

## **Workshop: Linguistics and Translation Studies in Dialogue** **Astrid Schmidhofer, Rossella Resi (University of Innsbruck)**

---

Despite translation studies being a relatively new field compared to the broader discipline of linguistics, and its occasional classification as a subset of applied linguistics due to its practical attributes, both disciplines have mutually influenced each other over the years. These interdisciplinary contributions have been enriching both fields, fostering a deeper understanding of language and communication in a globalized world. An extensive overview of mutual interaction between the two disciplines over the years can be found in Kuhiwczak & Littau (2007).

The contribution of linguistics to translation studies is multifaceted. Linguistics provides the metalanguage required to describe the most fundamental tool of translation, which is language. Linguistics crucially shaped early theories of translation (Siever, 2015), which continue to exert a certain influence even today. Discourse analysis within translation studies benefits from linguistic insights, enhancing our understanding of how meaning is constructed and conveyed across different cultural and social contexts (Munday & Zhang, 2017). Conversely, contrastive studies, fueled by translation practice, allow linguists to compare languages to identify similarities and differences at various levels (since Hatim, 1997). Such studies are also being supported by corpus linguistics, which provides the theory for the compilation and analysis of corpora, which are frequently used for contrastive analysis and analysis of translator behavior. This comparative work deepens our knowledge of language structures and contributes to linguistic theories.

A rather recent evolution, which connects the two disciplines, is the evolving synergy between computational linguistics and translation studies, which can also present challenges. Notably, advancements in computational linguistics may overshadow the unique human contributions that are intrinsic to the languages processes, including translation. However, while the first generations of Machine Translation relied heavily on linguistic rules, LLM and NMT approaches have transcended traditional linguistic structures (Sánchez-Gijón et al., 2015), creating a divide between computer science on one side and linguistics and translation on the other side. Given this evolution, which seem to disprove the limitation of machine in producing language (but not of encapsulating it into rules), researchers in translations and linguistics may need to collaborate or engage in dialogue more than ever.

In addition to these challenges brought on by recent developments, translation studies and linguistics continue to nourish their respective fields of study with interesting approaches.

### **Terminology at the interface**

#### **Rossella Resi (University of Innsbruck)**

---

While Terminology might be considered an independent science, it also entails a rather unique level of interaction with other intellectual disciplines. It relies heavily on dialogue with experts in the countless domains in which terminology is defined and studied. It also draws inspiration from, and serves, a variety of adjacent academic fields with which terminologists share common competences, for example linguistics and translation studies (RaDT 2020, TerminOrgs 2016 Drewer et al. 2018).

These disciplines interact in diverse ways, both synchronically and diachronically. Terminology is studied and applied in various ways across different environments, reflecting national perspectives, academic settings, and levels of engagement in the private sector (overview in Korkas und Rogers 2010, RaDT 2020). The development of these interrelationships over time reveals how terminological science has both served and benefited from recent technological advancements (Prieto 2024).

This work aims to provide an overview of the role of science based Terminology across this interface, exploring how it interacts with translation, linguistics, and associated technological advancements that are destined to evolve further. Terminology was traditionally relegated to a pre-process in computer-assisted translation, but it is now increasingly involved in downstream processes, such as Terminology Enforcement, and Retrieval-Augmented Generation (RAG) (Resi and Lang, in press). There is an emerging trend for the use of terminological metadata and this has the potential to promote

terminological awareness and elevate the perceived value of the field well beyond its historical roots. Examples are presented of these developments over time, highlighting the past and future evolution of Terminology and its practitioners, particularly in light of recent changes within the language industry.

### **How lexicography, terminology studies and specialised translation can benefit from the use of large-language models**

**Barbara Heinisch (Centre for Translation Studies, University of Vienna, Austria)**

---

The advent of large-language models (LLMs) such as GPT-4 has opened new avenues for research and practice in lexicography, terminology studies and specialized translation. Terminology is the critical nexus that connects these three fields, serving as the backbone of precision and clarity in professional (multilingual) communication. This contribution explores how LLMs can alter the management and application of terminology across these interconnected fields while simultaneously addressing sustainability issues.

In lexicography, LLMs can facilitate the creation and maintenance of specialized dictionaries. These models can automatically extract and define lemmata from large corpora, identify usage patterns and provide contextual examples, thus streamlining the lexicographic process. By analyzing real-world language use, LLMs help ensure that dictionary entries are accurate, up-to-date and reflective of contemporary discourse within specific domains. The contribution will also explore how text-generative LLMs might reshape corpus design and challenge traditional notions of what constitutes 'authentic text' as more and more content will be produced by means of text-generative LLMs.

In the realm of specialized translation, the consistent and accurate use of terminology is essential. Depending on the language pair, LLMs can support translators by providing context-aware terms that align with established terminological standards. They can also help translators navigate complex subject matter by offering suggestions for equivalents and usage examples across languages, thereby improving both the efficiency and accuracy of the translation process.

Terminology studies, which focus on the systematic collection, analysis and dissemination of terms, can benefit from LLMs' ability to handle large datasets and identify patterns across languages and domains, including providing definitions, finding variants or visualising entire concept systems. These models can also assist in building and updating (multilingual) terminological databases or analyze terminological consistency within and across documents to allow for terminology control.

Generally, by leveraging LLMs, professionals in lexicography specialized translation and terminology studies can enhance the quality and consistency of their work.

### **Das literarische Übersetzen als Brückenschlag zwischen Linguistik und Literaturwissenschaft sowie Theorie und Praxis**

**Katrin B. Karl (Universität Bern, Institut für Slavische Sprachen und Literaturen)**

---

Im Sommer 2023 fand an der Universität Bern ein russisch-deutsches Übersetzungsprojekt, durchgeführt von einem Literaturwissenschaftler und einer Sprachwissenschaftlerin statt. Dieses bestand aus insgesamt drei Komponenten und setzte sich zum Ziel, einen mehrfachen Transfer herzustellen. In einer im Co-Teaching durchgeführten Vorlesung wurden die Studierenden mit den wichtigsten Konzepten zur Übersetzungstheorie aus beiden Disziplinen vertraut gemacht. Im angeschlossenen praktisch orientierten Seminar standen literarische, bislang nicht ins Deutsche übersetzte Texte russischsprachiger Gegenwartsautor:innen im Mittelpunkt. Diese wurden zunächst gelesen, literaturwissenschaftlich eingeordnet und dann in Gruppenarbeit vom Russischen ins Deutsche übersetzt. Der Fortgang der Übersetzungen wurde im Seminar besprochen, wodurch die Sitzungen einen Werkstattcharakter bekamen. In ihnen standen konkrete Übersetzungsprobleme im Mittelpunkt, die mit Hilfe linguistischer Ansätze eingeordnet wurden. Die dritte Komponente stellte eine organisierte Lesereihe dar, zu der die jeweiligen Autor:innen eingeladen und die Texte im Original und der bereits erarbeiteten Übersetzung einem Publikum präsentiert wurden.

Das Ziel des Projekts war es zu demonstrieren, in welcher Weise das literarische Übersetzen dazu

beitragen kann, die beiden häufig als separat angesehenen Disziplinen der Literatur- und Sprachwissenschaft so miteinander zu kombinieren, dass aus ihnen ein erfahrener Mehrwert in der Auseinandersetzung mit einem Text entstehen kann. Daneben sollte durch die eigene Übersetzung und die Outreach-Komponente Studierenden der Slavistik vor Augen geführt werden, welcher praktische Nutzen aus ihrem Tun entstehen kann.

Die im Lauf des Projekts gesammelten Erfahrungen und entstandenen Übersetzungen werden aktuell in eine Buch-Publikation überführt. In dieser werden neben Informationen zu den Autor:innen Einblicke in die gemeinsame Übersetzungsarbeit und die darin erfahrene Verzahnung von Literatur- und Sprachwissenschaft gegeben sowie die Texte in zweisprachiger Form publiziert.

Der Vortrag stellt das Projekt in seinen unterschiedlichen Zielsetzungen und Verlauf dar und illustriert anhand von konkreten Sprachbeispielen die komplexe Verzahnung beider Disziplinen bei der Erschließung der übersetzten Texte.

### **The translation of compounds and the status of collocations**

**Lisa Marie Lang (Universität Innsbruck), Veronika Hudáková (Pavol Jozef Šafárik University in Košice)**

---

German is a language with many nominal compounds. In most other languages, various types of collocation are used to render them in translation. In Slovak, nominal compounds such as *zemeguľa* ('earth sphere') and *rýchlovlak* ('fast train') occur, but in most cases the translation of German nominal compounds requires collocations consisting of nouns with relational adjectives or genitive nouns. Some examples of German compounds and their Slovak equivalents are given in (1).

- |     |    |                |                |                    |
|-----|----|----------------|----------------|--------------------|
| (1) | a. | Benzinpreis    | cena benzínu   | ('petrol price')   |
|     | b. | Literaturpreis | literárna cena | ('literary award') |

In (1a), a genitive noun renders the non-head of the compound, in (1b) a relational adjective does this. In the Slovak grammatical tradition, expressions such as the two in (1) are classified as collocations. We investigated for nominal compounds with German *Preis* ('price' or 'award') and its Slovak equivalent *cena* to what extent the compounds in German can be translated along the two patterns in (1) and to what extent Slovak collocations with genitive nouns and relational adjectives can be translated as German compounds. As a basis for the German compounds, we used the list of analysed compounds described in Henrich & Hinrichs (2011), which in its 17<sup>th</sup> edition includes 115,000 compounds. We extracted 284 compounds with *Preis* as the head and 132 compounds with *Preis* as the non-head. For Slovak collocations, we used Ďurčo & Majchráková (2017), who give 181 collocations with *cena*. In some cases, the interpretation of a compound is not immediately obvious. Thus, German *Negativpreis* can mean 'price you pay when selling something' or 'award for a particularly bad performance'. We analysed the meaning and the translation of the sets of German compounds and Slovak collocations in order to discover any generalizations to predict which Slovak collocations should be translated as compounds in German and which type of collocation should be used in Slovak for translating a German compound.

### **The presentation of word meaning through ostensive aids for translation and language learning**

**Laura Rebosio (University of Innsbruck)**

---

In lexical resources for translation, pictures, illustrations (drawings), schemes, terminological diagrams, frames, videos and other visual aids are of prime importance and in several cases, they are the sole really useful tool for finding the proper correspondence in the target language: e.g., the German translation for IT *fiore* or FR *fleur* can be both *Blume* and *Blüte*, but these have a different semantic distribution. In a technical translation (botanical books) into German the terms must be used distinctively. Despite the usefulness of illustrations during the translation process, a translator can currently count only on some paperback, multilingual visual dictionaries but not on any online resource which is conceived and built in a systematic, ontologically coherent way for this purpose (BabelNet 5.3 2024 or The Visual 2024 being two exceptions but displaying some problems).

There are cases in which “a picture tells a 100 words” and others in which linguistic descriptions are most appropriate for grasping the gist of a word or of a sentence. For instance, depicting spices does not tell us much about their actual flavor, and the difference between *sweet paprika* and *harissa* is better described than depicted (the Treccani 2003 visual dictionary, *Il Visuale*, however, seems to ignore the poor usefulness of pictures for these and other cases).

On the other hand, describing the parts of a flower or of a bicycle, and describing particularly “visual-friendly” actions and activities like *nudging someone*, *chopping onions* or *chopping wood* is probably not the most effective way to convey their meaning – especially, if we consider a translator or a learner as the ideal user of a dictionary. In these cases, ostensive aids are a more valuable help.

In my contribution, I would like to give an overview of the current typological aids available for different word classes in current resources, analyse theoretical aspects regarding the choice of “the right picture”, discuss the difficulties of dealing with polysemy and the benefits/drawbacks of prototypical representations.

This study is part of my PhD project, which is being carried out within the framework of *PhraseBase*, which has the *Phrase-based Active Dictionary* (PAD) (DiMuccio-Failla/Giacomini 2017a; 2017b; 2022) as one of its components. In the PAD, lexical units are treated phraseologically, i.e. they consist of the lemma and its valency structure (semantically specified, cf. Hanks 2013). Such lexical units allow for less semantic ambiguity, giving polysemy some boundaries. Still, presenting these lexical units visually and in a cognitively-oriented way remains a challenge to be addressed (cf. Giacomini/DiMuccio-Failla/De Martin Pinter 2022 for a first approach in this direction).

### **Translation as a tool in the cross-linguistic investigation of compounding**

**Pius ten Hacken (Universität Innsbruck)**

---

Compounding is one of a range of concepts in linguistics for which it is hard to find a good definition. Linguistic data do not come with their classification attached to them. There is a certain degree of agreement among linguists as to which expressions belong to compounding, but in borderline cases, linguists will have different, sometimes conflicting intuitions as to whether an expression is a compound.

There are at least two reasons why it is interesting to approach such situations from a cross-linguistic perspective. First, when compounding occurs in several languages, it can be proposed as a potentially universal phenomenon. Secondly, when data from several languages can be used to delimit compounding, we can gain a more precise view of its nature and its position in language as a whole.

For a cross-linguistic study of compounding, translation is an essential tool. When we aim for a definition with a validity that transcends the boundaries of a single language system, we have to make sure that corresponding expressions in different languages are classified in the same way. Without translation, it is impossible to determine the correspondence between expressions.

Compounds are words. Translating words in isolation is problematic. Instead, in translation studies, we rather work with texts. In this way, the intended sense of a lexical expression is generally retrievable from the context in the source text and we can use the context to redistribute the information among the words in the target text. When we translate compounds in isolation, we cannot do this.

In my paper, I will show how it is possible to use translation in a theoretically and practically acceptable way in the research of compounding. The approach is based on a careful observation of linguistic and translational properties of expressions and can be extended to other theoretical phenomena.

### **Binominal constructions in Slavic and Romance languages: a corpus analysis of original texts and translations**

**Inga Hennecke (Tübingen), Christina Clasmeier (Münster)**

---

Binominal constructions are complex lexemes creating (new) complex nominal concepts by combining two nominal concepts (Pepper et al. 2023: 7), such as Polish *książka telefoniczna* [book telephone-ADJZ] ‘telephone book’ (NA = noun + relational adjective) or French *toile d'araignée* [web of

spider] ‘spider web’ (N prep N = noun + preposition + noun). In our talk we critically discuss the concept of binominals and their role in Contrastive linguistics and Translation studies. We present a corpus-based analysis of original texts and translations in Slavic and Romance languages, which share basic types of binominal constructions. However, considerable differences in frequency and distribution of binominal constructions exist within the Slavic languages as well as between Slavic and Romance languages.

In our project we investigate the influence of the source language (Polish) on the translation of binominal constructions in closely related and more distant target languages.

Our analysis contains three steps. First, we locate typical binominal constructions in Polish (based on Clasmeier & Hennecke 2024) and analyze their frequency and distribution in a belletristic corpus. Second, we carry out a comparative analysis of parallel belletristic corpora of one closely related Slavic language (Czech) and one more distant Romance language (French). Third, we analyze Polish binominal constructions and their Czech and French translation equivalents in a parallel translation corpus.

We hypothesize that (potential) differences in the frequency and distribution of binominal constructions between the original texts and the translations are due to the translation process. Furthermore, we expect the type of binominal constructions of the source language to show a stronger influence on translation equivalents in the closely related language than in the more distant language.

We argue that insights from Contrastive linguistics are useful and necessary for an appropriate interpretation of particular linguistic structures in Translation studies.

### **Constraints on the use of bare/*des* plural subjects in Romance: how (digital) translations can inform linguistics**

**Petra Sleeman (University of Amsterdam)**

In Germanic languages indefinite plural subject NPs can be used bare, i.e. without an overt determiner, in the function of topic or focus. Germanic languages differ from Romance languages, for which it has been claimed that the bare subject is only allowed if it is stressed or if it is modified, which would be a special type of focalization, as expressed by Suñer’s (1982) Naked Noun Constraint (see also Salem 2010; Leonetti 2013):

- (1) \**Turistas* / *TURISTAS* / *Turistas curiosos llegaron a la ciudad.* (Spanish)  
tourists / TOURISTS / tourists curious arrived in the city  
‘Tourists / TOURISTS / Curious tourists arrived in the city.’

French does not have bare nouns, but uses indefinite plural *des*-NPs for foci.

Both for bare subject nouns and for *des*-subjects it has been observed in the literature that they occur essentially in written texts (Englebert 1992; Oosterhof 2008; Leonetti 2013). I selected all the bare plural subjects in a Dutch novel. I investigated their translation into Spanish, Italian, Portuguese and French by the official human translators and by three digital translators (Google Translate, DeepL, ChatGPT).

The results for all of the four translators (human and digital) confirm that focalization largely enhances the use of indefinite noun phrases, whereas for the topics definite articles are used. The presence of a modifier substantially increases the use of bare or indefinite subjects in Romance by the translators, especially in Spanish, Portuguese and Italian, confirming the part on the presence of a modifier in Suñer’s Naked Noun Constraint. However, the global results also show that in non-modificational contexts bare or indefinite subjects are not totally excluded. A closer inspection of the contexts in which bare/*des*-subjects are used shows that they are broad focus contexts instead of narrow focus contexts. This goes against the Naked Noun Constraint

Although this is the general picture that emerges from the translations, there are some differences between languages or translators.

## **Erfolge und Misserfolge von MT und GenAI bei der Übersetzung des Modalverbs *sollen* ins Spanische**

**Astrid Schmidhofer (Universität Innsbruck)**

---

Nach der Präsentation der Ergebnisse einer Korpusstudie zur Übersetzung des deutschen Modalverbs *sollen* ins Spanische anhand einer Auswahl aus dem Korpus PaGeS (corpuspages.eu) wurde diese nun um die Performance von GenAI und MT bei der Übersetzung dieses Modalverbs erweitert. Dafür wurde ein Subset mit 90 zufällig ausgewählten Items aus den ursprünglichen 543 herangezogen, wobei auf eine möglichst gleichmäßige Verteilung über die Kategorien des Modalverbs *sollen* nach Milan (2001) geachtet wurde.

Die ausgewählten Items wurden dann von zwei MT-Systemen (Google Translate und DeepL) sowie zwei GenAI-Systemen (Gemini und ChatGPT) ins Spanische übersetzt. Diese Übersetzungen wurden auf einer Likert-Skala von 1 bis 4 bewertet, deren Werte vorher angelehnt an Schmidhofer (2024) beschrieben wurden.

Die erhaltenen Ergebnisse wurden statistisch auf Unterschiede zwischen den unterschiedlichen Systemen sowie den unterschiedlichen Kategorien getestet. Ebenso wurde die Häufigkeit des spanischen Modalverbs *deber* in den Übersetzungen analysiert und mit der Bewertung der betroffenen Items verglichen. Dieses Verb wird häufig als Entsprechung von *sollen* erachtet, was jedoch nur bedingt richtig ist.