

Call for papers

**11th International Conference of the European Association of Chinese Linguistics (EACL-11)
September 25-27, 2020**

The 11th International Conference of the *European Association of Chinese Linguistics* (EACL) will be held in Olomouc, Czech Republic. It is jointly organized by the EACL executive committee, the Department of Asian Studies (<https://kas.upol.cz/en/>), the Sinofon project (<http://sinofon.cz/>), the Department of English and American Studies, and the Department of General Linguistics, Palacký University Olomouc, Czech Republic.

Invited speakers:

Huba Bartos (Research Institute for Linguistics, Hungarian Academy of Sciences, and Eötvös Loránd University, Budapest, Hungary)

Redouane Djamouri (CRLAO, CNRS, Paris)

Ziyin Mai (Chinese University of Hong Kong)

John Whitman (Cornell University)

We invite abstracts (in English) for 20-minute talks (plus 10 minutes for discussion) presenting *unpublished original research* in all areas of Chinese linguistics (excluding applied linguistics and teaching of Chinese). Abstracts should be submitted exclusively via the Easy Chair conference system, using the following link:

<https://easychair.org/conferences/?conf=eacl11>

Abstracts can be submitted to the general session or to one of the two panels:

“Corpus-based approaches to Chinese linguistics: towards increased empiricism” (cf. below)

“Recycling characters: The significance of phonetic loan characters and substitutions in the Medieval Chinese writing system” (cf. below)

Please indicate on the top of the page the option chosen: *General, P Corpora, P Writing*.

The abstract should not exceed one page (examples and references included), with 2.5 cm (1 inch) margins, single spacing and Times New Roman 12pt. Submissions must be anonymous and all author-related file properties must be removed. Please include 2 to 4 keywords right below the title of the abstract, including the relevant sub-domain (syntax, phonology, historical linguistics etc.). No more than one single-authored and one co-authored abstract per person will be accepted.

Papers that are included in the program must be presented by the author(s) and cannot be presented by proxies.

All presenters must be(come) members of the European Association of Chinese Linguistics; on-site registration will be possible at the conference.

Important dates:

Deadline for submission of abstracts: March 20, 2020 (GMT+1)

Notification of Acceptance: May 31, 2020

Further relevant information will be posted in due time on the conference website:

<https://sites.google.com/view/eacl-11/home>

Please direct any inquiries to: joannautseong.sio@upol.cz

Panel proposal

Corpus-based approaches to Chinese linguistics: towards increased empiricism

Anna Morbiato, Ca' Foscari University of Venice and the University of Sydney

Bianca Basciano, Ca' Foscari University of Venice

In the past decades, corpus-based research has been gaining momentum in contemporary linguistics. While corpora, intended as large collections of naturally occurring texts, have always existed, rapid advances in computation and technology have provided tools for faster and more effective corpus construction and consultation. Chinese makes no exception: corpus data are now considered among the main resource for many linguists, while large-scale surveys are beginning to be taken seriously as an important tool for linguistic investigation (Jing-Schmidt 2013: 2). Among the reasons beyond the increasing number of corpus-based studies is the availability of “a myriad of large and publicly available Chinese corpora” (Xu 2015), which include general purpose corpora, such as the CCL (Centre for Chinese Linguistics, Peking University) corpus or the BCC (Beijing Languages and Cultures University Chinese Corpus), interlanguage corpora, such as the BLCU International Corpus of Learner Chinese, and specialized corpora, such as the ZHTenTen simplified Chinese corpus mounted at Sketch Engine, the LDC (Linguistic Data Consortium at UPenn) or the ELRA (European Language Resources Association). The great advantage of corpora lies in the fact that they offer access to large amounts of authentic, naturally occurring language data produced by a variety of speakers or writers, thus providing more robust, statistically significant foundation for linguistic accounts and analyses. There is now considerable emphasis on the reliability of linguistic data, as many scholars now stress the need for a shift to a more empirical mode of investigation: such an approach “energizes theoretical endeavors in the field, as rigorous theoretical advances are grounded in solid empirical data” (Jing-Schmidt 2013).

While the number of corpus-based Chinese studies is steadily increasing, scholars note that most are oriented toward applied linguistics, with the compilation of frequency character/word lists and interlanguage Chinese studies being the most popular types of research (Xu 2015). Among the latest lexical frequency and word list projects, there are the latest national Chinese character list, i.e., the 通用规范汉字表 ‘A General Service List of Chinese Characters’ (released in 2013) and Xiao et al. (2009) *A frequency dictionary of Mandarin Chinese*. Corpus-based researches on second language acquisition and language pedagogy have also been increasing over the last couple of decades, with early projects at BLCU now developed into the BLCU International Corpus of Learner Chinese, and with other studies (Tao 2008, 2009; Xiao 2007; inter alia). On the other hand, scholars agree that corpus-based sentential/grammatical level research is practically negligible as compared with lexical studies. There have been some innovative corpus studies on morphological aspects of Chinese (Sproat & Shih 1996), grammar (Xiao & McEney 2008, 2010; Tao 2004), discourse/pragmatics (Jing-Schmidt & Kapatsinsky 2012), and historical linguistics (Ji 2010; Cook 2011). However, according to Xu (2015), apart from these notable exceptions, Chinese corpus-based theoretical linguistics studies are scarce and by no means the mainstream, partly due to the technological and methodological limitations connected with corpus interrogation. McEney and Xiao (2016) also hold that research in corpus-based descriptive grammar in Chinese is rather sporadic and fragmentary, and has focused on specific linguistic features of interest to individual researchers.

This panel wants to explore this promising, yet relatively underdeveloped area of inquiry: it welcomes proposals that integrate corpus tools with theoretical investigation of Chinese grammar in its various components, including:

- Morphology and semantics
- Event structure and argument alternations
- Syntax and information structure
- Pragmatics and discourse
- Diachronic studies

Studies may be both quantitative and qualitative, as well as synchronic or diachronic. Specifically, the panel aims to address the following questions:

- How can corpora improve current theoretical accounts of Chinese grammar in general?
- What do corpora reveal about the statistical relevance of linguistic phenomena and constructions?
- What are the limitations and the drawbacks of using corpora to investigate Chinese?

References:

- Cook, A. (2011). Recent developments in the use of the plural marker *men* in Modern Standard Chinese in Taiwan. *Chinese Language and Discourse* 2(1), 80–98.
- Ji, M. (2010). A corpus-based study of lexical periodization in historical Chinese. *Literary and Linguistic Computing* 25 (2), 199–213.
- Jing-Schmidt, Z. (2013). *Increased Empiricism: Recent Advances in Chinese Linguistics*. John Benjamins Publishing Company.
- Jing-Schmidt, Z. and V. Kapatsinsky. (2012). The apprehensive: Fear as endophoric evidence and its pragmatics in English, Mandarin, and Russian. *Journal of Pragmatics* 44, 346–373.
- McEnery, T., and R. Xiao. (2016). Corpus-Based Study of Chinese. In *The Routledge Encyclopedia of the Chinese Language*, edited by S. Chan, 438–51. New York: Routledge.
- Sproat, R. and C. Shih. 1996. A Corpus-Based Analysis of Mandarin Nominal Root Compounds. *Journal of East Asian Linguistics* 5, 49–71.
- Tao, H. (Ed.) (2004). *Special Issue: Corpora, Language Use, and Grammar*. *Journal of Chinese Language and Computing* 14(2).
- Tao, H. (2008). The Role of Corpora in Chinese Language Teaching and Teacher Education. In *Issues in Chinese Language Education and Teacher Development*, edited by P. Duff & P, 90-102. Lester. University of British Columbia.
- Tao, H. (2009). Core Vocabulary in Spoken Mandarin and the Integration of Corpus- Based Findings into Language Pedagogy. In *Proceedings of the 21st North American Conference on Chinese Linguistics*, edited by Y. Xiao, 13–27. Smithfield, Rhode Island: Bryant University.
- Xiao, R. (2007). What can SLA learn from contrastive corpus linguistics? The case of passive constructions in Chinese learner English. *Indonesian Journal of English Language Teaching* 3(2), 1–19.
- Xiao, R. and T. McEnery. (2010). *Corpus-based Contrastive Studies of English and Chinese*. London/New York: Routledge
- Xiao, R. and T. McEnery. (2008). Negation in Chinese: A corpus-based study. *Journal of Chinese Linguistics* 36(2), 274–330.
- Xu, Jiajin. 2015. “Corpus-Based Chinese Studies.” *Chinese Language and Discourse. An International and Interdisciplinary Journal* 6 (2): 218–44.

PANEL PROPOSAL:

“Recycling Characters: The Significance of Phonetic Loan Characters and Substitutions in the Medieval Chinese Writing System”

Convener: Christoph Anderl, Ghent University

The notion of “phonetic loan characters,” which is used to translate both the Chinese terms *jiǎjiè* 假借 (‘borrowing’) and *tōngjiǎzì* 通假字 (‘interchangeable characters’), is of great relevance for both understanding the processes involved in the formation of the early Chinese script, as well as for studying the actual practice of writing in the Medieval period, especially those materials extant in manuscript form. In this panel we will question this English term, and define and approach the phenomena it refers to from a diachronic perspective, taking into account both the linguistic theories on this topic that were current among scholars between the Hàn and the Sòng dynasties, as well as evidence found in actual specimens of writing. A thorough analysis of the processes involved in phonetic borrowing and substitution will not only contribute to a better understanding of the medieval Chinese writing system and the phonological system behind it, but will also be of help for philologists, editors, and other researchers dealing with texts in which this is an important feature.

For our understanding of phonetic substitution and borrowing, it is important to explore how contemporary scholars defined and explained these concepts, and how they dealt with the various types of substitutions of Chinese graphs encountered in Chinese writing. The understanding of (phonetic) loan graphs in Medieval China has been deeply influenced by Xǔ Shèn’s 許慎 reflections on the evolution of the Chinese script as described in the postface of his work *Shuōwén jiězì* 說文解字, where he introduced the term *jiǎjiè* 假借 as the sixth “category” of Chinese graphs. He uses this term mainly to explain the process of “loaning” a Chinese graph/character for a homophonous (or semi-homophonous) word which did not possess a written form yet. In his work he provides several examples, typically grammatical particles (*xūzì* 虛字) which derived their written forms from graphs representing a word with similar or identical pronunciation. For example, Xǔ Shèn indicates – without explicitly referring to the term – that *qí* 其 ‘basket’ (OC *gə; later written as *jī* 箕) was borrowed for the third person pronoun *qí* (OC *gə). In addition to semantically completely unrelated words sharing the same graph, however, Xǔ also cites some examples where the same graph is shared by words which are related to each other or derive from the same word family, e.g., *zhǎng* 長 (*traŋʔ) ‘grow; elder’ and *cháng* 長 (OC *Cə-[N]-traŋ) ‘long’; or the group *líng* 令 (OC *riŋ) ‘send’, *líng* (OC *riŋ-s) ‘issue a command’ (eventually also written as *mìng* 命 OC *m-riŋ-s). As such, Xǔ Shèn’s explanation of loaning is rather complex, including phonological and semantic features (e.g., semantic extensions).

Later scholars, when discussing this phenomenon, base themselves on Xǔ Shèn’s analysis, but also add their own ideas. For example, the Southern Táng scholar Xú Kǎi 徐鍇 (920–974) in his commentary on the *Shuōwén* was the first lexicographer to elaborate on Xǔ Shèn’s definition of *jiǎjiè*. Xú Kǎi does not seem to conceptually distinguish between a phonetic loan and a semantic extension and groups them, following Xǔ Shèn, under the category of *jiǎjiè*. However, practically, and based on his experience of actually *reading* medieval manuscripts he does recognize a difference between phonetic and semantic borrowings. For instance, he gives an example from the *Shān Hǎi jīng* 山海經 in which *jùn* 俊 (MC tswiH) is used to write *shùn* 舜 (MC sywiH) without identifying any semantic connection linking the two graphs. Despite this, Xú Kǎi still refers to these cases as “distant borrowing” and considers

them the result of the poor work of editors who lack proper understanding of the Chinese script. The later scholar Zhèng Qiáo 鄭樵 (1104-1162) discusses *jiǎjiè* primarily as a historical phenomenon, but at the same time, under a separate category, devotes some attention to phonetic substitutions that only make sense from the perspective of Medieval Chinese.

To what degree were these substitutions acceptable for him, and how do his views differ from those of earlier scholars?

The Medieval period was characterized by dramatic changes concerning the script, having to cope with significant changes in the phonological system of the spoken varieties of Chinese, the integration of thousands of new words of Indic origin, the appearance of countless new graphs and variants, the phenomenon of mass production of handwritten copies of texts, and – in the Late Medieval period – the appearance of the written vernacular. Although most medieval scholars were mainly concerned with phonetic borrowing and substitution as it occurred in older texts, manuscript evidence shows that phonetic substitution was an important feature of the writing system of the Medieval period. Although this highly complicated situation which made the reading of many text copies increasingly difficult was addressed by numerous scholars in their lexicographical works (most importantly, in the *yīnyì* 音義 and *zìyàng* 字樣 / *zìshū* 字書 genres), the phenomenon of phonetic substitution is only rarely directly dealt with, despite the compilers' great efforts in determining the correct character readings. In this panel we will also discuss medieval lexicographic works, such as the mid-Táng *Zhèngmíng yàolù* 正名要錄 (extant in Stein 388), which defines several categories of substitutions, constituting one of the earliest sources consciously reflecting on actual writing practices in Medieval Chinese manuscript culture.

What motivated scribes to use a specific phonetic substitution, often one sharing the same phonophoric, but sometimes also using seemingly completely unrelated graphs? Were these “choices” based on a conventionalized repertoire of replacements? Or sometimes created *ad hoc*? Or were they based on historical relations between graphs? And to what extent did “orality” (i.e., the copyist “hearing” the word while he is copying it) play a role?

The study of loan characters / graphs during the medieval period is likewise of great value for the research on the evolution of vernacular words and for our understanding of how colloquial lexical items and function words gradually emerged in written forms, and what “shapes” they assumed before they were committed to more or less standardized forms. The use of different characters for the colloquialisms of the respective contemporary versions of Chinese does often not only reflect the use of graphical variations but can also inform us about the change of the phonetic value of a word in a diachronic perspective (e.g., the various forms of writing the precursors of 什麼 between the Late Táng and Sòng dynasties). In addition, phonetic loan characters and “local homophonous readings” can be an invaluable source for the reconstruction of dialect readings, such as the Northwestern dialect of the Medieval period. As such, the question to what extent and by what means graphs were manipulated in order to be capable of “expressing speech sounds” will be repeatedly addressed.



In addition, a thorough investigation and understanding of the various phenomena of phonetic loaning is highly important for editing medieval Chinese manuscript texts and the choices an editor has to make in rendering variant graphs and loans. For example, editors of the Buddha biography preserved in Stein 3906 rendered the text passage to the left as 嚶玖從 (a phonetic rendering of the phrase *kjiat kjuwX trfūwŋ* 結九重 ‘composed of nine layers’). The “normalization” of the dialect graph 嚶 into the “standard” 嚶 is problematic here, since it “edits away” the phonophoric 繫 (MC *kjiat*) which is essential for the understanding of the phonetic substitution.