

概念史研究的词汇史视角与数字资源

沈 国 威

The Perspective of Lexical History and Digital Resources in the Study of Conceptual History

SHEN Guowei

概念史の研究は、語彙史研究の知見を基礎にしている。一方、語彙史の研究は、それが近代新語・訳語の場合、概念史による方向付けがなければ、用例の羅列に陥りがちである。概念は言語によって具現し、言語によって伝播し、普及する。東アジアの概念史は、他でもなく漢字によって西洋の近代的新概念を受容する歴史である。本稿は、概念史研究と語彙史研究の相関性と方法論における相違を日中言語接触と語彙交渉の視点から論じている。コーパスの有効性に関してはこれまでに初出例の発見、意味用法の変遷を記述することに限られていたが、使用頻度の変化を捉えることも語彙史の研究にせよ、概念史の研究にせよ非常に重要である。デジタルリソースもそのために今までと異なる機能を備えなければならないと指摘している。

キーワード：概念、観念、概念範疇、語彙化

1. 概念史研究与词汇史研究

近年来，概念史的研究引起了极大的关注。不同领域的学者在概念史研究的名义下聚集，协力合作，取得了可喜的进展。反映最新研究成果的刊物有《东亚观念史集刊》（台湾政治大学）、《亚洲概念史研究》（南京大学）等，专著有《观念史研究》（金观涛·刘青峰，香港中文大学 2008）、《严复与科学》（沈国威，凤凰出版社 2017）、《东往东来：近代中日之间的语词概念》（陈力卫，社会科学文献出版社 2019）、《什么是概念史》（方维规，三联书店 2020）等。学衡尔雅文库计划出版概念史丛书，笔者刚刚脱稿的《科学》对 science，从译词形成和概念引介的角度做了讨论，特别对中日的先行者西周与严复的情况进行了考察。所谓概念史，究其实质而言就是近代概念形成史，是东亚如何用汉字接受西方的新概念，并建构大同小异的各自的近代意识形态体系的历史。什么是概念？人的感官感受自然界的森罗万象，在头脑中形成影像（image），我们即使闭上眼睛，也有残像浮现在脑海里。相似的，或同类的事物的影像——世界上是不存在绝对相同的事和物的，所以影像也不完全相同——多次重叠，轮廓化、抽象化。这样的重叠影像我们暂且称之为“概念范畴”（就是概念的轮廓），概念形成的过程就是外部世界在我们的头脑里的“影像化”的过程。

概念范畴无论在外部还是内部都没有清晰的边界，存在着一个连续的灰色区域。我们一方面

需要把某一个概念范畴同其他概念范畴区别开来，如“科学”与“宗教”不同；另一方面也需要把聚集在某一概念范畴内的相似的，或同类的事物加以区分，如在人类的知识这一范畴里，“科学”与“文学”“艺术”“技术”“美术”等不同。概念分类的任务是由语言来实现的。语言对外部世界进行切割，其结果是在概念范畴外部和内部确立了边界，这一过程被称之为“范畴化”。范畴化是一种命名行为，人们用固定的声音指代抽象化了的影像，使概念范畴和特定的语音形式发生了关联。人们在使用语言进行交流时，说者发出一个代表概念的声音，听者受到声音的刺激，在自己的头脑里再现影像，理解说者所欲表达的概念。命名的好处是，一、可以指称对象物，并将其从同类的事物中区别开来。例如我们称一种知识体系为“科学”，那么“科学”就不再是“文学”或“艺术”；二、将对象物加以抽象化。例如实际上有各种各样的杯子，我们甚至可以说，世界上没有两个完全一样的杯子，但是各式各样杯子之间的差别在“杯子”这一名称下都被舍像了。王国维说“事物之无名者，实不便于吾人之思索”点明了命名的重要性。¹⁾

我们还常常听到另一个术语“观念”。“观念”和“概念”有何不同？这两个术语都是和译词（“观念”是被改变了词义的佛经词），在早期的英和字典里分别对应 idea, concept, 但在现在的英和辞典里两者的区别并不分明，在汉语里这两个术语基本上袭用了日语里的用法。而在实际的翻译活动中，例如严复几乎不用“概念”，只用“观念”。王国维在评论日人之译语未必皆精确者也时说，译 idea 为“观念”，观念者，谓直观之事物。其物既去，而其象留于心者，这时还“谓之观，亦有未妥”，然而 idea 本身也有这种错误，这个词在古希腊语里是“观”的意思，源自五官，故称为“观”；所“观”的物体消失后，影像还存在，就称之为“念”。至于把 conception 译为“概念”，如果用荀子的词，可以称之为“共名”（《论新学语之输入》）。王国维在这里说的是译词的“理据”问题，理据即“物之所以名”。²⁾ 使用已有的语言成分创制新词必然发生“理据”问题，理据反映了造词者（通常是时代的先行者、启蒙家）的在理解、接受域外新概念时的思维方式。概念史研究常常通过对理据的分析复原观念化的某些过程；另一方面，词汇研究认为理据固然重要，但是理据的合理与否并不决定一个词的存亡。例如，严复在《天演论》中将 evolution 译为“天演”。尽管“天演”的立意（即“理据”）极好，但最终为日本译词“进化”所取代。³⁾

一个外来概念被译介的当初，常常呈现百家争鸣的状态，当时的人们用“混乱”、“不统一”来描述这种状态。混乱的译词在语言社会或者政治力量的筛选下，最终会有两种结局，一是定于一尊，另一是各有所司。“科学”取代“格致、格物”是定于一尊的结果；而“概念”和“观念”则可以看作是各有所司的例子，即“概念”可以指称任何概念，不论大小、具体抽象，是上位词；“观念”指称经过意识形态化的概念，只有那些得以编入某一语言社会意识形态体系的概念才能成为观念，所以“观念”是下位词。⁴⁾ 意识形态化的过程姑且称之为概念的“观念化”。

1) 王国维：《论新学语之输入》，《教育世界》第四期（总第96号），1905年2月。

2) 索绪尔以后的现代语言学的一条基本定理是除了某些拟声拟态词以外，单纯词能指与所指的关系是任意的。但复合词则不可避免地存在理据。

3) 在借入日本译名的世纪之交，日语的现代词汇体系几近完成，就是说，中国得到的是日本意义系统整合后的结果，某些译词即使在日语里也不再是创制初始的意义了。

4) 概念即语言所包含的意义内容，与语言的形式相对应。形式与内容的关系即“能指”与“所指”的关系，是

2. 词汇史研究与近代关键词

“概念是代号，是思想的出口”，⁵⁾而词语则是概念的外壳；没有概念不足以形成思想，概念借助于语言而具形，没有语词则无法表达概念。故概念的传播与接受需要借助于词语完成。外来的新概念可以用一个说明性或比喻性的词组、一句话，甚至一篇文章来表达，也可以用一个词来表达。用一个词（包括使用既有词和新造词）表达一个概念叫“词化”，没有词化就没有观念化。现代语言学的一个基本观点是：语言无优劣之分，不管是科学技术高度发达的社会所使用的语言，还是原始农耕社会所使用的语言，只要有必要，就可以表达任何一种概念，但并不是所有的概念都能实现词化。⁶⁾在引入域外新概念时，能否发生词化要受到各种因素的左右。一般来说，在接受社会里出现频率高的概念比较容易词化，否则将停留在词组和短语等说明性（非命名性）表达的层面上。梳理19世纪英华字典和汉译西书的历史，可以发现，外来的新概念在引介初期常常采取词组或短语的形式来表达；词组、短语常常在反复使用中逐渐凝缩成一个词，最后完成词化。需要强调的是近代社会的核心概念，其词化常常是强制性的，这就是所谓的近代关键词。概念的词化提供了表达上的便利性，也蕴含着概念异化，即名实乖离的危险。我们用一个词指涉一个概念时，常常忽视了概念的真正内涵及其潜移默化的演变。被观念化（意识形态化）的语词尤其如此，人们按照想当然的“观念”行事，而不是按照思想或严格定义的概念去把握事实。“个人”、“自由”、“革命”等观念在中国语境中的异变引起了研究者强烈的学术兴趣；“个人、自由、革命”之所以成为独具中国特色的关键词不在于它们是 individual, freedom, revolution 的译词，而在于这些术语在中国意识形态体系中的定位，这种定位是以与其他词语的搭配组合关系体现出来的。

在中国，西方概念的容受可分为三个阶段，第一阶段，19世纪初叶至中日甲午战争；第二阶段，1895~1915年；第三阶段，五四新文化运动时期。这三个阶段同时也是近代关键词创制或借入、普及、定型的过程。第一阶段是传教士造词，并在有限范围内传播的时期；第二阶段是日本译名大量涌入汉语的时期；第三阶段是对在此之前存在的新词、译词进行整合，使其融入汉语词汇体系的时期。自马礼逊1807年登陆广州，新知识的引介就面对一个译词创制的问题。但在翻译内容、造词方法上，第一阶段都不能说取得了成功。第二阶段，甲午之役，老大帝国败给蕞尔岛国，亡国亡种的危机加深。而在此之前扮演传播西学主角的传教士淡出中国政治舞台，⁷⁾虽有严复的孤军奋战终不能满足中国吸收西方新知识的迫切需要，国人只好将目光转向日本。留学日本、翻译日文书籍的热潮帮助汉语从日语获得了大量的新词译词，并由此迅速完成了现代汉语词汇体系的重构，进而实现了书面语的言文一致。汉语不但从日语接受了新词，还刷新了旧词

语言符号系统研究的对象，故语言研究更多地使用“概念”这一术语。

5) 方维规：《概念史研究方法要旨》，《新史学》，北京：中华书局，2010年，第3卷，第3~20页。

6) 自然语言并不为所有的概念都准备一个词，如汉语的“兄、弟、姐、妹”等概念在英语中就是用词组形式表达的。

7) 柯文（Paul A. Cohen）。费正清编《剑桥中国晚清史》上卷，中国社会科学出版社，1985年，第634页。柯文列举的第一个原因是传教士本身逐渐与政治问题拉开距离。亦参见任达著《新改革与日本——中国1898~1912》，李仲贤译，江苏人民出版社，1998年，第12页。

词义，近代关键词在第二阶段与原词更相吻合的现象正是日语影响的结果。第三阶段是对新出现的词汇成分（自造的或外来的）进行整合，使其融入汉语词汇体系的时期。大量语词被观念化成为近代关键词就是在被称为第三阶段的五四期间及其以后的一段时间内发生并完成的。

在近代研究史研究中，一个特别需要注意的问题是，近代概念在原生地，即西方也是逐渐发生、发展、定型下来的。这是重大人文概念往往是由综合向分析发展的自然属性所决定的。恰如严复所说“翕以合质，辟以出力，始简易而终杂糅”（《天演论》自序），他又用“由简入繁”、“物变所趋，皆由简入繁，由微生著”（《天演论·导言二》）等言辞加以描述。译词的混乱正是反映了近代概念形成的不同阶段、不同的信息来源、不同的转译者的理解、表述上的差异以及误解所造成的滞后、重叠、模糊等现象。例如，传教士山雅各（J. Sadler）1904年翻译出版了《哲学源流考》，原著为 *The Pioneers of Science*. 1893, by Sir Oliver Lodge, 内容是自然哲学。而在1904年“哲学”已经作为 philosophy 的译词普及定型，山雅各的用法不可避免地引起了混乱。

“文学”作为与“科学”紧密相关的概念也是一个饶有兴味的事例。人类的知识（knowledge; philosophy）可以分为文字性和非文字性两大类。Literature 是存在于文字中的知识和想象力的全部成果，在相当长的历史阶段，这是一个无所不包的概念。但是19世纪中叶，在日常的狭义用法上，将 literature 分为：文学和实证科学，前者是 polite literature（纯文学）或法语 belles-lettres（纯文学）的同义词，后者即 science。另一方面，非文字的 art 是运用知识或力量达到预期目的的知识领域，二分为精致（fine art）和实用（useful art）两类。19世纪末 useful 向 science 靠拢，最终确立了 technique 的概念范畴。1842年出版的 *A Dictionary of Science, Literature, and Art* 就反映了人们对概念分化现象的描述，而在19世纪末的 *Webster's International Dictionary* 1898上，literature 的同义词还依旧包括了 science; learning; erudition; belles-lettres 等。

作为向中文世界引介的努力，英国传教士艾约瑟（Joseph Edkins, 1823-1905）在《六合丛谈》创刊号上（1857.1）发表了《希腊为西国文学之祖》（Greek the stem of western literature）一文，这是艾氏介绍西方文学知识的一系列文章“西学说”（Western Literature）的第一篇，其“文学”已经不再包括实证科学的内容了。但是美国林乐知（Y. J. Allen, 1836-1907）在翻译森有礼的英文著作 *Education in Japan*, 1872时，惯性地使用了“文学”的旧的意义，译为《文学兴国策》（1896）。王国维说：“近年文学上有一最著之现象，则新语之输入是已。”（《论新学语之输入》1905）其中的“文学”则包括了人文科学的全部。

在日本，《英和对译袖珍辞书》中“文学”= grammar，《附音插图 英和字汇》（1873）中，“文学”是 belles-lettres; humanity; literature 的译词，“文学”= literature 与罗存德的《英华字典》（1866-1869）相同，belles-lettres; humanity 源于何处待考。有过暂短在华传教经历的美国长老会医疗传教士，平文（J. C. Hepburn, 1815-1911），在其《和英语林集成》中所展示的“文学”对译情况如下：

词条	第一版 1867	再版 1872	第三版 1886
文学	Learning to read, pursuing literary studies, especially the Chinese classics.	同左	Literature; literary studies, especially the Chinese classics.
Literature	学问，文，文道	同左	学问，文，文道， <u>文学</u>

“文学”=literature 对译关系的建立是在《和英语林字汇》第三版（1886），这似乎是井上哲次郎《哲学字汇》（1881）中“Literature = 文学”的影响。西周在《百学连环》（1871-1873）中讨论了 literature 与“学”与“术”的关系，介绍了 belles-lettres; humanity 的概念，显示了与欧洲同步的理解，但西周的“文学”是文字、语言的意思，literature 译为“文章”。“文学”最终通过井上哲次郎的《哲学字汇》（1881）普及到整个日本社会。但是，在语文辞典上，如下所示，“文学”=“学问”的解释一直持续到 20 世纪初。

《言海》（1891）：①通过读书考究的学艺，即经史诗文之学（对于武术而言）；②语言学、修辞学、逻辑学、历史学等学之总称。

《日本大辞书》（1893）：①利用书籍进行研究的学艺，具体指与武术相对的普通知识。②语言学、修辞学、历史学等无形（抽象？）之学的总称。③英语 Literature 的译词，所有具有相应知识的普通人，使用他们所能够理解的言辞，给出最普通的观念，最能发挥宇宙真善美的学问。据此而言，语言学、修辞学、逻辑学、历史学、哲学、诗学等，这些以前曾被视作文学的学科，根据上述定义，逐渐脱落，仅留下了“诗学”。

《日本新辞林》（1897）：①利用书籍进行研究的学艺，指与武术相对的一般性学问。②语言学、修辞学、哲学、历史学等学之总称。

《ことばの泉》（1901）：①利用书籍进行研究的学术，与武术相对。②运用所有具有普通知识的人都能理解的文辞，为人们带来精神上快乐的学问，诗学、修辞学等。与科学相对。

我们发现在上述辞典的释义中“小说”还没有出现。在这一时期的各类辞典中“技术”一词也和今日的意义用法有较大的距离。“技术”“艺术”原为汉语古典词，在英华字典中对译 art；“美术”是日本译词，也用来对译 art。在《英华大辞典》（1908）里，尽管还有重叠之处，但大致形成了 literature = 文学；science = 科学；technique = 技术；fine arts = 艺术、美术的对译关系。但是，如下所示，在转译英文释义时，还有不尽人意之处。

Literature, In a special sense, that body of literary compositions which, to the exclusion of merely philosophical, scientific, and technical works, are occupied mainly with that which is spiritual in its nature and imaginative in its form, whether in the world of fact or the world of fiction, 文学、文章（特別意義、除哲理及科学外、凡神靈思想爲其資料、離奇變幻爲其形式、或實記或杜撰者、皆文学也）。而在今天可能会译为：从某种特殊意义上说，纯粹的哲学、科学和技术性的作品不是文学，无论是在现实的世界还是虚构的世界，文学的本质是精神性的，其形式是富有想象力的。

新语词的加入必然促成汉语意义体系的重组，语言社会赋予语词以联想、语体、评价等周边义以及意识形态的价值指向。语词具有体系性，即作为概念的名称的“词”不是孤立存在的，而是与其他的词语保持着这样或那样的关系，织成一个意义网络。任何一个词的出现消亡或意义用法的变化，都会引起该词汇体系内同一语义场中其他词语的变动，可谓牵一发而动全身。处于意

义体系中心位置的关键词尤其如此。正是由于这一原因，我们的考察不能仅仅针对个别语词孤立进行，而必须兼顾同一概念范畴中的其他成员。例如，在讨论“民主”观念时同时将“民权、民治、民政、共和、立宪、德谟克拉西”等纳入视野；“经济”与“经世、计学、富强、生计、生产力”等相关联。只有在特定的概念网络中才能准确地观察关键词的诞生、普及、定型。所以关于“科学”的讨论也不能和“文学、技术、艺术、美术”等词语分割开来进行。

3. 词汇史研究法

由上可知，近代概念史的研究与词汇史研究，密不可分，或者说后者是前者不可或缺的基础。

任何一种语言的词汇系统都具有自我调整的机制，可以容受外来概念，重构原有的意义体系。这一切所需要的只是时间。在中国第一阶段和第二阶段初期本土译词如果缓慢地接受洗礼，其结局也会更加自然。但是在世纪之交，日本译词决堤似的涌入汉语，中断了本土译词自然演化的进程。从世纪之初的各科学术到 20 年代的社会主义、共产主义思想体系，文艺戏剧理论，可以说整个近代知识体系的建构都与日本知识有着深刻的关联，即西学从东方来，马列主义从东方来。考察 1895~1915 年这一阶段的中国概念形成历史，日本知识是一个无法回避的问题。最大限度地利用日本的研究资源，包括相关的历史文献及研究成果会使我们的视角更加全面，这在分析比较东亚各国近代化不同的进程时（共性与个性），尤为重要。因为这不仅仅是一个译词创制、借贷的词汇史的问题，而是对东亚近代史的整体描述。

日本的近代词汇研究有着深厚的学殖。近代词汇体系的建立为始于文学领域的言文一致运动提供了坚实的基础。明治时代结束（1911）后不久，日本的学者们就开始反思以新词译词为主的近代词汇形成过程中的各种问题了。日本语言学界的所谓「国語史研究」就是关于日语的历时研究，「漢語（かんご）」一词在日语里意为“汉民族的语言”。⁸⁾ 明治维新前后，随着大量汉字新词译词的增加与新知识的普及，「漢語」这一术语又被用来指称那些如汉字熟语和读汉字音的词等在日语中具有外来的，即来自中国的特征的词语。在日本第一部近代国语辞典《言海》中，「漢語」明确地作为「和語」即日语固有词的对立项，成为日语词汇体系中的一个重要的组成部分。⁹⁾ 而第一次把「漢語」作为研究对象展示在我们面前的是日本国语学家山田孝雄的《国語の中に於ける漢語の研究》（=日语中的汉字词之研究，1940）。¹⁰⁾ 在对日语中的汉字词使用状况做了分析之

8) 如前所述，日语「語」是 word 的译词，「漢語」即使用汉字的词语。一般人的「漢語」意识是：用汉字书写，读汉字音（即日语接受古汉语发音后形成的发音体系），而对造词者的因素则不作特殊的考虑。作为日语研究术语的「漢語」，其严格的定义涉及到许多具体的语言事实，这里不做进一步的展开。有兴趣的读者可参阅陈力卫《和製漢語の形成とその展開》，东京：汲古书院，2001 年。

9) 《言海》卷末的「採収語類別表」中将词条分为“和语”、“汉语（即汉字词）”、“外来语”等三大类，并分别做出了详细的数量统计：和语 55.8%，汉语 34.7%，其它 9.5%。这个统计显示了明治 20 年代日语词汇的构成情况。

10) 本书是山田孝雄 1931 年（昭和六年）在东北帝国大学（今东北大学）的讲义录。

后，山田指出：现代日语如果排除汉字词，不论是日常生活中的寒暄，还是社会成员于公私之间的思想交流，都将陷于不可能之状态（4页）。在第一章“序说”中，山田对《言海》的编纂者将汉字词与其它来自欧美语言的外来词区别对待的作法表示赞同。他认为汉字词不仅仅停留在“外来词”或“借用语”的层面，已经深深嵌入日语语言组织的内部；汉字词已经失去了外国语言的特色，与纯粹的固有词语并无不同之处，大多数国民对汉字词也不再抱有外国语的意识了。¹¹⁾ 山田指出：从数量上看汉字词占日语词汇的将近一半，尤其是在表达抽象意义的领域比重更大；因此研究汉字词就是研究日语的另一半。尽管在言文一致等语言问题上，山田具有强烈的保守和国粹主义倾向，¹²⁾ 但是对日语中汉字词存在的事实做了客观的描述，是第一位将汉字词作为日语研究对象的国语学学者。山田在他的著作中讨论了：汉字词流入的历史、汉字词的认定范围、汉字词的特点、汉字词的词形、汉字词的渊源、汉字词在日语中的现状以及汉字词对日语的影响等问题。例如，在第六章“源流之观察”中，山田专门考察了因西学翻译而产生的汉字词之后指出：日语中的汉字词来源于，一、中国的典籍；二、佛经翻译；三、汉译西书。尤其是第三点首次由山田明确点出。在山田以后，「漢語」遂作为国语学上的术语被广泛使用。汉字词的研究也成了日语词汇研究中的一个重要方面。

山田的著作首先提出了汉字词的问题，对汉字词的历史，现状做了细致的描述。而关于近代汉字译词新词的具体研究则多开始于第二次世界大战之后。广田荣太郎的《近代译语考》（1969）是一本论文集，该书收入了“恋爱、蜜月、新婚旅行、接吻、悲剧、喜剧、世纪、常识、俱乐部、冒险、探险”等词语的考证论文。广田所利用的文献资料包括：报纸、杂志、各类翻译书籍、英汉辞典、英和辞典等。他所采用的是典型的国语学词源考证的方法，即，寻找发现首见书证，考察新词由出现到普及、定型的全过程。广田是较早有系统地利用英华辞典考证近代汉字新词的国语学家。

与广田不同，东北大学教授佐藤喜代治在他的《国語語彙の歴史的研究》（1971），《日本の漢語》（1979）等著作中，主要通过对比江户时期等近代以前的日本知识阶层，如吉田松阴、桥本左内等人著述的考察，探索日语接受汉字词的历史。佐藤致力于阐明江户时代汉字词的使用与现代日语词汇之间的关系，在个别汉字词语源记述上做了很多开拓性的工作。

自山田以后，汉字、汉字词的研究从传统的汉学家之手转到了国语学家之手，成为国语学研究的一个重要方面。以近代新词为例，日本学界对汉字新词的产生、接受、普及、定型等的研究成果较多，而对日制汉字词的输出关注较少。另外，汉字词研究需要处理大量汉语典籍，在汉学教育相对削弱的今天，研究的难度无疑是增大了。

11) 山田将外来成分按照融入自语言的程度分为：纯粹的外语、狭义的外来语、借用语、归化语四类。借用语和归化语的区别在于能否实现某种形态变化。

12) 李妍淑《「国語」という思想》，岩波书店，1996年；子安宣邦《汉字论》，岩波书店，2003年等。如子安所指出：使用如“侵入，归化”等一系列与国家有关的比喻来谈论汉字词是山田的一个特点。

4. 概念史研究与数字资源

基于词汇史的概念史研究始于某些词语在文献典籍中的书证，然而既有的工具书还不足以胜任提供首例书证的要求，故传统的皓首穷经式的书证查找成了重要的前期准备工作，概念史研究也因此受到了极大的限制。进入 21 世纪以后，随着电子计算机性能的迅速提高，数字资源的整备和使用也越来越广泛，极大地推动了词汇史、概念史的研究。在今天查找某一范围内的书证成了举手之劳，研究者是在拥有大量例证的情况下开始论证自己的假说，他们已经无法想象在没有数字资源的条件下做研究了。提供词语检索的主要数字资源有：日本国立国语研究所研制、开发的《杂志太阳语料库》（「太陽コーパス」2005）。¹³⁾ 内田庆市等研制的关西大学“近代汉语文献语料库”，该语料库收录了大量西人汉语研究的文献、教科书，并以图片和文字数据对应的方式展现，保证了检索结果的可确认性。特别值得一提的是台湾“中央研究院”近代史研究所研制的《英华字典资料库》，该资料库收录 1815 年至 1919 年间极具代表性的早期英华字典 24 种，其中全文输入 14 种（含 11 种英华字典、3 种华英字典），影像浏览 10 种，据此可以大致把握近代译词发生及定型的历史。中国近年，商业汉语典籍的数字化的势头迅猛，古典有《中国基本古籍库》《四库全书数据库》，近代媒体数字资源有《申报》（1872-1949）、《大公报》（1902-1949）、《东方杂志》（1903-1949）以及一系列收录了 19 世纪末 20 世纪初期报刊书籍的语料库。但是耶稣会士和新教传教士的译著，即前期汉译西书、后期汉译西书，还没有专项的电子数字资源。这两类文献近期有大量活字本出版，可以自作 PDF 语料库。

数字资源的另一个优点是足不出户就可以迅速发现，即时阅览世界各地图书馆所藏的珍贵文献资料。上述的 literature 的概念演变，可以在美国档案馆（Internet Archive）里找到 20 世纪上半叶为止的数十种重要英语辞典、百科全书，在日本国会图书馆可以下载主要国语辞典和百科全书。一系列辞典的释义变化勾勒了 literature 概念的形成与定型。提供珍贵文献的著名网站还有，台湾中央研究院“数位典藏资源网”、日本国立国会图书馆近代数字馆（近代デジタルライブラリー），东京大学、早稻田大学、庆应义塾大学、京都大学、关西大学等学术机关的图书馆也都有可免费下载的文献典籍资源可以利用。

数字资源可以帮助我们在相当大的范围内迅速地找到研究对象的首例书证，意义用法的变迁也可以据此详细描述。汉语由于自身的特点（不像西方语言那样分词），收录汉语历代文献的语料库常常只是一个未做语词切分和词性标注处理的数据库。这样的语料库在语言研究上的作用是有限的，¹⁴⁾ 在检索某一范围的语料中是否存在某一字符串时，检索结果中不可避免地会混入无效的结果。因此，对检索结果的后期整理非常重要。

概念史研究和词汇史研究都需要新的研究方法。使语料库和观念史研究深度结合并获得丰硕研究成果的是台湾政治大学金观涛、刘青峰的研究团队。金观涛等建构的《中国近现代思想史专

13) 这个语料库只是间隔地收录了 1895-1925 年间杂志《太阳》的语料。现在这个语料库扩充为《日语历史语料库》（「日本語歴史コーパス」2016），《明六杂志》（1874-1875）《国民之友》（1887-1888）等也成为收录对象。

14) 只能提供包含某些字符串的例句。单音节成分的检索由于结果数量太大，常常不具实际意义。

业数据库（1830-1930）》共收录了1亿2千万字节，主要文献群跨度100余年。这个数据库按照近代思想史研究的脉络以穷尽的方式收录相关文献，并从专家的视角对收录文献做了最适当的选择。研究者可以该数据库中检索到某些关键词的首例书证（结合大量古典文献的语料库，初始书证的确认可以更精确），并追踪其后的历史演变。金观涛·刘青峰将其研究方法归结为：提出一个思想史上的假说，然后根据语词使用的情况去证明；或者相反，从语词使用的变动情况去解读具体的历史事件。¹⁵⁾与以往的研究不同，他们试图通过对关键词的考察勾勒近代思想变迁的脉络，并“采用数据挖掘（data mining）的方法，把表达某一观念所用过的一切关键词都找出来，再通过核心关键词的意义统计分析来揭示观念的起源和演变”。¹⁶⁾金刘认为通过这种方法，“观念史就可以从思想史研究中分离出来，成为思想史研究的经验基础”。¹⁷⁾

然而需要注意的是，有很多动词、形容词在中国的典籍或汉译西书中可以找到用例，在词义上，也具有古今一贯性，即词源上并非“现代词汇”。尽管有较早的书证，但是使用频率一直不高，在某一概念范畴中是不折不扣的周边词汇。但这些词在19世纪末突然开始活跃起来了。¹⁸⁾这也就是说，对这些词首例书证几乎没有什么意义。与初始书证相比，使用频率变化的契机与时期才是应该格外加以关注的。传统的词源调查的方法，对于使用频率的把握无能为力，但是，近年迅速发展的语料库以及大数据研究法提供了新的可能性。大规模语料库的出现，不但可以发现首见书证，还可以捕捉使用频率的变化。随着更多的语料库的建立和扩充，词频调查也会更加准确。词汇使用频率会受到突发性话题、事件的影响，社会的流行也会引起某些词的频率剧增，尤其是人名、事件名、物品名等名词具有较强的流行词特征，相比之下，动词、形容词的频率增加，更主要的是崭新的表达方式的影响。

5. 结语

数字资源的出现彻底改变了词汇史、概念史研究的原风景，同时数字资源环境下的研究又展示了新的可能性。中日近代词汇交流史的研究已经取得了可喜的成果，¹⁹⁾那么中日比较概念史研究是否可行？东亚近代化的进程引发了西方新概念的容受问题，在汉字文化圈内，长期以来“汉字”是概念的唯一的外壳（载体），以至于近代以降，汉字以及汉字构成的新词、译词是容受、表达西方新概念的唯形式。词汇史、概念史的实质成为：“东方如何用汉字容受西方的新概念？”特别是19世纪以来的中日之间的文化交流和语言接触使西方新概念的容受及其“词化”

15) 金观涛·刘青峰著《观念史研究》，第253页。

16) 但是，从“一切关键词”向“核心关键词”的收敛过程及其动机是观念史的问题还是词汇史的问题？金刘似乎倾向于前者，而笔者主张这毋宁是词汇史的现象。

17) 金观涛·刘青峰著《观念史研究》，第3页。不仅是思想史，观念史还将成为近代诸（人文系）学术史研究的“经验基础”。

18) 现代日语中常用的「サ変動詞語幹」、「形容動詞」很多在《言海》（1888-1891）等19世纪的日语辞典中都没有收录。其时这些词还未能成为一般词汇。

19) 沈国威：《新语往还——中日近代语言交涉史》，北京：社会科学文献出版社，2020年。

的过程成为跨语言的事件。因此，无论是词汇史的研究还是概念史的研究都无法在汉语或日语等单一语言内完成，而这同时也使“概念形成史比较研究”成为可能。例如，日本的西周和中国的严复在引介、容受 science 这一概念的过程都发挥了极其重要的作用，处于不同语境中的二人甚至有着惊人的相似之处。源泉在于西方，二人存在着某种相同之处并不奇怪，但是同样来自于西方的“科学”在中国和日本是否走了相同的观念化道路？为什么会这样？则应该是思想史研究的重要内容。

新的研究目标必将对数字资源的环境提出新的要求。笔者认为以下 3 点需要集中力量加以解决。

1. 可检索文献群的扩充，包括文献数量和覆盖面；
2. 增加使用频率显示功能，这是考察概念普及、渗透不可缺少的功能；
3. 特殊数字资源的整备，例如“西周语料库”等个人文献群、“《东洋学艺杂志》语料库”等报刊杂志全文检索语料库等。

笔者期待着下一个十年，数字资源的质与量的飞跃。