.

(14) $\mathrm{GEN}_s[\mathrm{Speak}'(\mathrm{M}, \mathrm{German}, s)][\mathrm{SpeakWell}'(\mathrm{M}, \mathrm{German}, s)]$

- the relation between the antecedent and the consequent is the $>$ relation (modal universal quantification over the most normal worlds)

(15) Marie *umívá německy dobře.

(16) $\mathrm{GEN}[\mathrm{Know}'(\mathrm{M}, \mathrm{German})][\mathrm{KnowWell}'(\mathrm{M}, \mathrm{German})]$

- Kratzer's prohibition against vacuous quantification:
  - for Kratzer, the prohibition can be overcome by quantification over objects: *Always when Mary knows a foreign language, she knows it well.*
    - maybe possible but much harder in Czech

(17) Prohibition against vacuous quantification: For every quantifier Q, there must be a variable x such that Q binds an occurrence of x in both its restrictive clause and its nuclear scope. (Kratzer 1995:131)

- For the bare plural version – the generic operator can bind the variable introduced by the bare plural noun phrase:

(18) Češi bývají vysocí.

(19) $\mathrm{GEN}_x[\mathrm{Czech}(x)][\mathrm{Tall}(x)]$

- Prediction – generic marking should be more probable with stage-level predicates than with individual-level predicates (there is always a variable to be bound in SLP but not necessarily in ILP)

- ČNK search ([lemma=“*.ávat” & tag=”V……….I.*”]): *sedávat* ‘sit’, *hrávat* ‘play’, *dělávat* ‘do’, *vysedávat* ‘sit out’, *stávat* ‘stand’, *říkávat* ‘say’, *zprostředkovávat* ‘mediate’ (from the most frequent verbs; the rest are secondary imperfectives, the more productive usage of *-va-*, it seems, like *dávat* ‘give’, *zůstávat* ‘stay’, *dostávat* ‘receive’)
    - But 7 are SLP generic verbs; no ILP generics

- more empirical arguments come from:

1. ECM constructions: only SLP generic verbs can be embedded in ECM constructions;
2. depictives: only for adjective predicates, so no generic marking is possible;
3. lifetime effects: maybe related to the distant past reading of generic sentences with ILP predicates, but not with SLP predicates: *Tady stával dům' 'There used to be a house here' (ILP generic) vs.* Tady Petr hrával fotbal* 'Peter used to play football here' (SLP generic);

- Prediction of Kratzer's analysis (but also of other approaches to generics): NPIs should be acceptable only in the restrictor of generic sentences:

(20) Ochotníci se sebemenším talentem hrávají
amateurs with slightest.INS talent.INS play.GEN
cimrmanovské hry.
Cimrman.PL.ACC plays.ACC
'Amateur actors with the slightest talent play Cimrman plays.'

- Since the NPI is in the restrictor $\text{GEN}_{s,x}$[amateur$_x$ ...NPI...][...Cimrman play$_s$...] part

- vs. NPI in the consequent:

(21) *Cimrmanovské hry jsou hrávány
Cimrman.PL.ACC plays.ACC are played.GEN
ochotníky se sebemenším talentem.
amateurs.INS with slightest.INS talent.INS
'Cimrman plays are played by amateur actors with the slightest talent.'

- Bad because the NPI is in the consequent part
  GEN$_{s,x}$[...Cimrman play$_s$...][amateur$_x$ ...NPI...]
- zooming out: the SLP/ILP distinction can be manipulated in syntax (Kratzer (1995)) and in information structure (Jäger (2001))

- The general picture: in this Kratzer-style analysis, genericity should license NPIs if they occur in the restrictor part of the generic statement
- The generic operator ranges either over individual variables (bare plural ILP predicates) or situation variables (SLP predicates)
    - In the experiment, we used more SLP predicates – they have higher compatibility with generic marking
    - Thus, the GEN operator bound both the situation variable and the free variable introduced by the bare plural noun phrase

Back to the experiment

- Our results starkly contrast with the English data (Gajewski (2016)), where the experiment did not show the effect of genericity on NPI licensing in definite descriptions;
- But only the effect of the number of the head NP;
- Similar to the contrast:

(22) a. The students who have **any** books on NPIs are selling them.
b. *The student who has **any** books on NPIs is selling them.

- Why the difference?

- Both singular and plural definite descriptions are Strawson Downward Entailing (SDE) in English (Gajewski and Hsieh (2014)):

(23) a. The student has arived.
b. There is a unique salient German student.
c. $\models$ The German student has arrived.

(24) a. The students have arrived.
b. There are salient German students.
c. $\models$ The German students have arrived.

- Prediction: both sg and pl should license NPIs equally well (but they do not)

- It seems like the correct generalization concerns genericity rather than number:
- Also Gajewski suggests (based on older literature) that plural sentences with NPIs express generalizations that suspend existential presuppositions
- Existential presuppositions of definite descriptions are claimed to be the reason why, e.g., *both* does not license NPIs:

(25) a. Every student/the students who saw anything should report it to the police.
b. *Both students who saw anything should report it to the police.

- Also pointing in the same direction:

(26) a. The student who has ever grasped this theorem knows how hard it is.
b. The students who have ever grasped this theorem know how hard it is.

- From Hoeksema (2008), cited by Gajewski (2014)
- Which would again point to genericity as a factor in NPI licensing
- Which was found in our experiment but not in Gajewski's

- We hypothesize that this follows partially from the predicate's morphological marking by *-av-* (dedicated genericity marking in Slavic languages)
  - Which allows us to test the effects of genericity more directly;
  - And allows us to avoid the confound of number marking;
  - But also from the fact that we provided a context, which made the generic reading more salient (unlike Gajewski);

Thank you for your attention!

Appendix

- A note on FCI/NPI licensing by genericity
- In the system of Barker (2018), FCIs are licensed similarly to NPIs: they signal narrow scope
  - But unlike NPIs, FCIs require modal or generic meaning
- Unified system, so we need not precisely tease apart NPIs and FCIs and decide whether *sebemenší* is an NPI or FCI
- But it seems more like an NPI:

1. It does not have the usual FCI inferences:

(27) You can take any fruit.

a. $\diamond(f_1 \vee f_2 \vee \ldots f_n) \rightsquigarrow \diamond f_1 \wedge \diamond f_2 \wedge \cdots \diamond f_n$

(28) Můžeš si vzít sebehorší ovoce.
can.2SG REFL take worst Fruit.ACC
'You can take the worst fruit.'

a. $\diamond(f_1 \vee f_2 \vee \ldots f_n) \not\rightsquigarrow \diamond f_1 \wedge \diamond f_2 \wedge \cdots \diamond f_n$

- Very odd, if grammatical at all

2. It has an existential (not universal) interpretation:

(29) Pokud věříš jakékoliv teorii, tak
if believe.2SG any theory.ACC then
piš beletrii.
write.IMP fiction.ACC
'If you believe any theory, then write fiction.'

a. $(\ldots \forall \ldots) \rightarrow ()$

(30) Pokud věříš sebehorší teorii, tak
if believe.2SG worst theory.ACC then
piš beletrii.
write.IMP fiction.ACC
'If you believe the worst theory, then write fiction.'

a. $(\ldots \iota/\exists \ldots) \rightarrow ()$

- Existential or possibly unique interpretation (the worst theory you can think of)

References i

Asher, Nicholas, and Michael Morreau. 1995. "What Some Generic Sentences Mean." *The Generic Book*, 300–338.

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

Brilleman, SL, MJ Crowther, M Moreno-Betancur, J Buros Novik, and R Wolfe. 2018. "Joint Longitudinal and Time-to-Event Models via Stan." https://github.com/stan-dev/stancon_talks/.

Chierchia, Gennaro. 2013. *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford University Press.

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.

Gajewski, Jon. 2011. "Licensing Strong Npis." *Natural Language Semantics* 19: 109–48.

———. 2016. "Another Look at NPIs in Definite Descriptions: An Experimental Approach." *Negation and Polarity: Experimental Perspectives*, 307–27.

Gajewski, Jon, and I-ta Chris Hsieh. 2014. "Comments on Negative Polarity Items in Definite Description." *The Art and Craft of Semantics: A Festschrift for Irene Heim*. MITWPL Cambridge.

References iii

Jäger, Gerhard. 2001. "Topic-comment Structure and the Contrast Between Stage Level and Individual Level Predicates." *Journal of Semantics* 18 (2): 83–126. https://doi.org/10.1093/jos/18.2.83.

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

Kirkpatrick, James Ravi. 2019. "Essays on Genericity." PhD thesis, University of Oxford.

Kratzer, Angelika. 1995. "Stage-Level and Individual-Level Predicates." In *The Generic Book*, edited by Greg N. Carlson and Francis Jeffry Pelletier, 125–75. University of Chicago Press.

Ladusaw, William Allen. 1979. *Polarity Sensitivity as Inherent Scope Relations.* The University of Texas at Austin.

References iv

Nishiguchi, Sumiyo. 2003. "Non-Monotonic Negativity." In *Language, Information and Computation: Proceedings of the 17th Pacific Asia Conference, 1-3 October, 2003, Sentosa, Singapore*, 204–15. Waseda University.

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