



Universität Tübingen · Wilhelmstr. 50 · 72074 Tübingen

Referát doktorského studia
Filozofická fakulta Univerzity Karlovy
nám. Jana Palacha 2
116 38 Praha 1

Slavisches Seminar

Prof. Dr. Tilman Berger
Slavische Sprachwissenschaft

Slavisches Seminar
Wilhelmstraße 50
72074 Tübingen
Telefon +49 7071 29-76733
Telefax +49 7071 29-5924
tberger@uni-tuebingen.de
www.slavistik.uni-tuebingen.de

Tübingen, 14. November 2023

**Gutachten über die Dissertation von Frau Mgr. Martina Rybová zum Thema
„Die Beziehung zwischen Evaluativität und Indefinitheit am Beispiel der zusammengesetzten
Indefinitpronomina im Tschechischen“**

Die vorliegende Dissertation beschäftigt sich mit einem Thema, das aus der Sicht einer traditionellen strukturalistischen Linguistik ungewöhnlich und merkwürdig gilt. Denn traditionell werden Indefinitpronomina in der Grammatik als Einheiten beschrieben, denen die Kategorie der Unbestimmtheit eigen ist, da denkt man zunächst nicht an Evaluativität, die man eher in den Bereich der Pragmatik einordnen würde. Aber die Tatsache, dass man über ein solches Thema eine Dissertation schreiben kann, zeigt eben, dass sich unser Wissen über Indefinitpronomina in den letzten Jahren sehr weiterentwickelt hat. So ist es nun auch kein Zufall, dass die Dissertation im Kontext des von Björn Hansen geleiteten DFG-Projekts „Zusammengesetzte Indefinitpronomina in slavischen Sprachen. Ein Beitrag zur semantischen Karte der Indefinitheit der zweiten Generation“ entstanden ist. Denn genau dieses Projekt hat die Aufmerksamkeit auf die vielfältigen zusammengesetzten Indefinitpronomina in verschiedenen slavischen Sprachen gerichtet. Und die vorliegende Dissertation nimmt sich die tschechischen Indefinitpronomina vor, die unter verschiedenen Gesichtspunkten, vor allem unter dem der Evaluativität, untersucht werden.

Schon an dieser Stelle möchte ich sagen, dass hier eine hervorragende Arbeit vorliegt, in der die Verfasserin auf einem hohen theoretischen Niveau arbeitet und gleichzeitig eine Vielzahl von empirischen Daten umfassend analysiert hat. Sie erweitert damit unser Wissen über die Bedeutung und Funktion der tschechischen Indefinitpronomina in wesentlicher Weise und gibt auch eine Reihe von Anregungen für weitere Forschungen.

Im Folgenden werde ich nun auf die sechs Kapitel der Arbeit genauer eingehen, kommentiere dabei einzelne Punkte, formuliere Fragen für die Verteidigung der Arbeit und schlage am Ende meines Gutachtens eine Bewertung der Dissertation vor.

Im **ersten** Kapitel „Einleitung und Zielsetzung“ stellt die Verfasserin kurz das Thema vor, stellt auch den Bezug zu dem DFG-Projekt her und skizziert, wie sie vorgehen wird. Das zweite und dritte Kapitel sollen sich mit dem Phänomen der Evaluativität und mit der Kategorie der Indefinitheit und den tschechischen Indefinitpronomina beschäftigen, im vierten Kapitel werden Forschungsfragen und-Hypothesen formuliert, das fünfte Kapitel befasst sich mit den in der Arbeit angewandten Methoden und das sechste Kapitel stellt die konkrete Analyse von elf zusammengesetzten Indefinitpronomina vor.

Das **zweite** Kapitel beschäftigt sich, wie bereits gesagt, mit dem Phänomen der Evaluativität. Die Verfasserin gibt hier einen breiten Literaturüberblick, der mit Bally (1913) beginnt, dann aber vor allem eine Reihe neuerer Arbeiten referiert. Ich muss gestehen, dass ich nur einen kleinen Teil dieser Arbeiten kannte, und ich finde die Art und Weise, in der die Verfasserin die verschiedenen Konzepte mit Bezug auf sechs Hauptaspekte beschreibt (vgl. hierzu die Tabelle auf S. 62f.), sehr beeindruckend. In Abschnitt 2.2 kommentiert die Verfasserin dann kurz verschiedene evaluative Ausdrucksmittel, hier hat mich ein wenig überrascht, dass sie das an deutschen und nicht an tschechischen Beispielen tut. In Abschnitt 2.3 gibt sie schließlich ihre eigene Definition von Evaluativität und positioniert sich hier der von ihr referierten älteren Forschung. – Insgesamt ist dieses Kapitel sehr informativ und gut gelungen. Ein wenig verwundert hat mich allerdings, dass die Arbeiten von Kateřina Veselovská, die sich mit dem Tschechischen beschäftigen, ganz in den allgemeinen Kontext integriert sind und nicht gesondert betrachtet werden, und dies, obwohl sie (anders als die meisten anderen zitierten Studien) im weiteren Verlauf der Arbeit öfter zitiert werden.

Leider findet sich in diesem Kapitel aber auch der einzige Punkt, den ich an der vorliegenden Arbeit wirklich kritisieren muss. Und zwar wird die Leserin bzw. der Leser in einer Fußnote auf S. 27 ganz kurz und gewissermaßen *en passant* darüber informiert, dass die vorliegende Arbeit überwiegend auf den Daten des Webkorpus „csTenTen v 17“. Diese Entscheidung halte ich zwar für völlig richtig (!), aber ich meine, dass die Verfasserin sie schon zu Anfang der Arbeit genauer begründen und vor allem mehr Angaben über konkurrierende Korpora liefern sollte. Entsprechende Informationen werden zwar im Unterkapitel 5.3 „Verwendete Korpora“ (S. 173–176) nachgeholt, ungefähr in der Mitte der Arbeit. Und auch dort werden im Wesentlichen die Korpora beschrieben und die eigenen Überlegungen der Verfasserin angeführt, ein Verweis auf die nicht ganz kleine Literatur zu Webkorpora fehlt aber völlig.

Das **dritte** Kapitel beschäftigt sich mit dem Begriff der Indefinitheit und mit dem Forschungsstand zu tschechischen Indefinitpronomina. In den ersten vier Unterkapiteln wird der Begriff der Indefinitheit, ihre Rolle bei Indefinitpronomina und die Beziehung zur Evaluativität diskutiert. Diese Unterkapitel sind relativ kurz, es wird aber, wenn ich es richtig sehe, alle wesentliche Literatur zum Thema besprochen. Vermisst habe ich hier aber eine Auseinandersetzung mit der Abgrenzung zwischen dem unbestimmten Artikel und Indefinitpronomina. Bekanntlich gibt es im Tschechischen keinen unbestimmten Artikel, in Übersetzungen aus Artikelsprachen wird er aber manchmal durch Indefinitpronomina (insbesondere *nějaký*) wiedergegeben. Wenn man darüber spricht, dass sich „die meisten Beiträge zu dieser Diskussion“ eher „mit Nominalphrasen mit einem unbestimmten Artikel als mit der Referenz anderer indefiniten Ausdrücke“ befassen (S. 70), wären einige kurze Bemerkungen zu dieser Problematik aber angemessen.

Das fünfte Unterkapitel gibt den Forschungsstand zu den tschechischen Indefinitpronomina wieder, die insgesamt nicht so gut erforscht sind, wie es eigentlich wünschenswert wäre. Bemerkenswert ist hier das große Gewicht, das Wörterbüchern beigemessen wird. Dies würde man nicht erwarten, wenn man Indefinitpronomina als ein Phänomen der Grammatik ansieht, mit Hinblick auf die Evaluativität und vor allem auch die Untersuchungsmethode, die die Verfasserin im fünften Kapitel vorschlägt, ist das aber ein sinnvoller Schritt.

Das sechste und siebte Unterkapitel beschäftigen sich mit der Identifizierung und Inventarisierung neuer potenzieller Indefinitpronomina im Tschechischen. Das ist ein sehr interessantes und innovatives Thema, bei dem diese Arbeit wirklich Bemerkenswertes leistet. Die Verfasserin geht hier im Detail auf die unterschiedlichen Faktoren ein, die bei der Identifizierung zu beachten sind, insbesondere Fragen der Getrennt- und Zusammenschreibung, aber auch der Interpunktion. Alle Betrachtungen werden mit Korpusdaten untermauert, die sehr sorgsam erhoben werden. Die Verfasserin gibt immer auch die Suchformeln an und erläutert diese, falls nötig.

Insgesamt ist dieses Kapitel, das im achten Unterkapitel mit Überlegungen zur Auswahl der Indefinitpronomina, die in der Dissertation genauer untersucht werden sollten, abschließt, sehr gut gelungen. Ich erlaube mir trotzdem den Hinweis, dass gerade hier die Beschränkung auf das Korpus *csTenTen* nicht übermäßig plausibel ist und hätte begründet werden sollen. Stattdessen wird auf S. 100 lakonisch gesagt, dass die Suche in diesem Korpus erfolgen soll, das zum letzten Mal in der erwähnten Fußnote auf S. 27 erwähnt wurde.

Das **vierte**, relativ kurze Kapitel enthält die Forschungsfragen und Hypothesen, die im weiteren Verlauf der Dissertation untersucht werden sollen. Die Fragen ergeben sich aus den beiden vorhergehenden Kapiteln, Es geht zunächst darum, ob zusammengesetzte Indefinitpronomina (im Weiteren ZIP) in evaluativen Umgebungen vorkommen, und dann um die Interaktion der ZIP mit den evaluativen Einheiten, um den Vergleich verschiedener Indefinitpronomina und um typische Muster, die hier zu beobachten sind. Und auch die Hypothesen liegen vor dem Hintergrund der Fragen nahe, auch wenn nicht erklärt wird, wie die Verfasserin genau zu ihnen gekommen ist. Das ist aber vielleicht bei Hypothesen auch gar nicht üblich.

Das **fünfte**, wieder etwas längere Kapitel befasst sich mit den Methoden, die in der Arbeit Anwendung finden sollen. Hier macht die Verfasserin einen wirklich sehr originellen und innovativen Vorschlag, den sie an späterer Stelle der Arbeit dann als eine Art „Crowdsourcing“ bezeichnet (S. 214). Sie kombiniert eine Korpusanalyse mit einer raffinierten Verwendung von Wörterbuchdaten. Konkret wird die Evaluativität der Einheiten, die in Verbindung mit den ZIP auftreten, anhand einer raffinierten Analyse von Wörterbuchdaten beurteilt, aber auch viele andere Gesichtspunkte werden berücksichtigt.

Nun aber zu den einzelnen Schritten: Auf einen kurzen Literaturüberblick im Unterkapitel 5.1 folgt zunächst eine Beschreibung der Kollokationsanalyse (Abschnitt 5.2). Diese beginnt mit einem kurzen Durchgang durch die Literatur, wesentlich ist aber die Einführung und Definition weiterer Begriffe wie der ‚Kolligation‘ und der semantischen und evaluativen Prosodie. Wie auch sonst in der Arbeit werden die Begriffe und Verfahren auch anhand von Korpusbeispielen untersucht, wobei mich freilich ein bisschen gewundert hat, dass nun auf einmal polnische Beispiele herangezogen werden. Nicht dass mich das wirklich stören würde, es unterstreicht den slavistischen Charakter der Arbeit und die Qualifikation der Verfasserin, aber etwas merkwürdig ist schon, wenn auf den Seiten 170 und 171 polnische Beispiele vorkommen und danach wieder nur tschechische. In Unterkapitel 5.3 stehen die Angaben zu den verwendeten Korpora, von denen ich, wie schon zweimal gesagt, meine, dass sie deutlich vorher hätten angeführt werden müssen. Hier stehen nun endlich genauere Angaben zum Korpus *csTenTen* und den mit ihm verbundenen Problemen – fasziniert hat mich der Bericht über die „Kasino“-Texte, die mir offengestanden bisher unbekannt waren. Dass der Ablehnung des Tschechischen Nationalkorpus nur weniger als eine halbe Seite gewidmet ist (vgl. S. 176), finde ich ein wenig problematisch, kann die Entscheidung der Verfasserin, so zu verfahren, aber letztlich akzeptieren. In Abschnitt 5.4 trifft die Verfasserin die Entscheidung, jeweils immer nur die erste 500 Kollokationen zu berücksichtigen, dies wird mit statistischen Argumenten begründet, die ich gut nachvollziehen kann.

In Unterkapitel 5.5 wird der Umgang mit den Wörterbüchern vorgestellt. Hier finde ich die Darstellung etwas unübersichtlich, weil der Abschnitt, in dem die Verfasserin begründet, welche Wörterbücher sie ausgewählt hat, sehr tief „eingebettet“ ist (es handelt sich um Abschnitt 5.5.1.2.1!). Hier wäre es wirklich besser gewesen, den Wörterbüchern einen eigenen „höheren“ Abschnitt zu widmen. Ich selbst will hier etwas genauer auf die Auswahl eingehen, da ich sie teilweise problematisch finde. Zunächst ist schon gut nachvollziehbar, warum die Verfasserin das „Czech Subjectivity Lexicon“ an erster Stelle nennt, denn dieses Wörterbuch erfüllt die gestellten Anforderungen sicher am besten, aber es reicht bei weitem nicht aus, weil es nicht so viele Lemmata enthält. Die Verfasserin

hat sich dann an zweiter Stelle für den in den sechziger Jahren entstandenen „Slovník spisovného jazyka českého“ (SSJČ) entschieden, das zweite große Wörterbuch des Tschechischen. Das erste große Wörterbuch, den „Příruční slovník jazyka českého“ (PSJČ), der von 1935–1957 erschienen ist, verschmäht sie, weil das Wörterbuch veraltet sei, und kritisiert auch den SSJČ aus ähnlichen Gründen (und verweist vor allem auf „kommunistische“ Lexik). Sie ergänzt dieses Wörterbuch durch andere neuere Wörterbücher, u. a. den „Akademický slovník současné češtiny“ (ASSČ), der aber bisher bisher nur bis zum Buchstaben Č erschienen. Weitere kleinere Wörterbücher werden auch herangezogen, spielen aber keine so wichtige Rolle. – Ich habe längere darüber nachgedacht, ob die Verfasserin hier wirklich richtig entschieden hat. Das ASSČ ist ein Torso, mit dem man fast nicht arbeiten kann, während das PSJČ nicht nur vollständig, sondern auch elektronisch durchsuchbar¹ ist. Und auch das Veralten der Lexik würde ich nicht so tragisch sehen. Natürlich gibt es solche Phänomene, und man muss sie berücksichtigen, aber gerade die kommunistische Lexik nimmt im SSJČ (und vermutlich auch in den späteren Bänden des PSJČ) keinen so großen Anteil ein, dass man das Wörterbuch gleich ganz ablehnen müsste. Auf der anderen Seite verstehe ich auch das Anliegen nach Berücksichtigung möglichst modernen sprachlichen Materials sehr gut, und das passt auch zur Entscheidung für das Korpus *csTenTen* (und zum weitgehenden Verzicht auf das Tschechische Nationalkorpus).

Von den weiteren Abschnitten des Unterkapitels 5.5 ist insbesondere der Unterabschnitt 5.5.1.4 wichtig, in dem es um die Beschreibung der evaluativen Bedeutung der Kollokatoren geht. Da es der Verfasserin offenkundig sehr wichtig ist, dass hier skalierbare Bedeutungen vorliegen, setzt sie zwischen *negativ* (abgekürzt NEG) und *positiv* (abgekürzt POS) noch drei weitere neutrale Bedeutungen, von denen eine zur negativen und eine zur positiven Bedeutung neigt, während die dritte ganz in der Mitte steht. Dafür, welche Bedeutung ein konkreter Kollokator hat, wird jeweils eine größere Zahl von Kriterien angegeben. Der folgende Unterabschnitt 5.5.1.5 behandelt die evaluative Bedeutung von Kolligationen, bei denen die Verfasserin 20 Typen unterscheidet, der Unterabschnitt 5.5.1.6 beschäftigt sich mit „häufigen Wortverbindungen und Phrasen“, d. h. um Einheiten, die auf dem Weg zur Phraseologisierung sind, und in Unterabschnitt 5.5.1.7 kündigt die Verfasserin schließlich an, dass sie auch die ontologischen Kategorien der Kollokatoren getrennt erheben und dokumentieren wird.

Dieses monumentale Begriffsgebäude wird sozusagen gekrönt von einem ganz kurzen Abschnitt (5.5.2), in dem die einzelnen Schritte der Analyse aufgezählt werden. Und im letzten Abschnitt des fünften Kapitels (5.5.3) folgt eine kurze Diskussion der Methodologie, in der als die zwei Schwerpunkte der Untersuchung erstens ein „Crowdsourcing“ aus Wörterbüchern und zweitens ein ganzes Spektrum von Kriterien genannt werden. Etwas überraschend finde ich, dass hier auch ein Experiment zur Reliabilität der Methode erwähnt, bei dem eine „unabhängige Annotatorin“ jeweils zehn zufällige ausgewählte Kollokatoren jedes Modifikators bewertet hat. Über dieses Experiment hätte meines Erachtens mehr gesagt werden müssen.

Das **sechste** Kapitel der Arbeit, das mit 166 Seiten mehr als 40% der Dissertation einnimmt, beschäftigt sich mit der Analyse von elf ausgewählten ZIP anhand der in den vorhergehenden Kapiteln eingeführten Methode. Ich werde auf diese Detailanalysen nicht genauer eingehen (aber ich habe alles gewissenhaft gelesen!) und beschränke mich auf die Feststellung, dass die Verfasserin die Methode konsequent anwendet und auch nicht während der Analyse verändert (was ja in wissenschaftlichen Arbeiten oft genug vorkommt). Das Einzige, was ich hier nicht völlig ideal finde, ist die alphabetische Reihenfolge der ZIP. Mir ist zwar klar, was die Verfasserin hiermit bezweckt, sie will sicher zeigen, dass sie sozusagen unvoreingenommen an alle Fälle herangegangen ist, aber für die Leserin oder den Leser wäre es vielleicht doch praktischer, wenn ähnliche bzw. verwandte ZIP zusammen behandelt würden. Einen Vergleich nimmt die Verfasserin ja durchaus vor (in Abschnitt

1 Unter der URL <https://psjc.ujc.cas.cz/>.

6.12), und wenn schon vorher in der Arbeit deutlich wurde, dass die mit *ně*- gebildeten ZIP gewissermaßen neutral bzw. unmarkiert sind, hätte man vielleicht mit diesen anfangen können. Am grundsätzlichen Wert der Arbeit ändert das gar nichts, aber wenn man eine Interesse daran, dass die eigenen Ergebnisse zur Kenntnis genommen werden, sollte man sich auch ein wenig um Lesbarkeit bemühen.

In Abschnitt 6.12 werden Unterschiede und Gemeinsamkeiten der 11 behandelten ZIP genau und anhand von vielen Tabellen beschrieben und (auf S. 369) auch visualisiert. Hier zeigt die Verfasserin noch einmal ihren bewundernswerten Kenntnisse quantitativer Methoden. Relativ kurz und knapp ist Unterabschnitt 6.12.3, in dem die Forschungsfragen beantwortet und die Hypothesen bestätigt bzw. abgelehnt werden. Hier ergibt sich ein sehr gemischtes Bild, bei dem deutlich wird, dass die sprachliche Realität deutlich komplexer ist als zu Anfang. Aber das ist letztlich nicht so erstaunlich und zeigt auch, dass die Verfasserin ihre Methode wirklich objektiv angewandt hat.

Über folgende Fragen möchte ich bei der Verteidigung der Dissertation gerne diskutieren:

- Wie lässt sich die Einschränkung der Untersuchung auf das Korpus *csTenTen* begründen?
- Wäre die Einbeziehung weiterer, älterer Wörterbücher (die aber ganz vorliegen) wirklich so problematisch?

Ich komme nun zur Zusammenfassung. Wie schon zu Anfang gesagt, liegt hier eine hervorragende Arbeit vor, die auf hohem theoretischen Niveau arbeitet und eine Vielzahl von empirischen Daten analysiert. Die Verfasserin verfügt über sehr gute Kenntnisse der bisherigen Forschung und von quantitativen Methoden. Die Vielzahl der Methoden und Daten ist aber manchmal auch ein bisschen problematisch, wenn beispielsweise Kapitel zu sehr verschachtelt sind (das betrifft vor allem das fünfte Kapitel) oder wenn einzelne Punkte zu kurz abgehandelt werden (beispielsweise in Abschnitt 5.5.3).

Für den tschechischen Teil des Verfahrens stupe ich die eingereichte Dissertation vorläufig als *bestanden* ein. Für den deutschen Teil des Verfahrens bewerte ich sie mit

magna cum laude (1,0).

Prof. Dr. Tilman Berger



Universität Tübingen · Wilhelmstr. 50 · 72074 Tübingen

Referát doktorského studia
Filozofická fakulta Univerzity Karlovy
nám. Jana Palacha 2
116 38 Praha 1

Slavisches Seminar

Prof. Dr. Tilman Berger
Slavische Sprachwissenschaft

Slavisches Seminar
Wilhelmstraße 50
72074 Tübingen
Telefon +49 7071 29-76733
Telefax +49 7071 29-5924
tberger@uni-tuebingen.de
www.slavistik.uni-tuebingen.de

Tübingen, November 14th 2023

**Review of the dissertation by Mgr. Martina Rybová on the topic
„Die Beziehung zwischen Evaluativität und Indefinitheit am Beispiel der zusammengesetzten
Indefinitpronomina im Tschechischen“**

This dissertation deals with a topic that could be considered unusual and strange from the perspective of traditional structuralist linguistics. Traditionally, indefinite pronouns are described in grammar as units that are characterised by the category of indeterminacy, so one does not initially think of evaluativity, which one would rather classify in the field of pragmatics. But the fact that it is possible to write a dissertation on such a topic shows that our knowledge of indefinite pronouns has developed considerably in recent years. It is therefore no coincidence that the dissertation was written in the context of the DFG project headed by Björn Hansen entitled “Compound indefinite pronouns in Slavic languages. A contribution to the semantic map of second-generation indefinites”. It was precisely this project that focussed attention on the diverse compound indefinite pronouns in various Slavic languages. And this dissertation focuses on Czech indefinite pronouns, which are analysed from various points of view, above all from the point of view of evaluativity.

At this point I would like to say that this is an excellent work in which the author has worked at a high theoretical level and at the same time comprehensively analysed a large amount of empirical data. It significantly expands our knowledge of the meaning and function of Czech indefinite pronouns and also provides a number of suggestions for further research.

In the following, I will go into the six chapters of the thesis in more detail, comment on individual points, formulate questions for the defence of the thesis and propose an evaluation of the thesis at the end of my report.

In the **first** chapter, “Introduction and Objectives”, the author briefly introduces the topic, makes reference to the DFG project and outlines how she will proceed. The second and third chapters deal with the phenomenon of evaluativity and the category of indefiniteness and Czech indefinite pronouns, the fourth chapter formulates research questions and hypotheses, the fifth chapter deals with the methods used in the work and the sixth chapter presents the concrete analysis of eleven compound indefinite pronouns.

As already mentioned, the **second** chapter deals with the phenomenon of evaluativity. The author gives a broad overview of the literature here, beginning with Bally (1913), but then referring mainly to a number of more recent works. I must confess that I was only familiar with a small number of

these works, and I find the way in which the author describes the various concepts with reference to six main aspects (cf. the table on p. 62f.) very impressive. In section 2.2, the author then briefly comments on various evaluative means of expression; I was a little surprised that she does this using German and not Czech examples. Finally, in section 2.3, she gives her own definition of evaluativity and positions herself in relation to the older research she refers to. — Overall, this chapter is very informative and well done. I was a little surprised, however, that Kateřina Veselovská's work on the Czech language is completely integrated into the general context and is not considered separately, even though (unlike most of the other studies cited) it is cited more frequently in the rest of the work.

Unfortunately, this chapter also contains the only point that I really have to criticise this work. In a footnote on p. 27, the reader is informed very briefly and *en passant*, so to speak, that the present work is predominantly based on data from the web corpus “csTenTen v 17”. Although I consider this decision to be entirely correct (!), I believe that the author should have justified it more precisely at the beginning of the work and, above all, provided more information about competing corpora. Corresponding information is provided in subchapter 5.3 “Corpora used” (pp. 173-176), roughly in the middle of the work. And even there, the corpora are essentially described and the author's own considerations are cited, but there is no reference at all to the not entirely small literature on web corpora.

The **third** chapter deals with the concept of indefiniteness and the state of research on Czech indefinite pronouns. The first four sub-chapters discuss the concept of indefiniteness, its role in indefinite pronouns and its relationship to evaluativity. These sub-chapters are relatively short, but if I can see it correctly, all the essential literature on the topic is discussed. What I missed, however, was a discussion of the distinction between the indefinite article and indefinite pronouns. As we know, there is no indefinite article in Czech, but in translations from article languages it is sometimes rendered by indefinite pronouns (especially *nějaký*). When talking about the fact that “most of the contributions to this discussion” have dealt with “nominal phrases with an indefinite article rather than with the reference of other indefinite expressions” (p. 70), a few brief remarks on this problem would be appropriate.

The fifth subchapter presents the state of research on Czech indefinite pronouns, which are not as well researched as would be desirable. What is remarkable here is the great importance attached to dictionaries. One would not expect this if one considered indefinite pronouns to be a grammatical phenomenon, but with regard to evaluativity and, above all, the research method proposed by the author in the fifth chapter, this is a sensible step.

The sixth and seventh subchapters deal with the identification and inventory of new potential indefinite pronouns in Czech. This is a very interesting and innovative topic in which this work is truly remarkable. The author discusses in detail the various factors that need to be taken into account in the identification process, in particular questions of hyphenation and contraction, but also punctuation. All observations are underpinned by corpus data, which are very carefully collected. The author always gives the search formulae and explains them where necessary.

Overall, this chapter, which concludes in the eighth subchapter with considerations on the selection of indefinite pronouns to be analysed in more detail in the dissertation, is very well done. Nevertheless, I would like to point out that the restriction to the corpus *csTenTen* is not overly plausible and should have been justified. Instead, it is stated laconically on p. 100 that the search should be carried out in this corpus, which was mentioned for the last time in the footnote on p. 27.

The **fourth**, relatively short chapter contains the research questions and hypotheses to be investigated in the further course of the dissertation. The questions are derived from the two previous

chapters. The first question is whether compound indefinite pronouns (hereafter ZIP) occur in evaluative environments, and then the interaction of ZIP with the evaluative units, the comparison of different indefinite pronouns and typical patterns that can be observed here. The hypotheses are also obvious against the background of the questions, even if it is not explained exactly how the author arrived at them. But perhaps this is not usual for hypotheses.

The **fifth**, again somewhat longer chapter deals with the methods to be used in the work. Here the author makes a very original and innovative suggestion, which she later describes as a kind of “crowdsourcing” (p. 214). She combines a corpus analysis with a sophisticated use of dictionary data. Specifically, the evaluativeness of the units that occur in connection with the ZIPs is assessed by means of a sophisticated analysis of dictionary data, but many other aspects are also taken into account.

But now to the individual steps: A brief overview of the literature in subsection 5.1 is followed by a description of collocation analysis (section 5.2). This begins with a brief review of the literature, but the introduction and definition of further terms such as ‘colligation’ and semantic and evaluative prosody are essential. As elsewhere in the work, the terms and procedures are also analysed using corpus examples, although I was a little surprised that Polish examples were suddenly used. Not that this really bothers me, it emphasises the Slavic character of the work and the author's qualifications, but it is a little strange that Polish examples are used on pages 170 and 171 and then only Czech ones again. Subchapter 5.3 contains the information on the corpora used, which, as I have already said twice, I believe should have been mentioned well in advance. Here, finally, there is more precise information on the *csTenTen* corpus and the problems associated with it – I was fascinated by the report on the “casino” texts, which, frankly, were previously unknown to me. I find it a little problematic that only less than half a page is devoted to the rejection of the Czech national corpus (cf. p. 176), but I can ultimately accept the author's decision to proceed in this way. In section 5.4, the author makes the decision to only consider the first 500 collocations in each case; this is justified with statistical arguments that I can well understand.

Subchapter 5.5 presents the use of dictionaries. I find the presentation here somewhat confusing because the section in which the author explains which dictionaries she has selected is very deeply “embedded” (it is section 5.5.1.2.1!). It would really have been better to devote a separate “higher” section to the dictionaries. I myself would like to take a closer look at the selection here, as I find it problematic in places. First of all, it is easy to understand why the author mentions the “Czech Subjectivity Lexicon” in first place, as this dictionary certainly fulfils the requirements best, but it is far from sufficient because it does not contain so many lemmas. In second place, the author then opted for the “Slovník spisovného jazyka českého” (SSJČ), the second major dictionary of Czech, which was compiled in the 1960s. She does not want to use the first large dictionary, the “Příruční slovník jazyka českého” (PSJČ), which was published from 1935-1957, because the dictionary is outdated, and criticises the SSJČ for similar reasons (and refers above all to “communist” lexis). She supplements this dictionary with other newer dictionaries, including the “Akademický slovník současné češtiny” (ASSČ), which has so far only been published up to the letter Č. Other smaller dictionaries are also consulted, but do not play such an important role. — I spent a long time thinking about whether the author really made the right decision here. The ASSČ is a torso that is almost impossible to work with, whereas the PSJČ is not only complete, but also electronically searchable¹. And I wouldn't see the obsolescence of the lexicon as such a tragedy either. Of course there are such phenomena, and you have to take them into account, but the communist lexis in particular does not take up such a large proportion in the SSJČ (and presumably also in the later volumes of the PSJČ) that you would have to reject the dictionary outright. On the other hand, I also understand very well the desire to include as much modern linguistic material as possible, and this also fits in

1 Under the URL <https://psjc.ujc.cas.cz/>.

with the decision in favour of the *csTenTen* corpus (and the decision to largely dispense with the Czech National Corpus).

Of the other sections of subchapter 5.5, subsection 5.5.1.4, which deals with the description of the evaluative meaning of the collocators, is particularly important. Since it is obviously very important to the author that there are scalable meanings here, she places three further neutral meanings between *negative* (abbreviated NEG) and *positive* (abbreviated POS), one of which tends towards the negative and one towards the positive meaning, while the third is placed right in the middle. A larger number of criteria are given for the meaning of a specific collocator. The following subsection 5.5.1.5 deals with the evaluative meaning of collocations, in which the author distinguishes 20 types, subsection 5.5.1.6 deals with “frequent word combinations and phrases”, i.e. units that are on the way to phraseologisation, and in subsection 5.5.1.7 the author finally announces that she will also collect and document the ontological categories of collocators separately.

This monumental conceptual structure is crowned, so to speak, by a very short section (5.5.2) in which the individual steps of the analysis are enumerated. And the final section of the fifth chapter (5.5.3) contains a brief discussion of the methodology, in which the two focal points of the study are firstly “crowdsourcing” from dictionaries and secondly a whole range of criteria. I find it somewhat surprising that an experiment on the reliability of the method is also mentioned here, in which an “independent annotator” evaluated ten randomly selected collocators of each modifier. In my opinion, more should have been said about this experiment.

The **sixth** chapter of the thesis, which at 166 pages takes up more than 40% of the dissertation, deals with the analysis of eleven selected ZIPs using the method introduced in the previous chapters. I will not go into these detailed analyses in detail (but I have read everything conscientiously!) and will limit myself to stating that the author applies the method consistently and does not change it during the analysis (which happens often enough in scientific works). The only thing I find not entirely ideal here is the alphabetical order of the ZIP. It is clear to me what the author is trying to achieve with this, she certainly wants to show that she has approached all the cases with an open mind, so to speak, but it would perhaps be more practical for the reader if similar or related ZIPs were dealt with together. The author does make a comparison (in section 6.12), and if it had already become clear earlier in the work that the ZIPs formed with *ně-* are to a certain extent neutral or unlabelled, it might have been possible to start with these. This does not change the fundamental value of the work, but if you are interested in having your own results recognised, you should also make a little effort to ensure readability.

In section 6.12, the differences and similarities between the 11 ZIPs discussed are described in detail and visualised (on p. 369) using numerous tables. Here the author once again demonstrates her admirable knowledge of quantitative methods. Subsection 6.12.3, in which the research questions are answered and the hypotheses confirmed or rejected, is relatively short and concise. A very mixed picture emerges here, in which it becomes clear that the linguistic reality is significantly more complex than at the beginning. But this is ultimately not so surprising and also shows that the author has applied her method in a truly objective manner.

I would like to discuss the following questions during the defence of my dissertation:

- How can the restriction of the study to the *csTenTen* corpus be justified?
- Would the inclusion of other, older dictionaries (which are available in their entirety) really be so problematic?

I now come to the summary. As I said at the beginning, this is an excellent work that works at a high theoretical level and analyses a large amount of empirical data. The author has a very good know-

ledge of previous research and of quantitative methods. However, the large number of methods and data is sometimes a little problematic, for example when chapters are too convoluted (this applies particularly to the fifth chapter) or when individual points are dealt with too briefly (for example in section 5.5.3).

For the Czech part of the procedure, I provisionally classify the submitted dissertation as *passed*. For the German part of the procedure, I assess it as

magna cum laude (1,0).

Prof. Dr. Tilman Berger