

Riccardo Pozzo*, Timon Gatta, Hansmichael Hohenegger,
Jonas Kuhn, Axel Pichler, Marco Turchi, and Josef van Genabith

Aligning Immanuel Kant's Work and its Translations

Abstract: This chapter discusses using CLARIN to edit Kant's work and to consider how to align it with its translations, with special attention to Chinese. *Kangde* 康德 is the two-character phonetic loan that renders Kant's name in Chinese. We have chosen *Kangde* 康德 as the name for our vision to express the challenge of setting up the new edition of the *Druckschriften* and their Chinese translation in the form of aligned corpora, thus opening up the way to further alignments with versions in other languages. From a philosophical-historical and cultural-political perspective, the chapter presents the idea of aligning two parallel corpora of around 1,580,000 German words and the corresponding characters in Chinese. The project is curiosity-driven and lays the foundations for investigating Kant's philosophy and discussing it in a global context, a long-term effort that relies on the synergies among philosophy, computational linguistics, machine learning, translation studies, and China studies. The idea of the alignment is to offer unrivalled material for historical-philosophical investigations and serve as a viable infrastructure to be scaled up to other languages. To date, few aligned corpora exist that connect German and Chinese philosophical texts. The tools are not statistically implemented. As suggested by Franco Moretti's notion of distant reading, experimentation on meaningful patterns in philosophical corpora is a step towards making new machine learning technologies usable for tackling issues in the humanities. Looking forward, we focus on the assumption that philosophers ought to explore new technologies to rethink conventional ways of interpreting texts in the humanities.

*Corresponding author: **Riccardo Pozzo**, Department of History, Humanities and Society,
Tor Vergata University of Rome, Rome, Italy, e-mail: riccardo.pozzo@uniroma2.it

Timon Gatta, DSPFS, Tor Vergata University of Rome, Rome, Italy, e-mail: timon.gatta@uniroma2.it

Hansmichael Hohenegger, Italian Institute of Germanic Studies, Rome, Italy,
e-mail: hohenegger@studigermanici.it

Jonas Kuhn, IMS, University of Stuttgart, Stuttgart, Germany,
e-mail: jonas.kuhn@ims.uni-stuttgart.de

Axel Pichler, IMS, University of Stuttgart, Stuttgart, Germany,
e-mail: axel.pichler@ts.uni-stuttgart.de

Marco Turchi, ICT, Bruno Kessler Foundation, Povo-Trento, Italy, e-mail: turchi@fbk.eu

Josef van Genabith, MLT, German Research Center for Artificial Intelligence, Saarbrücken,
Germany, e-mail: josef.van_genabith@dfki.de

Keywords: China studies, contemporary Chinese, corpus linguistics, critical editions, digital humanities, history of philosophy, intercultural philosophy, Immanuel Kant, multilingual philosophical corpora, translation studies

1 Introduction

Let us start with a thought experiment. Imagine a first-generation diaspora youth (*huaqiao* 华侨) who studies philosophy at a European university. At a certain point, she might be expected to read Kant's *Grounding of the Metaphysics of Morals*, first translated into the language of the country she lives in – for Europe's official languages are as many as 24 (Schlüter and Hohenegger 2020) – then in the German original and the English rendering, say, of Mary Gregor. Let us assume that through the library of her university or one of the e-corpora, she finds access to the same text in Chinese, say, in the fourth volume of Li Qiuling's 李秋零 (2003–2019) translation. At this point, she might be able to start a discussion on Kant in her Chinese-speaking environment (Wen Haiming 2012). In turn, fellow students would appropriate the fourth-century BC philosopher of human nature Mengzi 孟子, through the references indicated by her. In the end, by referring Kant's German to texts in Chinese, English, and possibly other languages, our imaginary classroom might start thinking together on batches of multilingual concepts. Eventually, they would come to grasp some key tenets of global significance on the autonomy of human nature (Tu Weiming 2010). This is something philosophers today might want to take advantage of (Pozzo 2020a: 57).

This chapter is about a corpus construction project (see Hajičová et al. 2022). It is based on our experience when using the constellation of resources offered by CLARIN to edit Kant's work and consider how to align it with its translations, with special attention to Chinese. The results of the corpus construction are yet to come. However, thinking of the multilingual dialogues that are to take place in the coming years – first and foremost, the 25th World Congress of Philosophy of Rome in 2024 – what we wish to offer here is the unfolding of a vision, spelling out the single stages of a procedure to follow. The challenge is of setting up in the form of aligned corpora the new edition of the *Druckschriften* and their complete Chinese translation (Li Qiuling 2003–2019), thus opening the way to further alignments such as with the Cambridge Edition of the works of Immanuel Kant (Guyer and Wood 1992–2016), the Russian translations coordinated by the Institute of Philosophy of the Russian Academy of Sciences (Tuschling and Motroshilowa 1994–2020), and many other translation endeavours (Schlüter and Hohenegger 2020). However, because few aligned corpora connect German and Chinese, we remain focused on Kant in Chinese.

This chapter aims to reach out to social sciences and humanities (SSH) scholars who are not used to combining resources with metadata in order to analyse and enrich them with linguistic tools, as well as to scholars of non-SSH disciplines, more generally, those researchers who “are not just consumers of data and tools, but also providers” (Maegaard Bente, Van Uytvanck, and Krauwer 2017: 5–6) who are encouraged to share their data and tools with others, thus enhancing familiarity with approaches that allow the communities of CLARIN users to benefit from text corpora for philosophical research in multilingual and multicultural contexts (see Draxler et al. 2022). After describing the state of the art, the remainder of this chapter is about editing Kant's re-established polygraphy for systematic comparison of translations and analysing the evolution of contemporary Chinese philosophical terminology in relation to Kant's work.

2 State of the art

Information technology is revolutionizing how we approach texts and practice philosophical inquiry. The vision of *Kangde* is about transformative effects on methodologies in the history of philosophy. In this context, we argue that the time is ripe for a paradigm shift from thinking of texts to thinking of corpora, which is an issue that connects with hard, theoretical questions such as how to conceive of philosophical works within the infosphere (Blair et al. 2011; Floridi 2019; Romele 2020; Pozzo 2021). Philosophers have always been strenuous advocates of the close reading of texts and champions of the centrality of text. However, they have also been among the first to seize the opportunity to profit from the distant reading of corpora for the history of ideas, the history of scientific terminology, the translation of philosophical texts, and the translation of studies (Gregory et al. 1967). “Distant reading,” says Franco Moretti, “is a condition of knowledge,” for it allows one “to focus on units that are much smaller or much larger than the text: devices, themes, tropes – or genres and systems.” (Moretti 2013: 48–49) Texts that are findable, accessible, interoperable, and reusable (FAIR) are expected to engage readers in the coming years, while the fact that only a few recent translations of philosophical works are available via open-access on the internet ought to quickly become an issue of the past (Schäfer and Serres 2016).

Advances in technology enable the history of philosophy to exercise an influence beyond its narrowly understood disciplinary borders, to reach scholars of different disciplines worldwide and far into the future. However, individual scholars continue to lag behind and remain somewhat ill-equipped to deal with the challenges of the digital transformation we face in our globalized era. As Timothy Wil-

liamson (1998) has said, philosophy is a science, but not a natural science (mathematics is another example of a non-natural science; it is, rather, a language of rigorous demonstration). At its best, philosophy strives to be as systematic, rigorous, precise, accurate, critical, and evidence-based as its questions permit and to use the best methods available to answer them. We are only beginning to become aware that digital rights management is a key enabling technology. Considering current trends towards a *data-driven history of philosophy* as a branch of both philosophy and digital humanities (Betti et al. 2019), our point is that the future of the history of philosophy urgently depends on finding ways to bring about radical enhancements of the way we edit, store, annotate, access, and translate corpora.

When we propose to look into *corpora talking to each other*, we are aware of the objection that a corpus does not talk – only human beings who are reading and understanding texts that belong to a corpus can talk. The anthropomorphism is charming. However, it must not cover up crucial details in the act of encoding, which links the texts supposedly in *conversation*, namely the embedding of assumptions and implicit interpretations that make talking possible, but which also prejudice it. Users must understand what annotation entails, the discipline it imposes, the caution it requires of anyone using the results, and the amount of critical work on text analysis, concept modelling, so-called machine learning, and so on (see Lenardič and Fišer 2022). The case for extensive application of CLARIN corpora and tools on this scale is the occasion to consider their potentialities together with their heuristically stimulating and pragmatically sobering limitations.

One of the most dynamic projects in the construction of parallel text corpora of modern languages and the development of reliable tools for alignment and morpho-syntactic annotation of words is InterCorp (Bozzi 2015: 37).¹ The necessity and added value of providing easy access to complex, highly structured philosophical content through corpora that talk to each other have been highlighted in the literature (Pozzo 2016). The aim is to break new ground for knowledge organization systems that produce synergies while optimizing crosswalks for future translation projects involving Chinese, eventually to be applied to other languages (Pozzo 2020b).

An interesting precedent is the ERC-AdG-2009 project led by Cristina D’Ancona, “Greek into Arabic: Philosophical concepts and linguistic bridges” (G2A) which aligns passages from Plotinus’s *Enneads* with their ninth-century Arabic translation in the text known as *Theologia Aristotelis*. From the point of view of sociolinguistics, of particular interest are the sentences from the original text that would have been difficult to understand for those who lived and were formed in a different cultural environment and who, moreover, were dedicated to conveying

¹ <https://ucnk.ff.cuni.cz/cs/>

ideas, philosophical concepts, moral, and religious principles from one culture to another (Bozzi 2015). The idea of *Kangde* goes beyond the G2A in four ways. First, the extent of Kant's complete *Druckschriften* is far larger than the individual passages of Plotinus. Second, *Kangde* is meant to develop a research interface with functionalities for parallel view and search, and interfaces to other research tools and networks, which is planned to offer a wider spectrum of functions than the G2A Web App (a resource offered by ILC CNR, leading institution of CLARIN-IT and host of the ILC4CLARIN B centre).² Third, the access being tied to validated contemporary translations (starting with Chinese and potentially extended to other languages) the interface is expected to be used by philosophers in the years to come for new multilingual investigations, with a different impact from that of a scholarly discussion of a manuscript tradition. Fourth, the tackling of contemporary Chinese contributes to a living language's morphological and syntactic enrichment, while G2A is about ancient Greek and Arabic, which are dead languages.

We find an analogous endeavour in the project to translate the *Corpus Iuris Iustinianaeum* into Chinese (*Luoma fa* 罗马法), which has made considerable progress in China studies. Not only have 16 volumes been published so far (Schipani 1994–2001, 2001–2021; see Colangelo 2015), but most importantly, Chinese terms have been charged with new, more precise meanings. However, the *Luoma fa* 罗马法 does not offer users any interface. Instead, it remains in published volumes on paper, which means it is not open to annotation and represents only an initial stage of implementing the alignment of translations among corpora. As regards philosophical terms, Timon Gatta has pointed to the *linguistic-lexical development of contemporary Chinese*, which the gradual introduction of Western philosophical production, especially through published translations, has enriched with new terms: the main issue here is “to adequately conform the new discipline [of philosophy] to East Asia's millennial philosophical speculations about religion, moral habits, political and social behavior.” (Gatta 2020: 193, 194)

The methodology and tools are appropriate to achieve the objectives of a parallel consideration of Kantian texts in German and Chinese insofar as it is based on tools such as vocabularies, ontologies, concordances, frequencies – more generally, on the analysis of texts and corpora, which integrates quantitative and formal methods into the portfolio of methods in the history of philosophy and intellectual history. Generally, we take up the *text-corpus method*, which derives a set of abstract rules that govern a natural language from texts in that language, and explores how it relates to others (Baker 1993). We also take up

2 https://g2a.ilc.cnr.it/Teologia_Wapp/Home.xhtml

approaches from science and technology studies with regard to research infrastructures-based innovation. The scientific approach is empirical, it is about presenting Kant's writings in a digital edition and operationalize his terminology for corpus linguistic questions.

Using the CLARIN resource families fully enhances the fruitful interaction among the history of philosophy, computational linguistics, machine learning, translation studies, and China studies. To achieve this truly interdisciplinary vision, we aim to integrate the methodologies of five different fields, thereby pursuing a disruptive overarching approach. Methodology and tools are understood to play an enabling role. First and foremost, however, the group that advances *Kangde* relies on the methodology of the history of concepts in its global extension (Betti and van den Berg 2016; Pichler et al. 2020; Pozzo 2021). What is more, the group takes advantage of achievements that have proven to be particularly effective for the advancement of the history of philosophy from a global perspective, such as the English-French *Vocabulaire de Philosophie* (INIST 2018, a CLARIN lexicon),³ the *Lessico Intellettuale Europeo* (Gregory et al. 1967–2022), and the *Key Concepts in Chinese Thought and Culture* (Wang Lin and Han Zhen 2015–2021).

3 Edition

Due to the celebrations of the tercentenary of Kant's birth, the history of the editions of the work is expected to reach a turning point in 2024 when the Berlin-Brandenburgische Akademie der Wissenschaften (BBAW) and the De Gruyter publishing house will present the new complete edition of the published writings, that is volumes 1–9 of the Kant Academy Edition (BBAW 2022–2024; see Gerhardt 2007; BKGE 2016).

Before going into alignment issues, we are aware we need first to open up Kant's re-established polygraphy for systematic text analysis of conceptual networks, which is now feasible, for the current (and new) Kant Academy Edition – thanks to the efforts of the De Gruyter publishing house – has been reset as proprietary HTML files and offers rich material for experimenting with reflected text analytics and machine learning (BBAW 2022–2024). The editions sponsored by the BBAW started with the *Aristotelis Opera* edition of Immanuel Bekker in the nineteenth century (continued by Olof Gigon in the twentieth century), which was followed by – among others – the editions of Gottfried Wilhelm Leibniz and Wilhelm von Humboldt. In 1894, Wilhelm Dilthey initiated the Kant Academy Edition to

³ <https://www.ortolang.fr/market/terminologies/philosophie/v1.1>

provide reliable and complete texts for scholars and students. At Dilthey's time, the Kant-Kommission (in the predecessor of the BBAW) asked the editors to iron out most orthographic and syntactic variants. Since Kant's orthographical habits – so argued the editors of the first volume of the *Druckschriften*, which appeared in 1902 – are neither systematic nor consequential, the Kant-Kommission thought it better not disturb most readers with obsolete forms (BBAW 1968: 1.513). Hence, Kant's works from 1747 onward were rewritten using orthography and punctuation of Kant's works after the *Kritik der reinen Vernunft*, with the result that Kant's polygraphy was lost.

For this reason, the first move by the editors of Kant's *Neuedition* was to submit queries to CLARIN's historical corpora in order to check Kant's polygraphy and see whether variants were in use at the time. Hansmichael Hohenegger and Riccardo Pozzo have found numerous examples of Kant's polygraphy (BBAW 2022–2024). Let us just mention the many cases of oscillating orthography such as *ascendat/adscendat*, *caussa/causa*, *Cirkul/Cirke*, *drücken/drucken*, *excentum/exemptum*, *exsistentia/existentia*, *Heerde/Herde*, *kömmt/kommt*, *promptus/promptus*, *siehet/sieht*, *soepenumero/saepenumero*, *sumptum/sumtum* (BBAW 1968: 1.514–516). The old Academy edition accounts neither for oscillations in the use of *v* and *u* as in *vniuersalitas/universalitas*, nor in the use of *f* and *s* as in *vniuerfalitas*. Also interesting is Kant's consistent usage of *quum* for causality and of *cum* for togetherness, which marks a grammatical difference, although it does not belong to Classical Latin. Finally, the old Academy edition irons out most capitalizations that Kant evidently used to stress the term's meaning as a *terminus technicus*, as was pointed out previously by Johann Joachim Lange (1734: 372; see Hohenegger 2020). Concerning editorial decision making on reading a word as a typo or leaving it in the text on its own account, today it has become indispensable to use CLARIN's historical corpora, such as the LatinISE corpus⁴ and the Deutsches Textarchiv (1600–1900),⁵ as well as, obviously, the DWDS (*Digitales Wörterbuch der deutschen Sprache*),⁶ and among its tools the DTA-CAB (Deutsches Text-Archiv Cascade Analysis Broker).⁷

4 <https://lindat.mff.cuni.cz/repository/xmlui/handle/11372/LRT-3170>

5 <https://clarin.bbaw.de:8088/fedora/objects/dta:3503/datastreams/cmdl/content?asOf-DateTime=2019-09-30T09:20:47.158Z>

6 <https://www.dwds.de>

7 <https://kaskade.dwds.de/~moocow/software/DTA-CAB/>

4 Alignment

Being a user of CLARIN means having access to a whole intangible network of knowledge with specific areas of expertise.⁸ Moreover, the alignment itself is meant to use many of the CLARIN resource families, especially the parallel corpora insofar as they serve as training data for statistical machine translation systems. Parallel corpora make up the largest CLARIN resource family and are central to translation studies and contrastive linguistics. Many of them are accessible through easy-to-use concordancers that considerably facilitate the study of interlinguistic phenomena. CLARIN provides access to 86 parallel corpora, the majority of which are available for download from national repositories and through concordancers such as Korp, Corpuscle, and KonText. CLARIN offers access to 47 bilingual corpora, mostly containing European language pairs but also non-European languages such as Hindi, Tamil, and Vietnamese. 39 corpora are multilingual, five of which contain texts in more than 50 languages. Almost half of the corpora are sentence-aligned, which allows for easy comparative research.⁹ While overviewing the corpora that are already part of the CLARIN resources families, one cannot help seeing the amount of work still to be done for Chinese, which is present, for example, in MultiUN (Multilingual U.N. Parallel Text 2000–2009).¹⁰

The corpora alignment of the German *Urtext* with its Chinese translation will be carried out on the Kant Online platform. The platform is currently under construction.¹¹ Kant Online takes the *Kant-Lexikon* (Willaschek et al. 2015) as its nomenclature. The endeavour consists, in no small part, in the extraction of terminology. The study of terminology is indispensable for a non-arbitrary translation but also for producing non-arbitrary dictionaries. Hence, we should reconsider the possibility of a dictionary with nomenclatures of different granularity: from basic to very fine. To name an analogous undertaking, one can look at the Nietzsche Online platform (Nietzsche Online 2011), which provides access to the complete edition of Friedrich Nietzsche's work by Giorgio Colli and Mazzino Montinari together with almost all publications published by De Gruyter on Nietzsche. In addition to about 70 volumes of the Nietzsche edition, the platform offers access to monographs and reference works such as the *Nietzsche-Wörterbuch* (van Tongeren, Schank & Siemens 2004) and all issues of the *Nietzsche-Studien*: all in all, more than 110,000 book pages. The platform offers significantly

⁸ <https://office.clarin.eu/v/CE-2017-1093-ValueProposition-update2020.pdf>

⁹ <https://www.clarin.eu/resource-families/parallel-corpora>

¹⁰ <http://www.euromatrixplus.net/multi-un/>

¹¹ <https://www.degruyter.com>

more than the sum of its printed content. A philological apparatus that justifies critical choices between variants and historical-critical explanations that provide information about the content and context of the corpus makes it possible to combine the reconstructed text with a textual universe (Pozzo 2014). It should be noted that the Kant Online platform is expected to go beyond Nietzsche Online by providing advanced access and more processing tools for philosophical and linguistic research. Besides this, the interface offers datasets in several formats available to download for future research ventures, tools, and networks. Indeed, the German-to-Chinese interface on Kant Online is meant to be focused on bilingual corpora, which are not considered in Nietzsche Online. Finally, it is constructed for annotation around an adaptation of traditional concept analysis to computational methods, designed by digital humanities scholars to enable a *computational history of ideas* (Betti and van den Berg 2016).

Annotating Kant has been undertaken with increasing regularity over more than 50 years alongside the progress of computational linguistics. The start was given by the *Allgemeiner Kantindex* (Martin 1967; Roser and Mohrs 1992), which gives Kant's words in non-inflected form and is currently preserved within the *Korpora.org* platform.¹² A giant leap forward was achieved by Tullio Gregory et al. (1967–2022) and Norbert Hinske (1982–2019), respectively, with the *Lessico Intellettuale Europeo* (which since its inception used a markup language very similar to TEI and now uses TEI) and the *Kant-Index* (built on TUSTEP), which granted access to Kant's writings in lemmatized form with metadata and semantic annotations that are interoperable with regard to multilingualism (i.e., Kant's use of Greek, Latin, German, and French). The next giant leap forward is expected to be achieved by recontextualizing Kant within multilingual philosophical corpora around computational concept modelling. Once humanities scholars have agreed to study a corpus, such as the ones envisaged by *Kangde*, they first identify appropriate levels and categories of analysis; they then perform annotations on a subsample of the corpus that acts as reference data, which become the basis for “machine learning experiments with candidate model classes, including additional tools or data resources” (Kuhn 2020: 76).

The nine volumes of Kant's printed works, with their 1,580,000 words, offer material for a full lemmatization and a formidable basis for reflected text analytics. Starting from an *Urtext* of German lemmata, it is possible to create an induced network of concepts through which to pursue empirically verifiable hypotheses on meaning shifts over the centuries. Restoring Kant's *Urtext* requires the closest attention for annotation so that the surface text does not lose anything of the orig-

¹² <https://korpora.zim.uni-duisburg-essen.de/kant/>

inal richness while accounting for historical usages, with deeper layers that offer standardized tokens for horizontal investigation. Methods for theory- and data-driven corpus analysis enable scholars to formulate hypotheses regarding systematic patterns in the distribution of specific concepts in a corpus and test them empirically (Kuhn 2020). For example, one might try to verify a presumed tendency for a school of thinking to translate the term *A* as *A'* in the context of debate *X*, but as *A''* in other contexts. This is what happened with the first translation of some passages into French (in 1788, i.e., at the very end of the Enlightenment) from Kant's *Kritik der reinen Vernunft* (published in 1781 and again in 1787), when the word *Vernunft* was rendered as *raison* in some contexts and as *entendement* in others (Müller and Pozzo 1988). In this perspective, Chinese offers a particularly challenging state of the art. Some sinologists – one thinks first and foremost of Marcel Granet (1968: 7) – have maintained that the difficulty of mutual understanding between Western and Chinese cultures might lie in the impossibility of Chinese to express logically defined and precisely circumscribed concepts that are necessary for philosophical arguments. However, current understandable and faithful Chinese translations of many Western philosophical works – and the translation of Kant's work by Li Qiuling 李秋零 (2003–2019) is certainly one – show that this assumption is incorrect and biased by cultural preconceptions. This is where the idea of *Kangde* reveals its added value insofar as it provides computational concept modelling of Kant's terminology referred to a validated Chinese translation.

5 Western Grammar in Contemporary Chinese

This enterprise is about creating a multilingual textual database knowledge extraction program for enabling context-guided lexical analysis in the form of an *open-ended knowledge-based architecture* that provides users with access to datasets while including the corpus in the LLOD cloud.¹³ For instance, in the cultural exchange between China and the West, the history of philosophy can play a significant role, notwithstanding the difficulties of engaging with the mutual textual legacy. We are talking of momentous cultural exchanges that raise awareness of the need for a culturally sensitive approach to different traditions, including challenges related to cultural and religious diversity.

Tradi, *perpoliri*, and *transferre* are terms that express Cicero's commitment to bringing over philosophical texts from Greece to Rome. They are the foundation

¹³ <https://linguistic-lod.org>

pillars of the *translatio studiorum* from Greek to Latin, which lasted for centuries. *Transferre* and *translatio* lie at the root of neosemic creativity: under certain conditions, writes Quintilian, “*nesesse sit transferre aut circumire*” (*De institutione oratoria* XII, 10, 34). Tullio Gregory (2012: 6) has suggested one could inscribe in the hendiadys *transferre aut circumire* the history of all problems related to translating. Boethius was well aware of this, and so too was Cassiodorus in the sixth century AD, that is, in the decades that saw the rise and the fall in the Latin West of that final renaissance of Hellenism, which marked the sunset of the ancient world.

In contrast with Western languages, Chinese does not allow free use of any Greek or Latin etymology. The long and arduous process of defining a Chinese philosophical lexicon undertaken during the last decades of the nineteenth and the first half of the twentieth century is not a mere linguistic issue. It also involves issues of political and social acceptance of the influence of the West over China, its culture, and its way of thinking. This process did not only consist in introducing philosophy into China as a new branch of knowledge and making it acceptable to and consistent with the intellectual sensibility of the ruling class, while introducing new terms for new ideas (Pozzo 2018). The main issue was to adequately conform the new discipline of philosophy to East Asia's millennial religions, moral habits, political, and social behaviors (Gatta 2020).

As regards Kant studies in China, the Chinese Kant Society was established at Peking University in June 2019, as the final stage of a confrontation with Kant's works that has pervaded the entire twentieth century, and at the center of which was the philosopher Mou Zongsan 牟宗三, a leading figure of contemporary neo-Confucianism (Heubel 2016; Gatta 2022). Since Chinese scholars began to actively study and research Western culture at the beginning of the twentieth century, Kant was perceived as a challenge in systematic and lexical fields. These two fields were interconnected, so that different lexical renditions have helped Chinese scholars adapt and domesticate Kant's theories using words rooted in China's literary and philosophical traditions. The introduction, translation, and adaptation of Kant's philosophy in China have greatly influenced modern Chinese philosophy and have had a key role in the formation and standardization of a modern Chinese philosophical vocabulary.

Interestingly, we have started reflecting on *Kangde* due to the impact the alignment of corpora can have on the development of the so-called *Western Grammar in Contemporary Chinese-Xiandai hanyu ouhua yufa* 现代汉语欧化语法 (Masini 2009: 648–650; Gatta 2022: 8), which has been proven to produce not only terminological enrichment but also significant modifications – both morphological and syntactic – of Chinese grammar. Translation corpora such as those studied for *Kangde* provide an ample repertoire of translation strategies

(Zanettin 2014). The alignment itself can be tied to the existing anchor points: in the paratext, these are the pages of the original editions and the lines of the old (and new) Kant Academy Edition; and in the text, the pericopes, and the periods. For this purpose, we can use unsupervised sentence aligners for symmetrical and asymmetrical parallel corpora.

From the point of view of translation theory, *Kangde* is about encoding a source language (German) through the translational language (machine-operated) to a target language (Chinese) to be decoded. The reverse process is a feasible possibility. We know of two types of translation universals (Mauranen 2007): one shapes the process from the source to the target text (S-universals), while the other (T-universals) compares translations to other target-language texts. The distinctive features of translational language can be identified by comparing translations with similar native texts, thus throwing new light on the translation process and helping to uncover translation patterns, or what William Frawley (1984) has called the *third code of translation*. The most precious added value of the *Kangde* idea lies in facilitating access to validated translations of complex texts. To this purpose, orientation among CLARIN corpora, lexica, and tools includes the Sheffield Corpus of Chinese Annotation (of the Oxford Text Archive),¹⁴ GATE (General Architecture for Text Engineering),¹⁵ and the BilingBank (of TalkBank).¹⁶ *Kangde* ought to empower Chinese readers (and, indeed, Western readers) with automatically generated references for words, whose translation and definition they might otherwise have to look for in glossaries or vocabularies, “because graphically the term would not contain any clue as to its meaning” (Gatta 2020: 201; see Fan Bingqing 1926).

Translating Western philosophy into Chinese is a complex phenomenon involving the linguistic-lexical development of contemporary Chinese through the gradual introduction of Western philosophical production, especially through published translations (Masini 1993). For example, Timon Gatta has presented a selection of exemplary concepts that attest to the formative process of China’s philosophical lexicography (Fan Bingqing 1926). Western philosophical terms have reached standardized translations in Chinese through similar, yet not identical paths of explicitation, simplification, normalization, sanitization, and levelling out. Think, for instance, of the long history that has led to establishing the current Chinese terms for logic-*luoji* 逻辑, metaphysics-*xing er shang xue* 形而上学, and aesthetics-*meixue* 美学 (Kurtz 2011; Gatta 2020).

¹⁴ <https://ota.bodleian.ox.ac.uk/repository/xmlui/handle/20.500.12024/2481#>

¹⁵ <https://gate.ac.uk>

¹⁶ <https://biling.talkbank.org>

Translating Kant into Chinese offers a striking visualization of a *third code in motion* by means of increasingly successful adaptations of translated language to native language. As Timon Gatta has explained, the lexical renderings (phonetic loans or semantic loans) of Western concepts that Chinese translators have experimented with over the centuries were, initially, hardly capable of adequately expressing the richness of meanings and nuances of the original language. Given the difficulty in the Chinese language of embracing words from other languages, however, translators have been forced, step by step, to look for one- or two-character words that recall the original meaning of the foreign term, often with results that are anything but satisfactory (Gatta 2020: 200–201). For example, if the rendering of intellect-*Verstand-zhixing* 知性, has been established in all translations of Kant's three *Critiques* over the past 50 years (Gatta 2021: 95), the rendering of phenomenon-*Phänomen/Erscheinung-xianxiang* 现象 tells a different story, for it was seemingly established very early but underwent recent oscillations with, for example, Li Qiuling 李秋零 (2003–2019) who established a character that includes the meaning of appearing, of showing itself (Gatta 2021: 312). The few dozen cases in which Kant uses *Phänomen/Erscheinung* actually mean a 'surprising case' in the context of the antinomic nature of the higher faculties complicate the translation but help refine the terminological analysis (Hohenegger 2020: 346–349). This effect is even more pronounced in the case of the translations of transcendental-*transzendental-xiayan* 先验, which sparked a debate in Japan and China during the first decades of the last century, so that, even now, one finds different opinions about it (Gatta 2022: 177–191).

6 Forward look

Philosophy requires critical editions and hermeneutics for text interpretation, while translation studies require attention to history and trust (Rizzi, Lang & Rym 2019). A translation "is always an interpretation, as shown by the connection of terms with the synonymic values *interpretari*, *vertere*, and *transferre*" (Gregory 2012: 4). From this perspective, the ground-breaking element of our vision lies in letting *corpora talk to each other*, and not simply individuals born in different parts of the world. Corpora are constituted according to the type of the text, the theme to be translated, and the target language. They are search-accessible complete collections of traditions of texts, with corresponding dictionaries, thesauri, and reference works. They are instrumental in engaging with traditions in innovative ways.

This chapter shows how *corpora alignment provides further steps towards enhancing data-driven philosophical translation* (Frawley 1984; Mauranen 2007; McEnery and Xiao 2007; Xiao and Yue 2008; Xiao, Lianzhen & Yue 2010; Zanettin 2014; Pozzo 2016). *Kangde* belongs to the long history of traditional translation techniques and theories that go back to Latin translations of Greek. The questions that ought to be posed reflect the vast differences in culture, which have to be bridged between European philosophy, as represented by Kant, and traditional Chinese thought, which cannot be described as philosophical in the Western sense.

All translations are likely to show specific linguistic characteristics simply by virtue of being translations – characteristics that are caused in and by the process of translation. The effect of the source language on the translation is strong enough to make the translated language perceptibly different from the target native language. Consequently, translational language (*Translationese*) is at best a particular unrepresentative variant of the target language (McEnery and Xiao 2007). Translational language entails the elimination of ambiguities regarding the choice of one word over another and has four core patterns of lexical use: a relatively lower proportion of lexical words over function words, a relatively higher proportion of high-frequency words over low-frequency words, a relatively more significant repetition of the most frequent words, and a smaller vocabulary (Xiao, Lianzhen & Yue 2010). Centuries before machine translation, famous historical examples of token-to-token translations include William of Moerbeke's translations of philosophical, medical, and scientific texts from Greek into Latin, in particular, of many works by Aristotle, which he did at the request of Aquinas between 1253 and 1286. William's translations were literal (*de verbo in verbo*), faithful to the spirit of Aristotle, and without elegance, that is, without any attempt at diminishing the impact of both his rudimentary command of Greek and of the primitiveness of medieval Latin philosophical terminology, which shows that the embedding on which machine translation is based existed long before machines. While William of Moerbeke's Aristotle are texts written in what we call today translational language, the translations of Plato from Greek into Latin by Marsilius Ficinus between 1462 and 1484 represent a famous example of a literary translation that is quite close to the native target language. We recall William and Marsilius to make it clear where the challenge lies. Machine translation of philosophical texts today produces, at best, William's of Moerbeke translational language, while the idea of *Kangde* is to boost machine translation until it pushes the third code so as to mould the translation into the native language, that is, to make it as close as possible to the results achieved by Ficinus. It is important to note that the alignment of two or more philosophical corpora will add substantial numbers of datasets to enable machine translation, training, and data devel-

opment. Today, the role of machine translation in assisting with the translation of literary texts shows both limitations and potential benefits. A key challenge in literary translation is preserving the meaning (as in other domains such as technical translation) and the reading experience, which means that a literary translator must carefully select from possible options (Toral and Way 2015, 2018).

A close study of the Chinese translation of Kant's writings is useful in gauging the reception of Kant's thinking within the limitations of Chinese semantics. The value of the aligned corpora is also useful for the study of the mechanics of translations into very different linguistic environments, which could eventually be instrumental for computer-based translations. The great challenge remains the *protection of datasets under intellectual property rights* (IPR). Our idea is to tackle this challenge from the very beginning because, thanks to an administrative system that manages inclusion and consultation rights, we wish to settle IPR issues of the German and the Chinese texts for making them open access for users in an editorial setting that fully exploits both government-sponsored research (BBAW) and the efforts of two prestigious publishing houses (De Gruyter in Berlin and China Renmin Press in Beijing). The envisioned interface is meant to connect German and Chinese texts first. It is structured, however, to be scaled up to other languages. On top of boosting Kantian philosophical reception in China, straight from German into Chinese, *Kangde* aims to reach out to communities of practices that receive and confer datasets and tools to research infrastructures such as CLARIN. *The challenge of the sustainability of the Kangde endeavour can be effectively tackled by conferring datasets to CLARIN while reusing its corpora, lexica, and tools.* As Martin Wynne has made clear, CLARIN is “keen to deal with all non-European languages, including major world languages such as Arabic, Chinese, Russian, Japanese, etc.”¹⁷

7 Conclusion

Wrapping up, this chapter lays out some interesting use cases of corpora, corpus linguistics, computational linguistics, natural language processing, and their contribution to digital humanities. It suggests approaches that impact humanities research through digital media, artificial intelligence, data mining, and machine learning. In connection with the CLARIN resource families, the chapter fosters the adoption of FAIR data standards, which

¹⁷ <https://www.clarin.eu/blog/users-clarin-who-are-they>

stimulates the reuse and repurposing of available research data, thereby enabling scholars in the SSH – including the DH – to increase their productivity and open new research venues in and across disciplines that address one or more of the multiple societal roles of language: as a carrier of cultural content and information, both synchronically and diachronically, as a reflection of scientific and instrumental knowledge, as an instrument for human communication, as one of the central components of the identity of individual groups, cultures, or nations, as an instrument for human expression, as an object for study and preservation.

(ESFRI 2018: 213)

All things considered, then, this chapter engages research agendas that “illustrate the added value of well-supported access to the wealth of data types that are available for multiple languages are the research initiatives for the study of migration patterns, intellectual history, language variation across period and region, dynamics in mental health conditions, customer opinions and parliamentary discourse, just to name a few” (de Jong 2019: 123).

We are looking forward to fruitful cooperation between CLARIN and Chinese-speaking infrastructures, for our project is about *cultural innovation* (Pozzo et al. 2020) in very concrete terms. Philosophy is, in fact, one of the core SSH disciplines, for which widespread use of language data is central to many key methods. Last but not least, we will discuss the *Kangde* vision at two events of global impact planned for the year 2024, which will focus on the tercentenary of Kant’s birth: the 14th International Kant Congress in Bonn and the 25th World Congress of Philosophy in Rome.

Bibliography

- Baker, Mona. 1993. Corpus linguistics and translation studies: Implications and applications. In Mona Baker, Gill Francis & Elena Tognini-Bonelli (eds.), *Text and Technology*, 232–252. Philadelphia & Amsterdam: Benjamins.
- BBAW: Berlin-Brandenburgische Akademie der Wissenschaften (ed.). 1968. *Kants Werke: Akademie Textausgabe*. 9 vols. Berlin: De Gruyter.
- BBAW: Berlin-Brandenburgische Akademie der Wissenschaften (ed.). 2022–2024. *Immanuel Kant, Gesammelte Schriften. Abteilung 1 – Neuedition*. 9 vols. Berlin: De Gruyter.
- Betti, Arianna & Hein van den Berg. 2016. Towards a computational history of ideas. *CEUR Workshop Proceedings* 1681.
- Betti, Arianna, Hein van den Berg, Yvette Ortwin & Caspar Treijtel. 2019. History of Philosophy in Ones and Zeros. In Eugen Fischer & Mark Curtis (eds.), *Methodological Advances in Experimental Philosophy*, 295–332. London: Bloomsbury.
- BKGE: Bundesinstitut für Kultur und Geschichte der Deutschen im östlichen Europa. 2016. *300 Jahre Immanuel Kant: Der Weg ins Jubiläum*. Oldenburg: BKGE.
- Blair, Ann, Paul Duguid, Anja-Silvia Goeing & Anthony Grafton (eds.). 2011. *Information: A historical companion*. Princeton: Princeton University Press.

- Bourdieu, Pierre. 2002. Les conditions sociales de la circulation internationale des idées. *Actes de la recherche en sciences sociales* 145. 5–9.
- Bozzi, Andrea. 2015. Greek into Arabic: A research infrastructure based on computational models to annotate and query historical and philosophical digital texts. In Andrea Bozzi (ed.), *Digital texts, translations, lexicons in a multi-modular web application: Methods and samples*, 27–42. Florence: Olschki.
- Colangelo, Lara. 2015. L'introduzione del diritto romano in Cina: Evoluzione storica e recenti sviluppi relativi alla traduzione e produzione di testi e all'insegnamento. *Roma e America: Diritto romano comune* 36. 175–210.
- Draxler, Christoph, Alexander Geyken, Erhard Hinrichs, Annette Klosa-Kückelhaus, Elke Teich & Thorsten Trippel. 2022. How to connect language resources, infrastructures, and communities. In Darja Fišer & Andreas Witt (eds.), *CLARIN: The infrastructure for language resources*. Berlin: De Gruyter.
- ESFRI. 2018. *Roadmap 2018: Strategy report on research infrastructures*. Milan: ESFRI.
- Fan Bingqing 樊炳清. 1926. *Zhexue cidian* 《哲學辭典》 [Philosophical Dictionary]. Beijing: Business Press 商务印书馆.
- Floridi, Luciano. 2019. *The Logic of information*. Oxford: Oxford University Press.
- Frawley, William. 1984. *Translation: Literary, linguistic, and philosophical perspectives*. Wilmington: University of Delaware Press.
- Gatta, Timon. 2020. The translation of western philosophical terms in Chinese: The case studies of Logic, Metaphysics and Aesthetics. In Marina Miranda (ed.), *Dal Medio all'Estremo Oriente 2: Studi del dottorato di ricerca in Civiltà dell'Asia e dell'Africa*, 193–219. Rome: Carocci.
- Gatta, Timon. 2022. *Lo sviluppo del lessico filosofico nel cinese moderno*. Florence: Olschki.
- Gerhardt, Volker. 2007. Erschließung und Sicherung des kulturellen Erbes: Zur Aktualität des Forschungsprogramms der Akademien. In Annette Sell (ed.), *Editionen: Wandel und Wirkung*, 3–9. Tübingen: Niemeyer.
- Granet, Marcel. 1968. *La pensée chinoise*. Paris: Albin Michel.
- Gregory, Tullio. 2012. Translatio Studiorum. In Marco Sgarbi (ed.), *Translatio Studiorum: Ancient, medieval and modern bearers of intellectual history*, 1–21. Leiden: Brill.
- Gregory, Tullio, Antonio Lamarra, Cesare Pasini & Riccardo Pozzo (eds.). 1967–2022. *Lessico intellettuale europeo*. 129 vols. Florence: Olschki.
- Guyer, Paul & Allen B. Wood (eds.). 1992–2016. *Cambridge Edition of the Works of Immanuel Kant*. 16 vols. Cambridge: Cambridge University Press.
- Hajičová, Eva, Jan Hajič, Barbora Hladká, Jiří Mírovský, Lucie Poláková, Kateřina Rysová, Magdaléna Rysová, Pavel Straňák, Barbora Štěpánková & Šárka Zikánová. 2022. Corpus annotation as a feasible and scientifically beneficial task. In Darja Fišer & Andreas Witt (eds.), *CLARIN: The infrastructure for language resources*. Berlin: De Gruyter.
- Heubel, Fabian. 2016. *Chinesische Gegenwartsphilosophie: Zur Einführung*. Hamburg: Junius.
- Hinske, Norbert (ed.). 1982–2019. *Kant-Index*. 55 vols. Stuttgart-Bad Cannstatt: Frommann-Holzboog.
- Hohenegger, Hansmichael. 2020. Philologie und Übersetzung: Technische Ausdrücke in Kants philosophischer Sprache. In Gisela Schlüter & Hansmichael Hohenegger (eds.), *Kants Schriften in Übersetzungen, Archiv für Begriffsgeschichte Sonderheft 15*, 337–366. Hamburg: Meiner.

- Institut de lu2019information scientifique et technique-CNRS UPS76 (INIST). (2018. *Vocabulaire de philosophie* [Terminologie]. ORTOLANG (Open Resources and TOols for LANGuage) – www.ortolang.fr, v1.1, <https://hdl.handle.net/11403/philosophie/v1.1>.
- Jong, Franziska de. 2019. CLARIN: Infrastructural support for impact through the study of language as social and cultural data. In Bente Maegaard, Riccardo Pozzo, Alberto Melloni & Matthew Woolard (eds.), *Stay Tuned to the Future: Impact of Research Infrastructures for Social Sciences and Humanities*, 121–129. Florence: Olschki.
- Kuhn, Jonas. 2020. Computational text analysis within the humanities. In Nils Reiter, Axel Pichler & Jonas Kuhn (eds.), *Reflektierte algorithmische Textanalyse*, 61–106. Berlin: De Gruyter.
- Kurtz, Joachim. 2011. *The discovery of Chinese logic: Genealogy of a twentieth-century discourse*. Leiden: Brill.
- Lange, Johann Joachim. 1734. *Verbesserte und Erleichterte Lateinische Grammatica*. Halle: Waisenhaus.
- Lenardič, Jakob & Darja Fišer. 2022. The CLARIN Resources and Tool Families. In Darja Fišer & Andreas Witt (eds.), *CLARIN: The infrastructure for language resources*. Berlin: De Gruyter.
- Li Qiuling 李秋零 (ed.). 2003–2019. *Kangde zhuzuo quanji* 《康德著作全集》集 *Di yi jie* 第一节 [The Complete Works of Immanuel Kant: Section 1]. 9 vols. Beijing: China Renmin Press 中国人民大学出版社.
- Maegaard, Bente, Dieter Van Uytvanck & Steven Krauwer. 2017. *CLARIN Value Proposition*. Utrecht: CLARIN ERIC.
- Martin, Gottfried. 1967. *Allgemeiner Kantindex zu Kants gesammelten Schriften*, ed. Dieter Krallmann. Berlin: De Gruyter.
- Masini, Federico. 1993. *The formation of modern Chinese Lexicon and its evolution toward a national language: The period from 1840 to 1898*. Berkeley: Project on Linguistic Analysis.
- Masini, Federico. 2009. La riforma della lingua. In Guido Samarani & Maurizio Scarpari (eds.), *La Cina*, vol. III, *Verso la modernità*, 621–662. Turin: Einaudi.
- Mauranen, Anna. 2007. Universal tendencies in translation. In Margaret Rogers & Gunilla Anderman (eds.), *Incorporating corpora: The linguist and the translator*, 32–48. Clevedon: Multilingual Matters.
- McEnery, Tony & Richard Xiao. 2007. Parallel and comparable corpora: What is happening? In Margaret Rogers & Gunilla Anderman (eds.), *Incorporating Corpora: The Linguist and the Translator*, 18–31. Clevedon: Multilingual Matters
- Moretti, Franco. 2013. *Distant Reading*. London: Verso.
- Müller, Gerhard & Riccardo Pozzo. 1988. Charles Bonnet: Bonnet critico di Kant: Due *Cahiers* ginevrini del 1788. *Rivista di storia della filosofia* 43 (1). 131–64.
- Nietzsche Online (NO). 2011. <https://doi.org/10.1515/nietzsche>.
- Palmquist, Stephen. 1995. *A complete index to Kemp Smith's translation of Immanuel Kant's Critique of Pure Reason*. Oxford: Oxford University Computing Services. <http://staffweb.hkbu.edu.hk/ppp/indx/toc.html>.
- Pichler, Axel, André Blessing, Nils Reiter & Mirco Schönfeld. 2020. Algorithmische Mikrolektüre philosophischer Texte. In Nils Reiter, Axel Pichler & Jonas Kuhn (eds.), *Reflektierte algorithmische Textanalyse*, 327–372. Berlin: De Gruyter.
- Pozzo, Riccardo. 2014. “Nietzsche Online: A critical appraisal.” *Lexicon Philosophicum* 2. 337–341.

- Pozzo, Riccardo. 2016. Corpora that talk to each other. In Suwanna Satha-Anand, Kanit Sirichan & Lowell Skar (eds.). *Proceedings of the international symposium: Philosophies in dialogue: Bridging the great philosophical divides: 26–28 March 2015*, 235–245. Bangkok: Chulalongkorn University.
- Pozzo, Riccardo. 2018. 《主旨演讲: 东西方哲学: 创新, 反思与包容: 章含舟》— Zhuzhi yanjiang: Dongxi fang zhexue: Chuangxin, fansi yu baorong: Zhanghanzhou — Keynote address: East-west philosophy: Innovation, reflection, and inclusion. In Li Nian 李念 (ed.), 在这里, 中国哲学与世界相遇: 24 位世界哲学家访谈录 *Zai zheli, Zhongguo zhexue yu Shijie Xiangyu: 24wei Shijie Zhexuejia Fangtanlu – Interviews of 24 philosophers all over the world (Chinese edition)*, 379–387. Beijing: China Renmin Press 中国人民大学出版社.
- Pozzo, Riccardo. 2020a. Bilingualism and multilingualism in Chinese and Western philosophy. 中国学第八辑 *China Studies Quarterly* 8. 56–67.
- Pozzo, Riccardo. 2020b. Blick nach vorn: Kant-Übersetzungen und Korpora. In Gisela Schlüter & Hansmichael Hohenegger (eds.), *Kants Schriften in Übersetzungen, Archiv für Begriffsgeschichte Sonderheft 15*, 323–334. Hamburg: Meiner.
- Pozzo, Riccardo. 2021. *History of Philosophy and the Reflective Society*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110709292>.
- Pozzo, Riccardo, Andrea Filippetti, Mario Paolucci & Vania Virgili. 2020. What does cultural innovation stand for? Dimensions, processes, and outcomes of a new innovation category. *Science and Public Policy* 47 (3). 425–436. <https://doi.org/10.1093/scipol/scaa023>.
- Rizzi, Andrea, Birgit Lang & Anthony Rym. 2019. *What is translation history?* London: Palgrave.
- Romele, Alberto. 2020. *Digital hermeneutics: Philosophical investigations in new media and technologies*. London: Routledge.
- Roser, Andreas & Thomas Mohrs. 1992. *Kant-Konkordanz zu den Werken Immanuel Kants*. 10 vols. Hildesheim: Olms.
- Schäfer, Valérie, and Alexander Serres. 2016. *Histories of the Internet and the Web*. Bern: Infoclio. <https://doi.org/10.13098/infoclio.ch-lb-0006>.
- Schipani, Sandro (ed.). 1994–2001. *Corporis Iuris Civilis Fragmenta Selecta-Luoma fa yu xiandai minfa* 罗马法与现代民法. 6 vols. Rome-Beijing: Sapienza-Consiglio Nazionale delle Ricerche-Zhongguo Zhengfa Daxue 中国政法大学.
- Schipani, Sandro (ed.). 2001–2021. *Corporis Iuris Civilis Digesta-Luoma fa yu xiandai minfa* 罗马法与现代民法. 20 vols. Rome-Beijing: Sapienza-Consiglio Nazionale delle Ricerche-Zhongguo Zhengfa Daxue 中国政法大学.
- Schlüter, Gisela & Hansmichael Hohenegger (eds.). 2020. *Kants Schriften in Übersetzungen. Archiv für Begriffsgeschichte: Sonderheft 15*. Hamburg: Meiner.
- Shanghai Times Press. 2010. *Jinxiandai Hanyu ciyuan* 《近现代汉语辞源》 [Dictionary of Contemporary Chinese]. Shanghai: Shanghai Times Press 上海辞书出版社.
- Tongeren, Paul van, Gerd Schank & Herman Siemens (eds.). 2004. *Nietzsche Worterbuch*. Berlin: De Gruyter.
- Toral, Antonio & Andy Way. 2015. Machine-assisted translation of literary text: A case study. *Translation Spaces* 4 (2). 240–267. <https://doi.org/10.1075/ts.4.2.04tor>.
- Toral, Antonio & Andy Way. 2018. What level of quality can neural machine translation attain on literary text? <https://doi.org/10.48550/arXiv.1801.04962>.
- Tu Weiming 杜維明. 2010. *The global significance of concrete humanity: Essays on the Confucian discourse in cultural China*. New Delhi: New Delhi Centre for Studies in Civilizations.

- Tuschling, Burkhard & Nelly Motroshilowa (eds.). 1994–2020. *Immanuel Kant: Werke: Zweisprachige deutsch-russische Ausgabe*. 6 vols. Moskow: RAS Institute of Philosophy.
- Xiao, Richard & Ming Yue 名月. 2009. Using corpora in translation studies: The state of the art. In Paul Baker (ed.), *Contemporary approaches to corpus linguistics*, 237–262. London: Continuum.
- Xiao, Richard, He Lianzhen 何莲珍 & Ming Yue 名月. 2010. In pursuit of the third code: Using the ZJU Corpus of Translational Chinese in translation studies. *Using Corpora in Contrastive and Translation Studies* 2010. 182–214.
- Wang Lin 王琳 & Han Zhen 韩震 (eds.). 2015–2021. *Key Concepts in Chinese Thought and Culture*, 6 vols. Beijing: Foreign Language Teaching and Research Press 外语教学与研究出版社 and 6 vols. Singapore: Palgrave-Macmillan.
- Wen Haiming 温海明. 2012. *Chinese Philosophy*. Cambridge: Cambridge University Press.
- Willaschek, Marcus, Jürgen Stolzenberg, Georg Mohr & Stefano Bacin (eds.). 2015. *Kant-Lexikon*. 3 vols. Berlin: De Gruyter.
- Williamson, Timothy. 1998. *The philosophy of philosophy*. London: Wiley.
- Zanettin, Federico. 2014. Corpora in translation. In Juliane House (ed.), *Translation: A Multidisciplinary Approach*, 178–199. London: Palgrave-MacMillan.