

动词起始概念	先秦义类		古梵语义类		先秦占比-古梵语占比
	个数	占比	个数	占比	
变化	794	0.17	330	0.19	0.02
接触	469	0.10	152	0.09	0.01
社会交互	426	0.09	175	0.10	0.01
行动	384	0.08	124	0.07	0.01
运动	276	0.06	103	0.06	0.00
其他	2325	0.50	1213	0.49	0.00
总计	4674	-	1764	-	-

表 14 先秦汉语和古梵语动词义类的分布对比

根据上述统计数据，我们可以发现：

(1) 无论是名词义类还是动词义类，在总量上先秦汉语都略高于古梵语。但先秦汉语和古梵语的名词、动词义类在大多数概念类别中的分布比例基本一致。这从语义分布的角度进一步表明两种语言的概念系统大致重合，概念系统具有普遍性。

(2) 先秦汉语和古梵语在分布上存在一定差异的名词类别有：动作、物质、认知、动物、状态、数量等，存在一定差异的动词类别有：变化、接触、社会交互、行动等。这一定程度地体现出了概念系统的特异性。

4.3 新增节点的异同

先秦汉语和古梵语词汇语义系统在词网的整体概貌之下，无法映射而新增的词更能反映这两种古老语言在词汇上的共性和差异。经过对新增节点进行补充以及更细致的处理，先秦词网目前共新增了 120 个大的节点，包含 1844 个词汇。古梵语新增了 22401 个小节点，都是名词节点，大都是一些专有名词，且没有对新增的义类进行进一步的语义类的划分。因此进行新增节点的对比时，我们通过梵语词网新增义类的英语释义，尝试将其划入某一类语义类，然后再与先秦词网进行对比。

下面以“祭祀”类义类和“神话人物”类义类进行简单分析。

	新增义类示例	词型示例
先秦汉语	祭祀名，古代因征战出师而祭天	禩
	祭祀用的酒肉	胙
	古代祭祀时献酒三次，即初献爵、亚献爵、终献爵，合称“三献”	三獻
古梵语	a sacrificial rite connected with liquids (与液体有关的祭祀仪式)	रसकर्मन्
	a religious rite preparatory to a sacrifice or any solemn observance (为祭祀或任何庄严仪式做准备的宗教仪式)	स्वस्तिवाचन
	name of a particular sacrificial text (特定祭文的名称)	यजुस्

表 15 新增“祭祀”类义类示例

表 15 中这些“祭祀”类下的义类和词汇可以很直接地说明先秦汉语词汇和古梵语词汇都存在着很多祭祀类的词汇。这些词汇在 PWN 找不到对应的词汇，因此都新增了很多这类节点。祭祀时往往有一定的规范，包括祭祀的活动、仪式、祭品、祭文、祭器等，古人希望通过这些来表示崇敬并求保佑，反映了人类初期对自然界、祖先或宗教的崇拜。

	新增义类示例	词型示例
先秦汉语	传说中的天神名	東皇太一
	传说中的水神名	玄冥
	传说中的山神名	颯氏
古梵语	name of a Buddha (佛教的创始人佛陀)	देवराज
	name of a Bodhisattva (菩萨)	सर्वसत्त्वप्रियदर्शन
	name of a Brāhman (婆罗门)	शम्पाक

表 16 新增“神话人物”类义类示例

表 16 中这些“神话人物”类下的义类和词汇可以很直接地说明先秦汉语词汇和古梵语词汇虽然都有大量的神话人物类词汇，但在具体人物和产生背景上仍存在着很大不同。当时的先秦人民主要是想象存在着某一类掌管具体事物的神仙，从而掌管天地的运行，大多都是中国人民在当时不能解释很多现象的情况下想象出来的产物，比如“玄冥”为传说中的水神。而古印度是以宗教为中心的神话之邦，婆罗门教（印度教的前身）、佛教和耆那教先后兴起，让人们在不能理解世界上的各种未知事物时能够寻找安慰，在古梵语词网中，新增了很多这类节点，大多都是印度社会里被奉为神的宗教人物，例如佛教的创始人佛陀 Buddha (देवराज)。虽然崇拜的神有所不同，但这都说明了古代人民对自然的敬畏和崇拜。

综上，可以看出，通过对先秦词网和古梵语词网的初步的跨语言对比，我们可以大致探究这两种古老语言的词汇语义系统的异同。如果能够对两者进行更详细的对比分析的话，或者与更多建构了词网的语言进行跨语言对比的话，将会得到更多的词汇语义映射情况和词汇语义关系对比等方面的信息。

5 结语

先秦词网能够展示先秦词汇语义系统的整体面貌，满足古汉语信息处理和跨语言对比的研究需要，为探究先秦特色语义提供新的资源和方式。本文借助国际通用的词网框架，以《汉语大词典》中的先秦部分和 CCD 为基础资源，进一步在现已构建的先秦词网的基础上进行了校正，完善了先秦词网的构建体系，构建一个准确率和覆盖率更高的先秦词网，涵盖 43591 个词语、61227 个义项、17975 个义类。另外，本文还进一步从跨语言对比的角度出发，将先秦词网和古梵语词网进行对比，探讨了这两种古老语言的词汇语义系统的异同，进一步证明了先秦词网这一词汇数据库的应用价值。

参考文献

- [1] Tufis D, Cristea D, Stamou S. BalkaNet: Aims, methods, results and perspectives. a general overview[J]. Romanian Journal of Information science and technology, 2004, 7(1-2): 9-43.
- [2] Miller G A. WordNet: a lexical database for English[J]. Communications of the ACM, 1995, 38(11): 39-41.
- [3] Vossen P. Euro WordNet: general document[J]. 2002.
- [4] Zhang Y, Li B, Dai X, et al. PQAC-WN: constructing a WordNet for Pre-Qin ancient Chinese[J]. Language Resources and Evaluation, 2017, 51(2): 525-545.
- [5] Li M, Zhou Z, Wang Y. Multi-Fusion Chinese WordNet (MCW): Compound of Machine Learning and Manual Correction[J]. arXiv preprint arXiv:2002.01761, 2020.
- [6] Bizzoni Y, Boschetti F, Diakoff H, et al. The Making of Ancient Greek WordNet [C]//LREC. 2014, 2014: 1140-1147.
- [7] Zhang Y, Li B, Wang X, et al. Mapping word senses of middle ancient Chinese to WordNet [C]//2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT). IEEE, 2014, 1: 446-450.
- [8] ZHANG L, LI J, HU M, et al. Implementation of Chinese WordNet [J]. Journal of Northeastern University (Natural Science) vol, 2003, 24.

- [9] Huang C R. Sinica BOW: integrating bilingual WordNet and SUMO ontology[C]//International Conference on Natural Language Processing and Knowledge Engineering, 2003. Proceedings. 2003. IEEE, 2003: 825-826.
- [10] Xu R, Gao Z, Pan Y, et al. An integrated approach for automatic construction of bilingual Chinese-English WordNet [C]//Asian Semantic Web Conference. Springer, Berlin, Heidelberg, 2008: 302-314.
- [11] Huang C R, Hsieh S K, Hong J F, et al. Chinese WordNet: Design, implementation, and application of an infrastructure for cross-lingual knowledge processing[J]. Journal of Chinese Information Processing, 2010, 24(2): 14-23.
- [12] Wang S, Bond F. Building the chinese open WordNet (cow): Starting from core synsets[C]//Proceedings of the 11th Workshop on Asian Language Resources. 2013: 10-18.
- [13] Xu H, Chen S, Cai J, Cao L, Wan C, Li B. 2020. The Construction and Statistical Analysis of Pre-Qin Ancient Chinese WordNet [J]. International Journal of Knowledge and Language Processing, 2021, 11(3): 48-61.
- [14] Minozzi S. 2017. Latin WordNet, una rete di conoscenza semantica per il latino e alcune ipotesi di utilizzo nel campo dell'Information Retrieval. In Paolo Mastandrea, editor, Strumenti digitali e collaborativi per le Scienze dell'Antichità, number 14 in Antichistica, pages 123–134.
- [15] 王力. 汉语史稿.中册[M]. 科学出版社, 1958.
- [16] 杨合鸣. 诗经词典[M].湖北: 辞书出版社,2012.
- [17] 张颖杰.资源稀缺语言的词汇语义资源自动构建方法研究[D].南京大学,2017.
- [18] 石民,李斌,陈小荷.基于 CRF 的先秦汉语分词标注一体化研究[J].中文信息学报,2010, 2:39–45.
- [19] 张俐,李晶皎,胡明涵,姚天顺.中文词网的研究及实现[J].东北大学学报,2003(04):327-329.
- [20] 于江生,俞士汶.中文概念词典的结构[J].中文信息学报,2002(04):12-20+44.
- [21] 姚天顺,张俐,高竹.词网综述[J].语言文字应用,2001(01):27-32.
- [22] 吴思颖,吴扬扬.基于中文词网的中英文词语相似度计算[J].郑州大学学报(理学版),2010,42(02):66-69.
- [23] 袁毓林.面向信息检索系统的语义资源规划[J].语言科学,2008(01):1-11.
- [24] 郭锡良.先秦汉语名词、动词、形容词的发展[J].中国语文,2000(03):195-204+286.