3.1 GENERAL INFORMATION ABOUT PROJECT B05

3.1.1 Project title: Negative Polarity Items in non-negative contexts

3.1.2 Research areas

104-01 Comparative Linguistics, Semantics

3.1.3 Project leaders

Richter, Frank, Apl.Prof. Dr., 31.10.1967, German
Institut für England- und Amerikastudien, Goethe-Universität,
Norbert-Wollheim-Platz 1, 60629 Frankfurt am Main
+49-69-79832525; f.richter@em.uni-frankfurt.de

Sailer, Manfred, Prof. Dr., 27.05.1969, German
Institut für England- und Amerikastudien, Goethe-Universität,
Norbert-Wollheim-Platz 1, 60629 Frankfurt am Main
+49-69-79832526; sailer@em.uni-frankfurt.de

Is the employment of the project leaders at the institutions indicated contractually secured for the duration of the proposed funding period? Yes

Do any of the above mentioned persons hold fixed-term positions? No

Funding for the positions of the project leaders at the institutions indicated is covered by core support (state funds or similar): Yes

3.1.4 Legal issues

This project includes

1.	research on human subjects or human material.	Yes
	A copy of the required approval of the responsible ethics committee is included with the proposal.	No
2.	clinical trials.	No
3.	experiments involving vertebrates.	No
4.	experiments involving recombinant DNA.	No
5.	research involving human embryonic stem cells.	No
6.	research concerning the Convention on Biological Diversity.	No
7.	investigations involving dual use research of concern.	No

3.2 SUMMARY

The project aims at reconciling the common assumption of concentric Negative Polarity Item (NPI) licensing contexts (antimorphic ⊂ anti-additive ⊂ downward-entailing ⊂ ...) with apparently unlicensed exceptions and with the much more fine-grained distributional profiles observed in corpus-based research. Whereas classical theories mainly concern licensing by at-issue content, we investigate NPI licensing by non-at-issue content. This leads to a better understanding of the relation between explicit (morpho-syntactically marked) negation and implicit negation (by conventionalized association). The hypothesis is that strong NPIs require a strong (anti-additive) licenser, but the licenser can be at-issue or non-at-issue. Weak NPIs accept a weak (e.g., downward-entailing) licenser, though their licensing must follow from the computation of the at-issue content.

The project approaches its main hypothesis with data from primarily three languages and with different complementary research methods. Exploiting the insights in the distribution and character of a large set of German and Romanian NPIs docu-

mented in previous work, we plan to investigate quantitative and qualitative properties of NPIs in English, German and Romanian. With semi-automatic extraction methods, the initial data set of NPIs and their theoretical classification and quantitative documentation will be recast into a considerably extended searchable database of NPIs, their licensing environments and their corpus occurrence patterns in the three project languages. To complement the limitations of corpus data, systematically collected acceptability judgments by native speakers will be used to critically evaluate apparent NPI licensing patterns and to investigate the availability of relevant inferences. Where relevant, we will add data from French to our set of languages. Further corpus studies will target quantitative occurrence distributions of representative items in systematically selected licensing contexts. The corpus studies also aim to shed light on differences in distributional profiles between various registers and corpus types.

Based on solid empirical foundations the grammar-theoretic goal of the project is the formulation of a theory of NPI licensing with enriched semantic representations, i.e. encompassing at-issue and non-at-issue content, in a constraint-based grammar framework. This NPI licensing theory will be built on the one hand on occurrence patterns of individual NPIs with particular attention to occurrence patterns in overtly non-negative contexts and the role individual NPIs play in these patterns (especially in the case of corresponding NPIs in different languages). On the other hand, the analysis of licensing contexts and different types of negativity exhibited by these contexts will be central. Of particular interest are constructions such as *before* clauses, which are also known as environments of expletive negation. Building on previous literature and pilot studies of the PIs, it will be beneficial to include French as an additional project language here.

The interaction of the occurrence of expletive negation and of NPIs in such environments should be of great interest for insights into the nature of and relationships between negation forms, negative meaning, and the inferences they trigger alone and in combination.

By their very definition, NPIs are conceived of as single- or multi-word items of any syntactic category whose use depends on the presence of negation: they are oriented toward negative polarity. The study of their licensing in the presence of a diverse variety of morpho-syntactic elements associated with negation, their non-licensing in the presence of other morpho-syntactic elements also associated with negation, but especially their unexpected licensing in the absence of overt morpho-syntactic elements associated with negation, and the comparison of all three types of (non-) licensing across the three project languages is expected to elucidate the status of negation and its interaction with grammatical categories in human languages in general. **All of these investigations address the fundamental research question of Domain B, which is the interaction of negation and other grammatical categories.** These are represented in B05 by the diverse morpho-syntactic elements associated with elements of different kinds of negativity in NPI licensing.

Based on the results of the first funding period, the long-term goal for future funding periods is the development of a distributional theory of the notion of possible NPIs, i.e. a predictive theory of which kinds of lexical semantic and pragmatic properties facilitate the restriction to certain kinds of negative environments of single- and multi-word expressions. We plan to conduct experimental tests of this theory, based on distributional hypotheses about classes of NPIs in contexts without overt licensers, which presupposes a detailed licensing theory of a sufficiently large number of such items. In addition, we want to investigate the question of which aspects of NPI-licensing are represented by large language models, especially with respect to licensing with non-overt negation. The sensitivity of such models to certain distributional restrictions could also become interesting in the context of developing experimental items.

3.3   RESEARCH RATIONALE

3.3.1   Current state of research and preliminary work

Empirical challenges: Negative Polarity Items (NPIs) are expressions that typically occur in negated sentences and are banned from affirmative contexts. Even accounts that acknowledge diversity among NPIs and distinguish a higher number of distribution classes usually assume a subset-relation between, or a concentric ordering of, NPI licensing contexts—often defined in terms of entailment properties derived from logical negation (Zwarts, 1998; Giannakidou, 1998; Gajewski, 2007). NPIs are classified according to the strength of the licensing contexts in which they occur. As a corollary of this assumption, weak NPIs (*ever*) should occur in a superset of the contexts of strong NPIs (*lift a finger*). While a plausible hypothesis, it is critically challenged by empirical observations as follows:

Challenge 1 - Non-uniformity of licensers of the same class: Some NPIs can occur with some relatively weak licensers, though not with others of the same strength. For example, both German *kaum* 'hardly/barely' and *selten* 'rarely' are downward-entailing but not anti-additive. Nonetheless, such NPIs are found with the former, see (1), though not with

the latter. In the literature, NPIs with this behavior are typically classified as strong NPIs because they fail to be generally acceptable in weak licensing contexts.

(1) Was   den   Steirern   von   der Entpolitisierung   der   Schulen
     What   the   people.of.Steiermark of   the depoliticizaton   of.the   schools

     schon die   Ohren voll   gelogen wurden,   geht   kaum   auf   eine   Kuhhaut
     already the   ears full   lied   was   goes   barely   on   a   cowskin
     'It is barely bearable how much the people of Steiermark have already been lied to ...'
     (Neue Kronen-Zeitung, 10.03.1999)

Using fine-grained distributional profiles, Hoeksema (2013) makes analogous observations about more subtle contextual behavior of NPIs when comparing various Strawson downward-entailing contexts.

Challenge 2 - Contexts licensing strong but not weak NPIs: Sedivy (1990) notes that outside the expected negative licensing contexts, strong NPIs occur in sentences with auxiliaries with contrastive stress and with counterfactive modals, though weak NPIs cannot, see (2) and (3). Evidently the idea that strong NPIs occur in a subset of the contexts in which weak NPIs may occur cannot be maintained without qualifications. Even more critically, there is no obvious NPI-licenser in (2-a) and (3-a).

(2) a. I DO give a damn.
     b. *Bert DID ever kiss Marilyn Monroe.

(3) a. John should have lifted a finger to help Mary clean up.
     b. *John should have eaten any healthful tofu. (Sedivy, 1990: 98–99)

The reported NPI data suggest that there is some property, like downward-entailingness, that negation shares with some other expressions, which underlies the licensing of weak NPIs. However, there seems to be another, lesser studied, property that negation shares with the expressions and contexts in which we find strong NPIs. This could be some non-at-issue effect. **In line with the Neg-Plus Hypothesis, this points to the direction that negation has at least two properties that can occur independently in other, negation-like operators.** Alternatively, it could be the NPIs that impose certain constraints on the contexts in which they can occur, such as being downward entailing (for weak NPIs) and involving some sort of cancellation (for strong NPIs). Then, in **accordance with the Neg-Only Hypothesis, the similarity of negation and different types of NPI-licensing operators and contexts is not a constitutive property of negation but rather a** side effect from constraints on readings imposed by the context and/or an NPI.

Systematic, theory-driven data work: Richter et al. (2010b) presents the *Collection of Distributionally Idiosyncratic Items* (CoDII, www.english-linguistics.de/codii), comprising the largest available collections of NPIs with corpus examples and a linguistic classification strictly following positive corpus evidence at the time of data collection (165 German and 58 Romanian NPIs). CoDII lists 14 licensing contexts and provides syntactic information for the German data.

The two challenges apparent in the CoDII data are: (a), NPIs classified as strong are found with the simple downward-entailing licenser *kaum* 'barely' (see (1)), though not with other simple downward-entailing licensers; (b), CoDII provides a category "exceptions" which subsumes well-formed, attested cases of NPIs without an appropriate licenser such as (4), *bis drei zählen können*. This category of exceptions, while being a valuable resource for further study, is not analyzed in CoDII itself. So far, this rich source of data has not been systematically explored and described in grammar theory.

(4) Wer Skat spielen   kann,   kann   auf   jeden   Fall   bis   drei   zählen, kann   beim
     Who skat play   can   can   in   any   case   up.to   three   count   can   with
     Reizen   bis ins   Unendliche   gehen,   ist   in   der   Lage   aus
     tease   up.to   infinity   go   is   in   the   condition   from
     einem   schlechten Blatt   noch   etwas   Gutes   zu   machen
     a   bad   hand   still   something   good   to   make
     'Anyone who can play skat definitely knows what's what, ...'
     (Berliner Zeitung, 26.10.1989)

Based on semi-automatic NPI extraction from corpora, **Richter** et al. (2010a) report more observations that illustrate various challenges from usage data. The expression *den Wald vor lauter Bäumen sehen* 'see the obvious' has the distribution of a strong NPI and is traditionally classified as such. While over 90% of their corpus data are consistent with this

classification, **Richter** et al. (2010a: 102–103) also list examples without overt strong NPI licensers, including (5) with a concessive particle *dennoch* 'still' as the only plausible licenser. The modal contexts in (6) and (7) are known environments for weak NPIs, and questions as in (8) are sometimes subsumed—together with the previous contexts—under a category of nonveridical licensing environments. In the traditional picture of concentric NPI licensing, these should license even fewer, 'weaker' NPIs, but not strong NPIs such as *den Wald vor (lauter) Bäumen sehen*.

(5) Manchmal sieht M.L. vor lauter Bäumen dennoch den Wald.
    sometimes sees M.L. of all.the trees still the forrest
    'Occasionally M. L. nonetheless manages to see the obvious.'

(6) Doch wie immer sollte man zunächst einmal den Wald vor
    but how ever should one at.first the forest for
    lauter Bäumen sehen.
    all.the trees see
    'As always one should first note the obvious.'

(7) Hätte die Kommission eindeutige und anerkannte Prioritäten und
    had the commission clear and recognized priorities and
    könnte sie den Wald vor Bäumen sehen, hätten wir nicht
    could it the forest for trees see had we not
    diese Aussprache heute Nachmittag.
    this debate today afternoon
    'If the commission had clear and recognized priorities and if it could see the obvious, we would not have to meet this afternoon.'

(8) Sieht er dann den Wald vor lauter Bäumen?
    sees he then the forrest for all.the trees
    'Does he see the obvious?'

In extensive corpus studies, Hoeksema (2010, 2012) investigates fine-grained distributional profiles of NPIs in sets of licensing contexts. The profiles are obtained by counting the number of occurrences in lexically or syntactically defined NPI licensing environments. The distributional profiles of individual items are subject to diachronic change that can happen within a few decades. The classes of observed distributional profiles give us valuable insight into usage preferences. These actual classes inhabit only a small subspace of the theoretical combinatorial possibilities (12 observed classes out of 256 possible classes given the number of investigated environments in Hoeksema (2012)) and must thus be considered non-arbitrary.

Richter & Radó (2014) conduct questionnaire studies with acceptability ratings on a 7 point Likert scale in which they vary the licensing environment of strong and weak NPIs by providing strong or weak licensers and by studying licensing in simple clauses versus Neg Raising constructions with and without intervening quantifiers. These variations reveal the gradient property of acceptability of NPIs in systematically different environments, with the number of environmental features differing from ideal licensing conditions contributing to a slow acceptability degradation. Schaebbicke et al. (2021) conduct an acceptability study with 60 NPIs gathered from CoDII and observe seven clusters of NPI licensing. With their observations, they agree with the results of Richter & Radó (2014) that NPI licensing should be considered a gradient property rather than categorical. Two of their licensing patterns follow the licensing categories predicted by the traditional classification of NPI licensing, with various other effects being observed with the other five. These kinds of data and insights are so far only implicit in CoDII, but their explicit addition will be explored in the present project, as it contributes to our systematic documentation of the interaction of negation-like contexts and negation-sensitive items. It would add very valuable information for other researchers on NPIs as well.

Non-at-issue meaning: Puzzling corpus data like (5) seem to be uniformly tied to pragmatic licensing effects. Over the last decades, the degree of formalization of implicit meaning has increased substantively. Van der Sandt (1992) demonstrates how presupposition accommodation can be treated alongside anaphora resolution within Discourse Representation Theory (Kamp & Reyle, 1993; Kamp et al., 2011). Potts (2005), Schlenker (2012, 2013), Liu (2012), and Gutzmann (2013) develop an integration of conventional implicatures into the computation of meaning. Gutzmann et al. (2020) analyzes verum focus as a use-conditional operator that marks that a speaker wants to prevent the negation of the uttered proposition being integrated into the common ground. **Sailer** (2021, 2022) sketches how the integration of non-at-issue content into the semantic representation can be used in a theory of NPI-licensing: For example, the NPIs *give a damn* can be licensed by a negation not only in the at-issue part of the semantic representation of an utterance, but also in its

non-at-issue part. In (2-a), whereas the proposition *I don't give a damn* should be prevented from being integrated into the common ground, this proposition is exactly the implicit negative meaning postulated in Sedivy (1990). Consequently, the NPI is in the scope of a negation in the non-at-issue part of the semantic representation of the utterance, which is sufficient for it to be licensed. This provides an initial theoretical setting to capture the so far underdescribed contexts of NPI-licensing that we had mentioned above.

3.3.2 Project- and subject-related list of publications

Bard, Ellen Gurman, Dan Robertson & Antonella Sorace. 1996. Magnitude Estimation of linguistic acceptability. In *Language* 72 (1), 32-68. doi: 10.2307/416793.

von Bergen, Anke & Klaus von Bergen. 1993. *Negative Polarität im Englischen*. Tübingen: Narr.

Dimitriadis, Alexis. 2010. Matching needs and resources: How NLP can help theoretical linguistics. In Fei Xia, William Lewis & Lori Levin (eds.), *Proceedings of the 2010 workshop on NLP and linguistics: Finding the common ground, 22–24. Uppsala, Sweden*. Association for Computational Linguistics. URL http://www.aclweb.org/anthology/W10-2104.

Dimitriadis, Alexis & Simon Musgrave. 2009. Designing linguistic databases: A primer for linguists. In Martin Everaert, Simon Musgrave & Alexis Dimitriadis (eds.), *The use of databases in cross-linguistic studies* (Empirical Approaches to Language Typology 41), 13– 75. Berlin: Mouton de Gruyter.

Featherston, Sam. 2007. Data in generative grammar: The stick and the carrot. In *Theoretical Linguistics* 33 (3). doi: 10.1515/TL.2007.020.

Gajewski, Jon Robert. 2007. Neg raising and polarity. *Linguistics and Philosophy* 30. 289–328.

Giannakidou, Anastasia. 1998. *Polarity sensitivity as nonveridical dependency*. Amsterdam: John Benjamins.

Gutzmann, Daniel. 2013. Expressives and beyond. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond expressives* (Current Research in the Semantics Pragmatics-Interface (CRiSPI) 28), 1–58. Brill. doi: 10.1163/9789004183988_002. URL danielgutzmann.com/ work/expressives-and-beyond.

Gutzmann, Daniel, Katharina Hartmann & Lisa Matthewson. 2020. Verum focus is verum, not focus: Cross-linguistic evidence. *Glossa* 5(1). 51. 1–48. doi: https://doi.org/10.5334/gjgl.347.

Hoeksema, Jack. 2010. Dutch ENIG: From nonveridicality to downward entailment. *Natural Language and Linguistic Theory* 28(4). 837–859.

Hoeksema, Jack. 1998. Corpus study of negative polarity items. Html version of a paper which appeared in the *IV-V Jornades de corpus linguistics 1996-1997*, Universitat Pompeu Fabre, Barcelona. URL http://www.let.rug.nl/~hoeksema/barcelona.html.

Hoeksema, Jack. 2012. On the natural history of negative polarity items. *Linguistic Analysis* 38(1/2). 3–33.

Hoeksema, Jack. 2013. Polarity items in Strawsonian contexts: A comparison. In Eva Csipak, Regine Eckardt, Mingya Liu & Manfred Sailer (eds.), *Beyond "any" and "ever": New explorations in negative polarity sensitivity*, 47–77. Berlin: De Gruyter.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Iatridou, Sabine & Hedde Zeijlstra. 2021. The complex beauty of boundary adverbials: *In years* and *until*. *Linguistic Inquiry* 52(1). 89–142. doi: https://doi.org/10.1162/ling_a_00368.

Iordăchioaia, Gianina & **Frank Richter**. 2015. Negative concord with polyadic quantifiers: The case of Romanian. *Natural Language and Linguistic Theory* 33. 607–658. doi: 10.1007/s11049-014-9261-9.

Jin, Yanwei & Jean-Pierre Koenig. 2020. A cross-linguistic study of expletive negation. *Linguistic Typology* 25(1). 39–78. doi: https://doi.org/10.1515/lingty-2020-2053.

Jin, Yanwei & Jean-Pierre Koenig. 2021. The grammatical representation of expletive negation. In Stefan Müller & Nurit Melnik (eds.), *Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar, Online (Frankfurt/Main)*. 47–67. Frankfurt/Main: University Library. doi: 10.21248/hpsg.2021.3.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer Academic Publishers.

Kamp, Hans, Josef von Genabith & Uwe Reyle. 2011. Discourse representation theory. In Dov Gabbay & Franz Günthner (eds.), *Handbook of philosophical logic*, vol. 15, 125–394. Dordrecht: Reidel 2nd edn.

Kilpatrick, F. P. & Hadley Cantri. 1960. Self-anchoring scaling: A measure of individuals' unique reality worlds. In *Journal of Individual Psychology* 16, 158–173.

Krifka, Manfred. 1994. The semantics and pragmatics of weak and strong polarity items in assertions. In *Proceedings of SALT IV*. 195–219.

Krifka, Manfred. 2010. How to interpret expletive negation under *bevor* in German. In Thomas Hanneforth & Gisbert Fanselow (eds.), *Language and logos: Studies in theoretical and computational linguistics*, 214–236. Berlin: Akademie.

Lamoure, Nicolas. 2020. Nominal Coordination in Italian, French and German: Internal and external agreement of

B
05

D>N>&>N constructions. Doctoral dissertation. Frankfurt a.M.: Universitätsbibliothek. URL: https://nbn-resolving.org/urn:nbn:de:hebis:30:3-574727

Linebarger, Marcia Christine. 1980. *The grammar of negative polarity*: MIT dissertation. Cited after the reproduction by the Indiana University Linguistics Club, Indiana, 1981.

Liu, Mingya. 2012. *Multidimensional semantics of evaluative adverbs*, vol. 26 (Current Research in the Semantics-Pragmatics Interface (CRiSPI)). Leiden: Brill.

Nugent, William R. 2004. A validity study of scores from Self-Anchored-Type Scales for measuring depression and self-esteem. *Research on Social Work Practice* 14(3). 171–179. doi: 10.1177/1049731503257879.

Onea, Edgar & Manfred Sailer. 2013. All that clear? In Eva Csipak, Regine Eckardt, Mingya Liu & Manfred Sailer (eds.), *Beyond "any" and "ever": New explorations in negative polarity sensitivity*, 323–350. Berlin: De Gruyter.

Potts, Christopher. 2005. *The logic of conventional implicatures* (Oxford Studies in Theoretical Linguistics 7). Oxford: Oxford University Press.

Richter, Frank & Janina Radó. 2014. Negative polarity in German: Some experimental results. *Journal of Semantics* 31(1). 43–65. doi: 10.1093/jos/ffs023. URL http://dx.doi.org/10.1093/jos/ffs023.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of Lexical Resource Semantics. In Arne Beckmann & Norbert Preining (eds.), *ESSLLI 2003: Course material I*, vol. 5 (Collegium Logicum), 87–143. Vienna: Kurt Gödel Society Wien.

Richter, Frank & Jan-Philipp Soehn. 2006. *Braucht niemanden zu scheren*: A survey of NPI licensing in German. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar, Varna 2006*. 421–440. doi: 10.21248/hpsg.2006.23. URL http://cslipublications.stanford.edu/HPSG/2006/richter-soehn.pdf.

Richter, **Frank**, Fabienne Fritzinger & Marion Weller. 2010a. Who can see the forest for the trees? Extracting multiword negative polarity items from dependency-parsed text. *Journal for Language Technology and Computational Linguistics* 25(1). 83–110.

Richter, **Frank**, **Manfred Sailer** & Beata Trawiński. 2010b. The Collection of Distributionally Idiosyncratic Items: An interface between data and theory. In Stefaniya Ptashnyk, Erla Hallsteinsdóttir & Noah Bubenhofer (eds.), *Korpora, Web und Datenbanken. Computergestützte und korpusbasierte Methoden in der Phraseologie, Phraseografie und der Lexikografie*, 245–261. Hohengehren: Schneider Verlag.

Rizea, Monica-Mihaela & **Manfred Sailer**. 2020. A constraint-based modeling of negative polarity items in result clause constructions in Romanian. *Lingvisticae Investigationes* 43(1). 129–168. doi: https://doi-org.proxy.ub.uni-frankfurt.de/10.1075/li.00042.riz.

Sailer, **Manfred**. 2021. Minimizer negative polarity items in non-negative contexts. In Stefan Müller & Nurit Melnik (eds.), *Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar*. 348–368. Frankfurt/Main: University Library. doi: 10.21248/hpsg.2021.19. URL https://proceedings.hpsg.xyz/article/view/424.

Sailer, **Manfred**. 2022. Use-conditional licensing of strong Negative Polarity Items. In Daniel Gutzmann & Sophie Repp (eds.), *Proceedings of Sinn und Bedeutung 26*. 734–752. University of Bochum. URL https://bit.ly/Proceedings-of-SuB-26.

Sailer, Manfred & Eva Csipak. 2011. Negative Polarity Items im Englischunterricht. In Andreas Krafft & Carmen Spiegel (eds.), *Sprachliche Förderung und Weiterbildung — transdisziplinär*, vol. 51 (Forum Angewandte Linguistik), 139–156. Frankfurt: Peter Lang.

Sailer, **Manfred** & **Frank Richter**. 2021. Negative conjuncts and negative concord across the board. In Berthold Crysmann & Manfred Sailer (eds.), *One-to-many relations in morphology, syntax, and semantics*, 175–244. Berlin: Language Science Press. doi: 10.5281/zenodo.4729806.

van der Sandt, Rob A. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9. 333–377.

Schaebbicke, Katharina, Heiko Seeliger & Sophie Repp. 2021. The diverse landscape of negative polarity items: On the use of German NPIs as experimental diagnostics. *Journal of Psycholinguistic Research* 50. 1461–1486. doi: doi.org/10.1007/s10936-021-09793-0.

Schlenker, Philippe. 2012. Maximize presupposition and Gricean reasoning. *Natural Language Semantics* 20(4). 391–429. doi: 10.1007/s11050-012-9085-2.

Schlenker, Philippe. 2013. Supplements within a unidimensional semantics II: Epistemic status and projection. In Seda Kan, Claire Moore-Cantwell & Robert Staubs (eds.), *Proceedings of the fortieth annual meeting of the North Eastern Linguistic Society*, vol. 2. 167–182. Amherst, MA: Graduate Linguistics Student Association.

Sedivy, Julie. 1990. Against a unified analysis of negative polarity licensing. *Cahiers Lingistiques D'Ottawa* 18. 95–105.

Smith, Suzanne. 2022. A corpus-based study on time-period Negative Polarity Items. BA thesis, Goethe University Frankfurt.

Sorace, Antonella. 1992. Lexical conditions on syntactic knowledge: Auxiliary selection in native and non-native grammars of Italian. Ph.D. dissertation. University of Edinburgh.

Stevens, Stanley Smith. 1946. On the theory of scales of measurement. *Science* 103(2684). 677–680. doi: 10.1126/science.103.2684.677.

Tahar, Chloé. 2021. Apprehensive and frustrative uses of *before*. In *Proceedings of SALT 31*. 606–628.

Trawiński, Beata & Jan-Philipp Soehn. 2008. A multilingual database of polarity items. In European Language Resources Asso-

ciation (ELRA) (ed.), *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*. Marrakech, Morocco.

Trawiński, Beata, Jan-Philipp Soehn, Manfred Sailer & Frank Richter. 2008. A multilingual electronic database of distributionally idiosyncratic lexical items.

Zwarts, Frans. 1998. Three types of polarity. In Fritz Hamm & Erhard W. Hinrichs (eds.), *Plurality and quantification*, 177–237. Dordrecht: Kluwer Academic Publishers.

3.4   PROJECT PLAN

A key insight regarding the formulation of a theory of NPIs and the negation properties of their licensing environments is expressed in a recent experimental study by Schaebbicke et al. (2021). In light of the observed highly idiosyncratic licensing characteristics of individual NPIs which might deviate very considerably from the few highly researched token NPIs in the theoretical literature, a reliable identification of classes of NPIs in terms of their licensing requirements needs a firm empirical footing. Without close investigation, classifications might be misled by the peculiar behavior of a very small group of items after whose contextual negation requirements the classes are modeled.

The general lack of such an empirical basis is especially salient in the theoretically most deeply researched language, English. Schaebbicke et al. (2021: 1462) observe that "(theoretical) research often appears to limit its scope to the investigation of very few NPIs." While von Bergen & von Bergen (1993) list about 110 English NPIs, their collection has had little impact on the theoretical discussion. In fact, other sources list only around three dozen English NPIs, such as Huddleston & Pullum (2002: 823) or a file by John Lawler dated 2005 and still available online (https://websites.umich.edu/~jlawler/NPIs.pdf, accessed May 8th, 2023).

Huddleston & Pullum (2002) also comment that it is not possible to state a comprehensive list for English. The need for broad and careful empirical foundations is particularly pressing in light of the limitations of traditional introspective grammaticality judgments when faced with the very subtle, inherently gradient acceptability status of NPIs across contexts of various types of "negativity." This makes NPI classification susceptible to error due to researchers' preconceptions of which environments should be proper licensing contexts for certain sets of NPIs. Schaebbicke et al. (2021) rightly caution that the study of a large number of NPIs is a necessary first step to a better understanding of the nature of negative polarity sensitivity.

Moreover, the true extent of NPIs is yet unknown for any language, let alone the nature of their licensing environments. Building on our own prior research, these observations translate into a number of concrete measures:

We confront the empirical challenge by building our first of two work packages around a variety of methods, all oriented at the entirety of data to be studied: corpora serve NPI discovery, observation of naturally occurring licensing contexts, and the creation of quantitative profiles of NPI licensing. Experimental methods are necessary to address the limitations of data. They include acceptability investigations of unattested environments as well as understanding the status of apparently questionable corpus data. The findings inform the grammar architecture, which will in turn guide further empirical study. This framework of thinking about NPIs is reflected in the design of our two main work packages, each with three subsections:

The first work package is dedicated to the empirical foundations and employs different methods with the goal of securing the best possible overview of relevant licensing behavior of a very large set of items in three languages. The project languages were selected for both project internal and external reasons: CoDII already provides solid foundations with the currently largest collection of NPIs in its German section (165), and a sufficiently large collection for Romanian NPIs (58), for which we have both external and internal expertise available. A collection of English is a logical extension of CoDII, as English is on the one hand the language with the most extensive research of NPIs, but on the other hand, as mentioned above, this research usually only looks at a small fraction of the English NPI inventory. Consequently, an inclusion of English guarantees exposure of CoDII to the international research community. Building a large collection of English NPIs and making it systematically available will support the strengthening of the empirical basis of NPI research beyond our focus on German and Romanian. The structuring of data collections in the first work package feeds into the grammatical theory pursued in the second work package. The second work package is centered around the idea of describing NPI licensing in terms of enriched semantic representations for modeling at-issue and non-at-issue content in the unified format of a constraint-based grammar framework. This core theory will be applicable to English, German, and Romanian without any architectural differences.

A theory-driven combination of quantitative and qualitative data is necessary to avoid a misclassification of NPIs on the basis of infrequent but systematic occurrences in non-negative contexts. The systematic integration of statistically meaning-

ful amounts of data of this kind will go hand in hand with a thorough update of CoDII that is based on insights from other theoretically oriented database projects (Dimitriadis & Musgrave 2009; Dimitriadis 2010). The project is arranged in two major, closely related work packages, which differ in their primary goals and methods.

WP1: Triangulation of types of data collection The first work package will critically reflect on the types of available data, how they can and should be used in the project, and what is the best format to make the sum of all collected data on NPI licensing in the three project languages accessible to the research community in such a way that they are maximally useful for further qualitative and quantitative research and analysis. As we observed in the research rationale for this project, we assume that current theories of NPIs suffer from severe sparsity of relevant data by neglecting the large majority of existing NPIs whose diverse clustering behavior may defy classical classifications informed by only a possibly very peculiar set of items with their own idiosyncratic behavior. The way to remedy this situation is seen in a systematic broadening of the empirical base as a foundation for the theoretical analysis of WP2.

The work on extending the German and Romanian collections and on establishing an English collection will be informed by the patterns and categories we observe in our previously constructed collections. It will also benefit from the earlier work on NPI extraction from corpora. With CoDII being an established resource in the field, we already have substantive experience with a large NPI data resource among the PIs, and there is a rich basis of data, data analysis and structured resources to start with. NPIs in the three main languages of the projects (English, German, and Romanian) have a different degree of documentation: There is a large, systematic collection of German data which includes pointers to relevant non-negative contexts. For Romanian there is a solid base of empirical and theoretical work with our cooperation partners (Iordăchioaia & **Richter** 2015; Rizea & **Sailer** 2020), but closer work, including a Romance linguistics perspective, is required for the discussion of NPIs in non-negative contexts. While the literature on NPIs in English is extremely rich in studies of individual NPIs and isolated phenomena, a broad, systematic, up-to-date, theoretically interpreted corpus-based investigation of NPIs in English is a pressing desideratum.

Existing theoretical work on English and, to a lesser extent, on German, can guide theory development for Romanian; the data situation in German and, to a lesser extent, in Romanian, can guide methods and expectation on the search space for data acquisition in English. The different past research foci with varying emphasis on deep analysis of a few items vs. much wider data horizon elsewhere make the selected three languages ideal for our goals and will help us develop and test the cross-linguistic applicability of our empirical methods.

We will use a three-pronged approach of methods in order to achieve an empirically sound overview of NPIs and to construct a database for further future analysis:

- WP1.1 (Richter): Transformation and extension of CoDII
- WP1.2 (Sailer): Manually compiled quantitative and qualitative corpus profiles
- WP1.3 (Richter & Saier): Systematic collection of speaker judgments

WP1.1 (Richter): The first subsection of WP1 is concerned with a theory-driven extension of the CoDII database of NPIs. It is the foundation for the other two subsections of this work package, as well as for determining the grammar architecture of WP2. The German collection comprises 165 annotated and categorized entries with a documentation of attested corpus occurrences in 14 licensing environments (dating from the 2005–2008). From semi-automatic NPI extraction from corpora in **Richter** et al. (2010a) we know of 112 more German NPIs, in addition to others, manually collected since, that need to be integrated. We will use the patterns observed in the German and Romanian collections within CoDII and beyond to pursue theory-driven NPI candidate extraction with large present-day corpora, with a new focus on gathering data for non-at-issue licensed items. Based on more recent work on quantitative profiles with additional distinctions by Hoeksema and by Sailer, we extend and redesign the listed licensing contexts, also taking into account possible ways of classifying types of licensing by non-at-issue negation in the absence of overt syntactic or lexical licensing. The new database should also contain standardized quantitative profiles for all collections wherever this is feasible, as well as documentation of their corpus analysis. The present XML-based format of CoDII does not support easy exploration of the data it contains. Database design guidelines for linguistic resources (Dimitriadis, 2010; Dimitriadis & Musgrave, 2009) will be used to transform the three new collections for German, Romanian and English into searchable databases whose analyses and quantitative data can easily be exported and visualized with an appropriate software pipeline with standard tools.

The data format translation will of course also be applied to Romanian, to keep it formally synchronized with the German collection, and the new English NPI section will be modeled after the other two. Building up the English collection should benefit from experience with German and Romanian that could point toward semantic and pragmatic properties of NPI clusters to find corresponding elements in English. Finally, the semi-automatic NPI extraction methods previously used for German can be adapted to English to support the construction of the English database.

WP1.2 (Sailer): Following the method established in many studies by Jack Hoeksema, motivated in Hoeksema (1998), Sailer has been collecting quantitative corpus profiles of NPIs in addition to the qualitative profiles documented in CoDII up to now. The relevance and potential of detailed profiles becomes particularly clear when contrasting NPIs that seem to be similar in their meaning or their general classification in the literature. For example, Smith (2022) contrasted the English NPIs *in a million years* and *in ages*. They are both traditionally classified as strong NPIs, as in Krifka (1994). However, their corpus profile reveals considerable differences, with *in ages* showing gaps in some known licensing contexts of strong NPIs, such as *without* clauses and rhetorically used *wh* questions. Sailer & Smith presented these observations in a poster at *Sinn und Bedeutung 27, 2022*.

Similarly, while *any* and *ever* are both considered weak NPIs and can, indeed, occur in the same contexts, they differ considerably with respect to their distribution in these contexts. In quantitative profiles based on the BNC, Sailer & Csipak (2011: 141) found a dominance of occurrences of *ever* in *yes/no* questions and with comparative and superlative. In contrast to this, *any*, although possible in these contexts, is not that frequent in them.

Register or corpus type seems to play an additional role. For example, in a recent pilot investigation (slides available at: https://tinyurl.com/MucksZweig), Sailer found a considerable percentage of occurrences of the German minimizer NPI *einen Mucks machen* (a peep make) 'make a peep' in *before* clauses in the written corpora of the IDS, which are primarily newspaper texts. In the *Timestamped JSI web corpus German 2014–2021*, however, there is only a single occurrence of this NPI with *bevor* 'before.'

In WP1.2 we will systematize the manual extraction and annotation of qualitative and quantitative corpus profiles and develop an annotation scheme that can be used, first for an openly accessible documentation of both types of profiles, and second for a potential later automatic extraction method, informing the work on redesigning CoDII of WP1.1. The contexts considered in the project require an extension of the annotation scheme in Trawiński & Soehn (2008) and Trawiński et al. (2008) that underlies CoDII.

WP1.3 (Richter & Sailer): A classical—but certainly not invalid—reservation towards corpus data states that they are neither complete nor fully reliable, on the one hand, the absence of a given structure in a given corpus cannot be used as evidence in favor of its ungrammaticality, on the other hand its presence cannot be taken to mean that it is grammatical. This inherent deficiency of corpora has led many researchers interested in grammar of natural human languages, especially generativists, away from corpora towards introspection / speaker judgment as their main method of investigation. However, speaker judgments fall short in other aspects: the amount of questions possible is severely limited and hence they are not really suitable for large scale, explorative investigations into a phenomenon. The kind of enterprise aimed for by our project falls right into the blind spot of both these methods. On the one hand NPIs and the critical contexts for our project are not very frequent and thus are prone to fall prey to the unreliability of corpora (especially web corpora that cannot be controlled for nativeness). Whereas, on the other hand, the sheer number of questions resulting from the explorative nature of our enterprise makes speaker judgments not feasible. As the well-formedness of many of our data depends on specific contexts the number of variables would be too high to check by speaker judgments. Furthermore, some contexts are considered valid licensing contexts for strong NPIs only in particular readings. Typical examples are rhetoric versus information seeking readings of *wh* questions, or rule-like versus episodic readings of *if* clauses. For the lack of corresponding annotation in our corpora, this means that every sentence needs to be judged for its reading in order to classify the licensing type.

Generally, WP1.3 will address these concerns by combining both—corpora and speaker judgment—in order to cancel out the weaknesses of each. We will use corpus data for exploratory purposes, exploiting the fact that a corpus can be queried indefinitely, but turn to speaker judgments for validation of our findings (Lamoure 2020 for the same methodology). More specifically, to make the data and its classification as reliable, intersubjectively robust, and experimentally replicable as possible, we will integrate the following measures: The first measure to achieve this goal is to always include sufficient meta-data with the data point in our CoDII collections of NPIs. Second, we will establish a gender-balanced group of competent native speakers for each language, controlling for their linguistic biography. The group members will be tasked with judging extracted corpus sentences for acceptability along a 7 point scale. We chose an N-point scale as opposed to other methodologies (e.g., magnitude estimation task, cf. Stevens 1946; Sorace 1992; Bard et al. 1996) or thermometer method task (Kilpatrick & Cantril 1960; Nugent 2004; Featherston 2007) as it is inherently simpler and requires less instructions. The reason for choosing a 7 point scale is derived from a respective preference on the informants' side for scales equal to or greater than 5 (Bard et al. 1996: 45) and the hope to thus reduce quantization effects (i.e. errors caused by rescaling from one's preferred scale to the one imposed by the researchers). Moreover, informants will be checking available inferences (to determine the perceived reading of a sentence). We intend to commit ourselves to including in our database and in our publications only sentences for which we get at least five such judgments and to

B
05

report the results of these judgments. This will provide users of our database with an indicator for the acceptability of the given items and with a further feature for querying the database. To offer informants greater flexibility, testing will be done incrementally, in waves of about 20 items to reduce overloading participants, and by using an online platform.

In order to guarantee the availability of five informants at any stage, our informant groups should contain at least seven members per language. We will need permanent informant groups for English, German, and Romanian. We will also try to establish an informant group for French, which might be needed for WP2.1 as detailed below. The task of setting up the acceptability judgment groups for the project languages will be of high initial priority, and work with the group members will be actively maintained from the start and throughout. The activities with the Romanian group will require regular visits of project members to Romania. Dr. Lamoure will support the project in this working package, in particular with respect to data from Romance languages.

The main results of WP1 will be the following: We will expand and transform CoDII systematically with the goal of increasing its visibility in the NPI research community, and its usefulness and versatility as a resource for empirically well-founded NPI research. We will (i) convert CoDII into a searchable database for quantitative and qualitative analysis, (ii) update CoDII with the data from Richter et al. (2010a), (iii) create a database for English NPIs, and (iv) use (ii) and (iii) as the necessary basis to systematically extract, document, and analyze NPIs in non-negative contexts. As a result of this work, it should no longer be acceptable to base wide-ranging theories of NPIs on just a handful of examples from a single language.

WP2: NPI licensing in enriched semantic representations The second work package pursues the integration of empirical observations concerning the two challenges into a general constraint-based theory of NPI licensing. Meaning as part of licensing environments will be rendered as logical expressions of a higher order logical representation language. The main grammar-theoretical innovation of WP2 will be the creation of a constraint-based grammar architecture which unifies the representation of at-issue and non-at-issue content in a single framework such that an NPI licensing theory can reference both in the formulation of appropriate licensing conditions for different classes of NPIs. We will proceed in three subpackages:

- WP2.1 (Richter): NPI licensing in overtly non-negative constructions: *before* clauses and expletive negation
- WP2.2 (Sailer): Types of NPIs that are licensed in non-negative contexts
- WP2.3 (Richter & Sailer): Towards a constraint-based modeling of at-issue and non-at-issue NPI licensing

WP2.1 (Richter): It is natural to assume that non-negative contexts which license NPIs have some hidden negativity, as done explicitly in Linebarger (1980) and **Sailer** (2022). The same assumption has been made about contexts that allow expletive negation (see, for example Jin & Koenig (2020, 2021)). On the other hand, Iordăchioaia & **Richter** (2015: 639), for example, argue that NPIs cannot be licensed by expletive negation. This makes contexts with expletive negation an ideal area of study for our project. Rather than investigating expletive negation on its own, we will focus on a context in which expletive negation and NPI licensing in non-negative contexts can be studied together in the project languages. By doing so, we hope to learn more about the nature of negation in NPI licensing, covert NPI licensing environments on the one hand, and the negation requirements of NPIs on the other. *Before* clauses are known to license expletive negation in many languages, including German (Krifka 2010). Tahar (2021) convincingly identifies three types of *before* clauses, which she calls *consecutive* ('first X, then Y'), *apprehensive* ('X so that not Y') and *frustrative* ('(not) X until Y'). Consecutive *before* is factual and does not license expletive negation in any language. The other two uses allow for an expletive occurrence of the negative marker *ne* in French. In German, NPIs and expletive negation seem to be in complementary distribution in *before* clauses: we find an expletive negation with frustrative cases only (9-a), and NPIs exclusively in apprehensive uses (9-b).

(9) a. Denn die Unternehmensleitung von Goodyear Frankreich weigert
because the management of Goodyear France refuses

sich, mit den Gewerkschaftsleuten zu sprechen bevor diese nicht
itself with the union.people to talk before these not

die beiden Manager frei lassen.
the two managers free let
'Because Goodyear's management in France refuses to negotiate with the unions before they
release the two managers.' (Timestamped 2014–2021, German)

b. Wenn jemand in   den   Raum kommt, macht   man   sich
   when someone in   the   room   comes   makes   one   themselves

ein Bild   von ihm, bevor er auch nur   einen Mucks machen
a   picture of   ihm,. Before he even only a   peep   make
'When someone walks into the room, you form an impression of them before they
even make a peep.' (IDS corpora written German)'

For example, the German minimizer NPI *einen Mucks machen* 'make a peep' is not attested with weak NPI licensers like *few* but can occur in apprehensive *before* clauses as (9-b). English does not allow for expletive negation in *before* clauses, but shows *any*-type NPIs in apprehensive *before* clauses only, and hardly any occurrence of minimizer NPIs with *before*. The pattern in Romanian and French (added to this section of WP2 as a second Romance language) is further complicated by at least two factors: (i) the preverbal negative marker has expletive uses, (ii) the translation equivalents of many minimizers in German or English have marginal uses as neg-words in French, such as *ne ...dire mot* 'say nothing' (lit.: 'NE ...say word'). We plan to explore the distribution of expletive negation and various types of NPIs in *before* clauses in the three main project languages and in French. The context at hand exhibits considerable cross-linguistic variation and, for that reason, constitutes a potential challenge to theories of expletive negation and NPI licensing. Only through the careful study of a series of NPIs in the different types of *before* clauses outlined above can it be determined whether the contrasts reported here should be attributed to cross-linguistic differences in the meaning of *before* or to item-specific differences in the meaning of the NPIs.

In the preparation of WP2.1, Lamoure and Richter have conducted a pilot online processing study (self paced reading) together with Yvonne Portele (PI in C04) on the inferential effects of expletive negation in *before* clauses in German. Preliminary results suggest that (i) the position of the expletive negation in *before* clauses has no impact on the derived meaning and that (ii) reaction times may indicate the existence of different processing patterns. We are planning to use primarily corpus data in this work package, enhanced by native speaker evaluation (as described in WP1.3).

WP2.2 (Sailer): In the second section of work package 2, we will take the opposite perspective from WP2.1. Instead of starting from a given licensing environment, we will start from a set of NPIs and will study the occurrence patterns of individual NPIs to identify their distributional patterns. After all, systematic occurrences of NPIs in non-negative contexts put a big question mark behind any definition of an NPI as an item that must be in the scope of a negation or a "sufficiently negation-like" operator.

In a pilot study, Sailer uses quantitative and qualitative data on two NPIs from CoDII in the written corpora of the IDS. He compares the occurrence profile of *einen Mucks machen* 'make a peep' with *auf einen grünen Zweig kommen* 'be ok (financially, professionally, ...)' (lit.: 'get on a green twig.'). The data suggest that *einen Mucks machen* must be in the scope of a negation, but the licensing negation can reside in the at-issue content, in some relevant presupposition, conventional implicature, or generalized conversational implicature, in line with the analysis proposed in **Sailer** (2021, 2022). The expression *auf einen grünen Zweig kommen* shows a slightly, but remarkably different distribution in the corpus. It can occur in non-negative contexts such as inside purpose clauses (i.e., cases like *um auf einen grünen Zweig zu kommen* 'in order to get on a green twig.') Sailer proposes that the expression contributes a negative presupposition and is licensed in any context that does not contradict that presupposition. This looks similar to what was claimed for *all that A* in Onea & Sailer (2013), as in *It isn't all that hot*: Onea & Sailer (2013) argue that *all that A* asserts that the property A holds to a degree *d*, but presupposes that the speaker assumes that there is a high degree *d* such it is not hot to that degree but that someone believes it was. The preference of *all that A* to occur in NPI-licensing contexts follows from an avoidance of a contradiction between the presupposed and the asserted content. A similar approach can be pursued for *auf einen grünen Zweig kommen*. It is clear that detailed systematic studies of NPIs are required to determine which factors contribute to each item's predominant occurrence in negative or negation-like contexts.

A related issue arises with ambiguous items, which are not infrequent among the NPIs in CoDII. Some expressions seem to have NPI and non-NPI uses with distinguishable readings. The occurrence of NPIs in non-negative contexts makes the question of whether these readings can be subsumed under a single lexical description highly relevant for theories of NPIs. Lamoure and Sailer have conducted a preliminary study on the occurrence of *ever* with superlatives. They argue that uses of *ever* as in *the best book ever* should be considered universal uses of *ever* (as in *happily ever after* or *ever so gently*). This contrasts with the commonly considered, existential NPI cases like *the best book that I have ever read*. It is, then, an open question if there can be a single lexical description of *ever* that is not an NPI or whether there are sufficiently compelling reasons to postulate an existential NPI *ever* and a universal non-NPI *ever* as two separate lexical items. An analogous discussion for *until* is found in Iatridou & Zeijlstra (2021).

In WP2.2 such investigations of individual NPI candidates will be pursued systematically, primarily based on corpus data with additional evaluation by our informant groups.

WP2.3 (Richter & Sailer): In the third section of WP2 we will be working on covering the results of our empirical work in a theory formulated in a constraint-based grammar framework. Since the current state of research as well as our preliminary studies strongly suggest that there is an interaction of various types of meanings (asserted at-issue, presupposed, conventionally and conversationally contributed) and of properties of individual words or constructions, we employ a framework that provides simultaneous access to syntactic, semantic, and pragmatic information. As a consequence, principles of grammar can be formulated which take all of them into account simultaneously. For example, a licensing condition of an NPI may be determined by semantically or pragmatically given negativity within a certain syntactic context. *Lexical Resource Semantics* (LRS, Richter & Sailer (2004)) is a suitable framework for this purpose and has been implicitly assumed in our recent work on NPI licensing in non-negative contexts (**Sailer** 2021, 2022). We pursue the assumption that NPI data support new arguments to justify the integration of enriched semantic representations into a constraint-based grammar framework. To substantiate this claim, we will provide a formal account of the apparently non-negative contexts that proved to be relevant in the other sub-WPs of this project in terms of at-issue vs. non-at-issue content. This account will enable us to model various subclasses of NPIs and capture the distributional behavior identified and documented in WP1.

The insights collected in WP1 and WP2 will support a more informed perspective on NPIs. Questions that arise in the development of a formally precise architecture and the grammar-theoretic description of classes of NPIs can feed back into the investigation of (non-) negative environments (such as *before* clauses in WP2.1), or determine the choice of (groups of) NPIs targeted by the in-depth studies (as in WP1.2).

At the same time, the project will make an important contribution to the central research questions of the CRC. Restricting the scope of these questions to NPI licensing, the following potential answers can be considered the working hypotheses of B05:

QB.1-B05: What commonalities between negation and other NPI licensers are universal? All NPI licensing operators and contexts involve some kind of semantic negation, be it at-issue or non-at-issue. While morpho-syntactically marked clausal negation comprises relevant at-issue and non-at-issue properties, other NPI-licensing operators or contexts may have only one of them, which results in different NPI-licensing potential.

QB.2-B05: To what extent do negation and other NPI licensers behave similarly or differently within and across languages with respect to NPI licensing? We expect to find systematic cross-linguistic differences with respect to the contexts that are not strictly at-issue negative, as exemplified in WP2.1 with *before* clauses in French vs. German.

The answer to the two questions supports the Neg-Plus Hypothesis if we can show that the licensers share underlying properties. There is support for the Neg-Only Hypothesis, if the data are better described in terms of independently motivated context requirements of the NPI.

Cooperation and field work Based on previous research on Romanian negation and NPIs by Richter and Sailer and drawing on a wealth of empirical and theoretical studies on Romanian NPIs, we plan an intense cooperation with Dr. Monica Mihaela Rizea Casa (University of Bucharest, Romania) and with Dr. Gianina Iordăchioaia (HU Berlin). The project will also allow Lamoure and one of our project employees to develop research expertise in Romanian. This will give them a unique advantage in Romance studies in Germany, as Romanian is a lesser represented language in linguistics.

In **year 1 (2024)**, Rizea Casa should visit our research unit for one month during which we plan the future steps for fieldwork visits of Lamoure and of our project employee with a specialization in Romance linguistics. The two project members will, then, spend one month in Bucharest to prepare and carry out the first round of work with informants. In **year 2 (2025)** and **year 3 (2026)** each, the Romance-language project employee will spend one month in Bucharest to collect the data they need for their thesis and for the Romanian part of the project. In **year 4 (2027)**, the project will benefit from a second visit by Rizea Casa in Frankfurt. She will provide final critical feedback for our Romance-language project employee, and we will discuss the future cooperation in a potential second project phase. Dr. Rizea Casa's visits will be covered by the project's share of the budget for guests applied for in Project Z.

During the first project phase, we intend to invite experts on empirical and theoretical work on polarity items. We plan one extended visit of at least 3 days per year. Candidate guests are Jack Hoeksema (Groningen), Gianina Iordăchioaia (HU Berlin), Mingya Liu (HU Berlin), Katharina Schaebbicke (Cologne).

	2024				2025				2026				2027			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
WP1.1	■	■	■	■	■	■	■	■	■	■	■	■	■			
WP1.2	■	■			■	■			■	■			■			
WP1.3	■				■				■				■	■		
WP2.1		■	■							■			■			
WP2.2							■	■			■			■		
WP2.3	■			■				■				■		■		■
Rizea Casa		■												■		
Fieldwork			■		■				■							

Table 1. Gantt chart for the 2 planned work packages, visits and fieldwork

WP1.1 will be running through most of the project phase, incorporating the results of the other work package sections. The innovative extrapolation of the corpus extraction methods to English and Romanian is scheduled primarily for year 2 (2025), which is the reason for having less activity in the other working package sections co-coordinated by Richter. We plan to have the results of this section by Q3/2026 in order to make the resulting database publicly available during the project phase and to prepare publications on the theoretical and methodological aspects. **WP1.2** will always provide the corpus basis for the studies of the project. Therefore, it will be at the beginning of each year and precede the theoretical interpretation. **WP1.3** will evaluate the data harvested in WP1.2 and every fieldwork trip is primarily associated with research activities of this work package section.

The main part of **WP2.1** is scheduled for year 1 (2024). In year 3 (2026), we will reconsider these results in the light of the research on other phenomena, and plan to finalize a publication of the results. In year 4 (2027) we will carry out pilot studies for a potential continuation. The phenomena considered in **WP2.2** will be in the focus of year 2 (2025). In year 3 (2026) we will reconsider these results in the light of the research on other phenomena, and plan to finalize a publication of the results. In year 4 (2027) we will carry out pilot studies for a potential continuation. **WP2.3** will play a role in preparing the empirical work and in the theoretical interpretation of the data. Consequently, this section of WP2 will be mainly focused on at the beginning and the end of each project year and in the final year (2027).

All **fieldwork** trips are planned for the term break, around September in year 1 (2024), and in March for the other project years.

3.5 ROLE WITHIN THE COLLABORATIVE RESEARCH CENTRE

The project addresses central research questions of the CRC: First, NPI licensing and the definition of negation and negation-like operators are closely related (see QB.1 and QB.2). The data considered in the project presents fundamental challenges to existing characterizations of negation-like operators. Second, by contrasting explicit and implicit ways of expressing negation, the project can gain insights into what set of properties are associated with explicit negation marking (Neg-Plus Hypothesis). Third, building on CoDII, the most exhaustive publicly accessible collection of NPIs, the project contributes to the CRC's mission to provide a solid and systematic empirical basis for theoretical linguistics.

In interaction with A01 (Ebert/Fuhrmann) we will investigate the relationship between the canceling operations investigated in A01 and the non-negative NPI licensing contexts considered in B05. While they seem to overlap as in (2a) above, they might not be identical (see NPI-licensing in *before* clauses).

The project complements the work on NPIs in other projects: A02 (Meier/Weiß) look at derived adjectives in German that require *un*-prefixation, which they in part consider as NPIs below the word level. We envisage an exchange between the two projects at the methodological and the theoretical level as both take corpus profiles of potential NPIs as their empirical basis, and both pursue as one possible analytical strategy that the negation-requirement is a consequence of the interaction of different meaning components. For A04 (Sailer/Zeijlstra), NPIs are just one diagnostic reflex of Neg Raising. Since A04 investigates the extent to which Neg Raising is dependent on contextual rather than lexical or structural properties, its results on NPIs in Neg Raising are of interest to B05. Analogously, the investigation of pragmatic licensing of NPIs in B05 can help put findings related to NPIs in Neg Raising into a larger context. Finally, B01 (Bader/Bargmann/Webelhuth) looks at the alternation of *nicht ein* 'not a' and *kein* 'no' in German, which is also found in NPIs—in particular in minimizers such as *einen Mucks machen* 'make a peep.' This group of NPIs will certainly be included in the ones looked at in detail in