

数字人文视野中的 文本与文本性问题反思

陈 静

[提要]作为 20 世纪文艺理论中两个颇为核心的概念,文本与文本性从内涵和外延上都具有着异常的丰富性和开放性。无论是从语言学讨论作为语言人工物的文本,从新批评的角度来讨论作品与文本、文学性与文本性的关系,从新历史主义对历史与文本关系进行重新定义,还是从文本与书籍、文本性与互文性的角度来进行后结构主义的文本话语建构,亦或是从文化研究路径出发对社会阶级、文化现象、意义生产进行文本化反思与批判,文本与文本性在不同的思想与理论框架下具有着不尽相同的指涉意义与范围。尤其是在数字文化时代,作为电子文本的文本,其边界在信息技术环境中被进一步打开,被作为信息而重构,被作为数据而被计算机处理。这也为我们研究文本及文本性提供了新的可能和契机。

[关键词]文本;文本性;数字人文;物质性;电子文本

中图分类号: I01

文献标识码: A

文章编号: 1004—3926(2020)08—0172—09

基金项目: 国家社会科学基金重大项目“西方美学经典及其在中国传播接受的比较文献学研究”(17ZDA021)阶段性成果。

作者简介: 陈静,南京大学艺术学院副教授,文学博士,研究方向: 数字人文。江苏 南京 210023

在千禧年前后的一个访谈里,德里达曾有过这样一段表述“我发现我又处在一种矛盾之中,不过我接受这种处境。一方面我说: 书已完结——从我说出‘书的终结’那日起——但同时我也支持拯救书以抵抗威胁着记忆、威胁着数字文化的某些新技术。这里我也试着同时进行两种相互矛盾的工作: 接纳新技术,接纳一切与之一同而来临的东西,将之当作某种运气来接纳,但同时也去发现与之具来的威胁,以我所能,以我的方式与之抗争,以求保存一切与书本文化相联的东西。不仅仅是卷轴的形式,还有阅读的时间,阅读的耐心以及阅读的个体空间,即一切与书文化相联的品质”。^{[1](P.8)} 今日读来,德里达言语间所表达出的矛盾心理,是信息时代人文研究者都能感受到的茫然与不安: 信息时代的到来是如此深刻地改变了文学艺术的生存语境,促使了基于新媒介、新技术、新方法和新语言的文艺创作与现象的出现,同时也在挑战与改变着我们介入、观察和研究文学艺术的方式,那么,该如何介入? 是无条件的接纳,亦或是尽我所能的去保留和抗争? J. 希利尔·米勒的答案是跟随,他的“文学终结论”唱响

了“建立在印刷技术、报纸以及印发《宣言》的地下印刷机和出版商的基础之上”的书籍文化不可避免的衰落。^[2]而马克·波斯特则更犀利地指出,德里达在宣告印刷时代终结的同时却以一种谨慎甚至是怀疑的态度徘徊在“深渊”的边缘“似乎他(德里达)不能确定,解构到底是哲学家的‘灰身灭智’(gray on gray),描画着印刷书写这个已逝时代的轮廓,还是先觉者舞动的星,在宣告电子书写的诞生。现今的情境是一个深渊,向里面窥视只会看到怪物”^{[3](P.141)}。如果说上个世纪下半叶,德里达在深渊边的彷徨与矛盾,或是因为新技术的发展尚处于初期,具有诸多不确定性,造成了认知上的困难。那么在 20 年后的今日,随着信息技术的高速发展和广泛应用,我们是否可以断言我们已经拨开笼罩在深渊之上的迷雾,看清了眼前的路了呢? 似乎并没有。我们还在经历着从书籍文化到数字文化的转型期,依然被如德里达“双重约束”(double bind)般的心态所束缚,在不断地反思与再认识中寻找着自己的路。

数字人文或许是这样一条路。首先,数字人文是信息技术、数字技术所带来的知识生产的新

形态。其次,“数字人文”从知识谱系上对已有的文学艺术理论进行了拓展,以多样性的“人文”的概念兼容了数字文化中的各种形式,比如电子文学、游戏研究等。这些数字文化形式具有“实践性”的特点,与数字人文学具有一致性,从而也造就了广泛意义上的数字人文学者,一个以实践性介入知识生产的广泛、多样的群体。

数字人文的实践性还体现在研究者们从不同的路径和角度运用数字技术对人文问题进行研究和呈现。在这个过程中,技术与人文的概念发生了交混,从而造成了一些看似相同实则差异很大的概念被不加区别地使用。“文本”就是其中一个具有代表性的概念。在20世纪之前文本尽管被认为是可以脱离特定媒介而独立存在的,但更大程度上还是被认为是一种基于印刷技术得以呈现在纸上并能被引用的符号。这些符号可能具有特定意义而且只能被在一定的媒介载体上才能被理解(作为手稿或者版本意义上的文本),或者并不具有某一种特定意义(作为作者意图或社会文化的呈现)。20世纪中期,“文本”作为一个关键词被打开,其脱离了印刷文化既定框架内的使用语境,而被赋予了更多的意义,比如文学中会质疑是否有一个文本存在,新历史主义将文本视为历史语境的产物而历史皆可成为被研究文本,(后)结构主义更愿意将文本从作品中解放出来,强调文本无处不在,文本与文本的关联性构成了意义的网络,而文化研究则视文本为一种元模型,将社会、文化、历史都视为可以被作为意指实践的对象。人文领域内关于这些的研究也不胜枚举。然而,我们还应该看到在20世纪中期以后,文本在信息科学、计算机科学领域也被广泛地大量使用。比如计算科学中文本多指文本数据,是一种可以被分解为段落、句子、词和最终单个字符的文本文档。与此同时,文本是可以数字形式被表示的,也就是说,文本中每个独立的字符都有对应的代码。比如最常见的ASCII字符集和Unicode字符集。而在数字人文领域中,多学科和跨学科的研究使得不同意义上的“文本”概念被大量同时使用,造成了概念上一定的混乱。比如在数字档案与编辑中,文本既具有文献学中内容对象物的意义,同时又具有标记文档的意义,两者相互关联但又不完全相同^[4]。数字人文在其发展过程中,两个重要的来源,即“人文计算”(humanities computing)和

“电子编辑”(electronic edition)都非常强调文献学意义上的文本,但前者更多的是从语言学的角度强调作为语料的文本之于计算的意义,而后者则将文本视为一种无法与其历史物质技术剥离的社会实践。两者都对20世纪以来后结构主义开启的、丰富多样的、理论化的“文本”保持了一定的距离^[5]。近年来,随着“文本标记”、“文本挖掘”、“文本分析”等数据分析技术的广泛应用,使得计算意义上作为语言语料的“文本”概念更加凸显。阿兰·刘因此将数字人文前身的“人文计算”称为“基于文本的数字人文”(text-based digital scholarship),指出更广泛数字人文应该包括人文计算,以及新媒体研究。而后者恰恰是沿着后结构主义、解构主义和批判理论路径与传播研究融合形成的一个跨学科研究领域^[6]。数字人文学者也对其进行了回应。比如霍伊特·朗和苏真在《文学模式识别:文本细读与机器学习之间的现代主义》一文中,开宗明义地提到他们试图建立一种介于(结合了)文学细读和机器学习的文学模式识别的分析方法,超越计算模型与批评理论对峙的僵局。此处不对该模式做评价,但想指出的是,这种努力也是本文想要试图实现的一种路径,即从问题出发,从文本理论、电子编辑、信息理论和数据分析的角度对数字时代的“文本”问题进行辨析,试图在数字人文视野下打破计算本位意识下将文本作为数据处理对象的狭隘观念,重新打开“文本”概念的复杂性,建立一条融合了媒介研究视角的数字人文文本研究路径,发现数字时代人文研究持续发问并进行回答的动力机制。

一、从文本到电子文本:被解放的文本

在书籍时代,作品-文本是文本概念的核心内涵。德里达所强调的文本与文本性恰恰是出于对基于印刷的文学观的批判。文学性不是一种自然本质,不是文本的内在物,而是指向了一种文本间的意向性关系。文本与文本之间进行了一种不断重复、折叠的游戏过程,而对文本的阅读或者说之所以认为这个文本具有文学性,则很大程度上依赖于这种文本之间的意向性的指向所体现出的社会性法则的意义。在这个意义上,形式与内容的区分本身已经没有意义了,因为,文本本身就处于不断地延宕和补充之中。后一个文本则是从不同的角度和不同的元素对前一个(或多个)文本的模仿,差异性并不在于文本本身的独创性,而在于

其所在文本的网络中的位置。

《从作品到文本》是罗兰·巴特在1971年发表的一篇文章,在其中,巴特将作品与文本作为两种不同的概念进行了比较性的说明,成为了“文本”理论的关键性著作之一。而“从作品到文本”也成为西方文学理论从现代范式到后现代范式的一种概括性描述,并为文本的解放奠定了基础。罗兰·巴特在《从作品到文本》中对“作品”与“文本”进行了比较。他从7个方面指出了两者之间的差异性:作品是物质的碎片,占据书的一部分空间,因而能够握在手中;而文本是方法论范围内的,是一个演示的过程,只存在于话语的运动中,因此能在语言中把握——作品才是文本想象的细微末节。与作品相比,文本不能进行简单的分类,其所涉及的是某种不适当的体验,接近了阐述的规则极限,与套语(Doxa)相背离。作品接近所指,因此是文学学科的、哲学的、有待解释的对象,像一般的符号一样发挥作用;文本则接近能指,是一种游戏,文本的逻辑是转喻的——作品的象征快被耗尽,而文本是激进的象征,没有中心,没有终结。作品是有源头的、遵循着一元论哲学的;文本实现了意义的不可化约的复数,是作为信息而存在,其复数依赖于意义交织的复数性,也就是互文性。作品具有目的论,可以推导至某个像作者这样的本源,作品可以比喻成以发展的方式成长起来的有机体;而文本则是一种网络,是系统联合的结果,因此没有本源可以追溯,作者在文本中也是客人。作品通常被认为是消费的独享,等同于书的阅读行为;文本则是一种游戏、活动、生产和实践,在意指实践中将书写与阅读结合起来。文本通向的是愉悦之路。^{[7] (P.156-163)} 尽管他是将书籍与作品等同起来,而忽略了两者之间的差异,但他确实看到了作品(书籍)所具有的物质性特点——占居一定的物质空间、可握在手里、可阅读的对象。同时他还指出作品是解释与学科的对象,有着目的论的本源的存在。与作品相对应的,罗兰·巴特提出了“文本”理论。与作品,这个封闭在一本书的外壳里面的书写的主体,一个已经完成的对象相比,文本在巴特看来,是属于方法论的,是只有在生产活动与意指实践中才能理解的。作品和文本之间的区别可以用一系列的差异的术语表示:事物和过程、已完成的产品和生产性、能指与所指、“真理”与“游戏”。

在《文本理论》中,巴特对古典意义上的“文本”和他所提出的“文本”进行了区分。他指出,在认识论的角度来看,文本在经典的意义中,是一个以符号为中心的概念体系的一部分。他将从斯多葛学派到20世纪中期的西方文明都称之为一种符号的文明。其原因就在于所指与能指的结合构成了封闭的古典的符号。在这个符号中,歧义、漂移都是不允许的。而“古典的文本也是同样,它关闭了作品,用它的字母把它串起来,吸引它到它的所指,它因此用两类操作来约束我们,这两类操作都试图修补那些因为上千个原因(历史的,物质的,人为的)而打入符号完整性的洞。这两类操作是修复和解释”^{[8] (P.33)}。巴特在此处强调的修复也就是文献学意义上的在不同版本之间存在的差异的比较,古典语文学或者文献学的一个重要的意义也就是试图找到“原作”或者作品的“原真性”,力图尽最大可能去再现作者的意图。不同的版本背后总有着一个原本的存在。这一点在我们前面讨论印刷术及书籍的时候也已经讨论过。而文本之所以能成为一个作品,也就在于这种意义的修复和解释过程中。用巴特的话来说“文本的精确性的储存由它的原初,它的意图,那种不得不坚持和再发现的权威的意义形成的。文本因此变成了所有解释学的那个对象。从能指的‘补偿’,人们可以自然的经过到达权威的能指的解释:当文本被一个,仅仅一个,意义,一个‘正确的’意义,一个权威的意义占据的时候,文本就叫做作品了”^[8]。继而,巴特将古典的文本的概念与一种形而上学、真理联系在一起。在此,又一次印证了我们之前所作出的判断,巴特与德里达之间的共识。而与德里达不同,巴特并没有就此指向一种“终结”,他认为这是一次符号的危机,是由尼采在19世纪宣布的,而现在(20世纪60年代)则在语言和文学的理论中再次开放了,由符号的意识形态批评和新的文本取代了文献学意义上的文本。巴特赋予了文本以新的生命,在危机之后的一种新的可能性。

德里达的文本观与巴特的文本观有着类似之处,都强调文本的开放性、流动性以及作为网络和互文性的文本。可以说,后结构主义的文本观是在印刷文化内部的一次颠覆,巴特及德里达等人将基于印刷的作品作为了文本的对立面,对其进行了批判,作者不再成为作品的本源,作品也不再

是文学的基点 相反地 ,文本被提升到了根本性的地位 这无疑是对文本的一次解放。但是 ,我们应该看到 ,后结构主义的文本观依然是在印刷文化的内部所进行的一次解构。巴特的“理想文本”、德里达的差异之文都只是基于一种隐喻意义上的理想化表达 ,而真正实现这种文本观的却是在超文本出现之后。超文本以及电子文学理论家们通常都会将这两人视为在前超文本时代就已经对超文本进行了理论阐发的理论家 ,并且不仅认为超文本是对他们的文本理论的现实化 ,而且在电子文本及文本性问题上都以两人作为理论源泉之一。

乔治·P·兰道(George P. Landow) 在他 2006 年最新版的《超文本 3.0》的导论一开始就写了这样一个小标题“超文本的德里达 后结构主义的纳尔逊?”(Hypertextual Derrida , Poststructuralist Nelson?)。纳尔逊是超文本的定义者 ,也是最早的实践家之一。兰道在这里将他们换了身份 ,就在于想突出德里达与超文本之间的关系。事实上 ,正如在本章一开始就说到的 ,兰道是将后结构主义者德里达、巴特以及超文本研发者的纳尔逊和凡戴姆一起视为一个范式转移的代表。

尽管兰道这里是说了“超文本的德里达” ,但事实上 ,应该加上巴特的名字。“罗兰·巴特在《S/Z》中提出的理想文本在计算机超文本中实现了。计算机超文本是由被多重路径、链条或踪迹电子链接起来的文字(或者图像) 块组成的 ,多重路径、链条或踪迹在一个开放 - 结局的、永远未完成的文本性中用链接、节点、网络、网页和路径这些术语来形容。”^{[9] (P.2)} 兰道将巴特理论中的术语与信息技术中的术语之间画上了等号。由此他在巴特的文本(性) 和超文本(性) 之间建立了一种联系。兰道还用巴特的术语“文片”(Lexia) 来形容在超文本中被链接起来的文字(或者图像) 块 ,并指出巴特在可读文本与可写文本之间的区分也显示基于印刷技术和电子超文本的文本的差异。那么超文本是如何生产出了巴特意义上的可写文本呢? 兰道强调了计算机与书籍的不同 ,同时他说明了超文本具有编辑、添加注释的能力 ,这就使得超文本体现出一种网络化的特质 ,同时呼唤了积极读者的参与^{[9] (P.5-6)}。而这种非常积极读者(Very Active Reader) 在兰道看来也同时是读者 - 作者(Readers - as - Writers) ,这不仅证明了

超文本技术具有积极的文化民主的意义 ,而且读者阅读超文本的过程也就成为了一个创造超文本的过程 ,从而实现了巴特理想化的可写文本。

在强调文本的开放性时 ,兰道在指出巴特、德里达以及巴赫金和福柯都使用了像链接(Link , Liasons) 、网(Web ,Toile) 、网络(Network ,Rèseau) 和编织(Interwoven ,S’y tissent) 这样的术语的同时 ,还强调了巴特与德里达对于超文本的不同意义。前者是强调可写文本和它的非线性 ,也就是网络性 ,而德里达则强调了文本的开放性、互文性和在一个特定文本的内部和外部之间差异的不适宜^{[9] (P.53)}。兰道主要指的就是德里达在《散播》中说到的对柏拉图的文本与由希腊语言系统编写的世界之间涉及一种虚拟的、动态的和边缘的方式。兰道认为德里达在这里实际上是在描述现在的超文本系统。同时兰道还从格雷戈里·厄尔默的“应用文字学”(Applied Grammatology) 角度出发 ,认为德里达(或许是提前) 已经认识到了一种新的、更自由、更丰富的文本形式。这个依赖于“离散阅读单元”(Discrete Reading Units) 的文本对互联网时期的我们的经验来说更真实。兰道将德里达对这种不连续的聚合的强调看做是和超文本有着精神上的契合之处 ,称之为“元文本”(Metatext) 。此外 ,兰道还用德里达的“去中心化”(Decentering) 的概念来形容读者在网络上浏览网页的体验。当读者在网络上浏览网页的时候 ,也就是身处一个文本的网络 ,而他们不断的从一个网页到另一个网页的过程 ,也就不断的在进行德里达意义上的去中心化的过程。兰道也将这一过程视为超文本的一个基础性特征 就是由链接的文本组成而没有主要的组织轴。“换句话说 ,元本文或者文档系统——这个整体在印刷技术中被描述成书籍、作品或者单一文本——是没有中心的”^{[9] (P.56)}。正是在这样一个去中心化的过程中 ,选择的权利掌握在读者的手中 ,因此读者在这里成为了选择的中心。有意思的是 ,兰道指出了一点 ,所有的超文本都允许读者自己去选择考察和体验的中心。这也是德里达强调的: 当文本的意义处于变动不居的延宕状态下 ,不断的等待着新的补充的出现 ,从而填补缺席的“空白”之处的时候 ,读者事实上成为了意义(仅仅是一种理解上的解读) 的生产者和散播者。但是 ,读者在非中心的文本中进行选择并不意味着个体的经验就是非中心的。事实上 ,

往往读者在其中进行的是有中心的阅读,只不过这个中心因读者个体而异。这一点也同时体现在超文本的非线性文本呈现和阅读上,即非线性的超文本允许线性的阅读。

二、从作者意图到编辑意图:被建构的文本

尽管巴特宣称“作者已死”,而超文本也在一定程度上实现了作者权民主化,但印刷文化中作者权威性的问题实际上并无法单纯地从理论的层面打破,更需要从文本,尤其是结合文本生产体制来审视文本与作者权问题。在印刷文化中,“作品”以及“文本”历史的与作者以及权威性联系在一起。这促使文本成为了一个有待解释的对象,无论它被解释成什么,都体现着作者的意图。甚至,一部文学理论史就是一部作者史。这也是哈罗德·布鲁姆在《影响的焦虑》中所努力证明的,“一部诗的历史就是诗人中的强者为了廓清自己的想象空间而相互‘误读’对方的诗的历史”^{[10] P.3}。而印刷术技术与书籍的物质性就被转换成了“作品”与“文本”在不同手稿与版本之间的差异。而对这种差异的分析则是为了进一步证明作者的意图对于“作品”与“文本”所产生的影响及其正确性和权威性。这样一种观点在电子媒介时代已经受到了挑战,被认为是一种形而上学的整体论思想。

艾斯潘·J·阿赛斯(Espen J. Aarseth)认为“什么是文本”的观念受到了阅读、书写和稳定性三个方面的共同作用:(1)文本是你所阅读的,你所看到的文字和句子,它们在你的头脑中产生意义;(2)文本是一个信息,充满了特定作者/文类/文化的价值和意图;(3)文本是一个固定的部分(开头、中间、结尾)序列,这个序列不能改变,尽管它的意义可能改变”^{[11] P.53}。他进一步指出,可以从两个由不同的规则统治的、内在独立却又相互影响的视角去看“文本”的观念:信息的(Informative)——文本作为技术的、历史的和对象和解释的(Interpretable)——文本是被单独地接收和理解的对象。阿赛斯认为文本的信息层面通常难以看到,因为是最明显不过的。文本的可视文字和空间,就是通常说的“手稿”,除此之外,还包括实践、结构和例行使用。^{[11] P.53}

阿赛斯以此为出发点,对于之前的文本概念进行了反思。与手稿联系在一起的,是文本的作者,也就是署名的问题。被署名了的作者被认为是文本的主人,但是作者的署名却并不是在文本

之内被发现的,而是在分类、书评以及书页的页眉或页脚中,他称之为“文本的,但不在文本内”(Of The Text, But Not In The Text.)^{[11] P.55}。在文本的传统中,阿赛斯指出,“一句话,我们更喜欢想象的完整的形而上学的对象,而不是我们观察到的稳定的版本。正如我们能够想象和解构一个理想的版本,完整无缺,我们的形而上学也保持了完整”。这种文本的观念正是我们在之前提过的,与作者的权威性联系在一起。阿赛斯认为这是我们的观念决定的,“不仅仅是我们对它的功能、意义或者形而上学的可靠性,也是由它看起来是什么做成的,什么情况下与我们遭遇而让我们得知它的存在。留下来被调查的,因此是文本性通过地点,解剖和临时性超越形而上学存在的可能性”^{[11] P.55-56}。阿赛斯将文本与一种形而上学联系在一起,而将新的文本的可能性赋予文本的构成要素和阅读语境。他同时认为,这种反思源于一场由信息时代的数字奇观所带来的“文本危机”(Crisis of the Text)。然而面对许多关于“书本已经死了”、“这是印刷晚期”、“未来,人人都是作家”的呼声,阿赛斯认为这些呼声往往由于频繁使用电子文本(Electronic Text)与印刷书本(Printed Book)这样的概念而使得分析难以细致下去。因为这种理论上的整体性,使得对于文本性的定义可以从这个整体性出发,而不需要考虑到实践、历史或技术^{[11] P.56}。

德里达与巴特都指出过,文学是一个学术体制内的产物。但是他们所强调的文学体制主要是基于马克思主义的物质基础。而在文本的层面上,这种学术体制的影响则主要是基于编辑等学术活动。这是一个更为具体化的操作,并且在长期的过程中逐渐形成的。在印刷文本数字化的过程中,计算机所具有的功能性模糊了创作、编辑与出版之间的界限,一方面促进了文学的民主化,另一方面也因为电子文本生产活动中,不再有原本在印刷文化中的作者和编辑之间的明确分工,从而使得原本被掩盖在文本后面的操作层面的意义被凸显了出来。

安娜·冈德(Anna Gunder)则在其论文,《形成文本,执行作品》(Forming the Text, Performing the Work)中从电子媒介和印刷媒介之间的关系入手,考察了编辑是如何在印刷目录研究语境中使用“作品”、“文本”和“文献”等基本术语的。她指出,编辑将作品设定为一个“抽象的艺术整体”,预

设了一个理想,然后通过对不同的修订本和副本的校对来满足这个理想化的预设。而这种预设也要服从于协商、挑战和交流准则以及文化的先决条件等外部语境中的活动。冈德还指出作品向着理想目标的过程是通过某种文本来实现的,而这个文本被称为一个特殊作品的特定符号系统。在讨论绘画时,人们是不能将绘画作品同画布、色彩、线条等构成作品的物质基础分离开的,而编辑在讨论文本时确实将文本视为从一个人工制品的物质具象中分离出来的对象。只有在文本的最低层面上,也就是在文献的层面上,才会被看做是与符号系统融合的、抽象的、表征的也是物质的人工制品^[12]。

冈德的这种看法,是与文学以及文学批评作为一种体制化的学术建构分不开的。在这样的专业活动中,作品与文本才可能作为一种稳定的、有标准可循的对象被固定下来,也只有这样一个意义层面上,作品和文本具备了被进行分析、分类和经典化的条件。值得注意的是,冈德所说的“文献”作为最低层面的文本与阿赛斯所说的文本的信息层面具有相似性。

而彼得·施林斯伯格(Peter Shillingsburg)在《计算机时代的学术编辑》(In Scholarly Editing in the Computer Age)一文则应证了冈德的观点。他在文章中将文本定义为“文字和标点的实际秩序,包含在任何一个物质形式中,比如手稿、校样或者书籍中”。他指出,文本包含在内且固定在物质形式中,但并不是物质形式本身,因而文本不受实践和空间的限制。这就使得施林斯伯格认为,在从印刷到电子媒介的转换中,文本是不会发生变化的。而海勒丝则将施林斯伯格的这种观点认为是印刷-中心的观念^[13]。因为这种观念所考察的文本依然还是限定在印刷文本时代,而没有考虑到电子媒介对文本的影响。

三、从信息到过程:被重构的文本

基于印刷文化的文本概念被打破后,如何建立一种新的文本观?如何身处深渊而坦然直视巨龙?N.凯瑟琳·海尔斯(N. Katherine Hayles)在《我母亲是计算机——数字主体与文学文本》(My Mother Was a Computer: Digital Subjects and Literary Texts)中,从“翻译媒体”(Translating Media)的角度对从印刷文档转变成电子文本的过程进行了考察。她指出,这样一个“翻译媒体”的过程也是一个对作

品、文本等概念进行反思的过程。而当下对于从印刷文档到电子文本的过程中,“大量的注意力放在了意义与语言和书目符码的关联上,但是几乎没有注意到意义与数字代码的关联”^[13]。海尔斯由此对文本(性)进行了反思,她指出在电子文本时代,学者们对于之前的作品及文本观进行了反思,但是他们的文本观都还是一种非物质性的文本观。因此,她认为“我们认为‘文本是什么的核心’的核心是相信‘作品’和‘文本’是非物质的建构,独立于它们所具体化的基质。我们迫切需要重新思考这个假设,因为只要它保持了完好无损,那么试图说明印刷与电子媒体的特殊性的努力就不可能实现”^[13]。

正是在这样一个前提下,海尔斯界定了这样一个定义:“一个具体化的文本的物质性是它的实体特性与它的意指策略的相互作用”^[13]。在这里,文本的具体化是指某种文本类型在某种特定媒介中的示例化,比如我们说电子超文本,那就是指超文本这种文本类型在电子媒介中的呈现。

而“物质性”这个概念则有两方面的意义。一方面,该概念延展出了单个物体的边界。是物体生产的社会、文化和技术的结果。首先,作为物质生产的书写技术也是一种社会和文化活动。“古代书写技术不仅是纸莎草、墨水和制作书卷的技巧,还有古代书写的风格和类型以及古代修辞的社会和政治实践。现代书写的技术不仅仅包括印刷技术,也包括当代科学和官僚的实践以及印刷读写能力(literacy)的经济和社会后果”^[14]。组成一个既定文本的物质性的并不纯粹是硬件或者可见的装置、设备。比如,我们在讨论一个电子超文本的物质性的时候我们并不关心主机的电源线之类的设备部件。同时,既定文本的物质性是否可见还取决于解释和批判的过程。比如对于某些读者而言,依赖于一定的批评理论或者从一定的视角就可以看到别的读者看不到的东西,这点在电子文本的代码层面上体现的就很明显。另一个方面,物质性是作为一种再概念化的,在文本的实体特征与它的意指策略之间的互动。比如在考察电子诗歌写作中的语言文本的物质性时,语言文本的实体特征就是代码,而它的意指策略就是作为持续的表演^[15]。

然而,与强调文本的物质性相对立的,在新的电子语境中对“文本(性)”概念进行重构的过程

中最普遍也是最为人们所接受的观点就是“文本是信息”。这种观点与控制论的“信息”的观点分享了一个理论的基础。控制论的创始人之一诺伯特·维纳(Norbert Wiener)就曾在其著作中论述过“艺术作为信息”的观点。在维纳看来,信息就是模式。艺术的价值就在于信息量的占有以及有效性上。传统的艺术作品中之所以有着大量的陈词滥调,就是信息的复制性使得在不同的艺术作品中却含有相同的信息。就像透视法在文艺复兴初期刚出现的时候,还是一种新鲜的方式,但是等到它作为一种基本的技法被广泛接受的时候,这种信息就不再被人们注意到。而“要使社会上的一般信息丰富起来,该信息就必须说出某种在本质上异乎社会上原先公共储藏的信息”^[16](P.102-103)。而伟大的经典中就有这样的大量的一般信息,这也是后来的人们会对莎士比亚的作品觉得乏味的的原因。但当人们从原来的陈词滥调中重新建立了与这位作家的在信息方面的联系后,人们会对其作品作出新的评价。维纳是从信息理论的角度重新阐述了艺术作品与社会的关系。他的认识的基础是信息是可以超越不同的事物而存在的,海尔斯所说的“信息失去身体”的意思。同时,信息反馈循环使得人们可以在不同的语境下对艺术品所具有的信息作出不同的理解。以此来看文学理论,事实上,不同的理论也是从不同的角度对文学作品进行了认识。阿赛斯很明显的接受了维纳的这样一种观点,他所说的信息层面的文本也是从这个角度去对文本进行了新的阐发。他指出“文本是一种一套矩阵间的交叉产品,这些矩阵包括语言的(刻画)、技术的(机械条件)、历史的(社会-政治语境)。因为这些变动的临时的稳定性,文本是可能终结和简化的过程”^[11](P.59)。文本的矩阵也就是一个信息的集合,语言的、技术的、历史的信息共同构建了文本。

特别是在计算机等数字化技术出现以后,这种“文本作为信息”的观念表现的更为突出。当文本被用字符集乃至数值来表达的时候,文本在语言符号和物理介质层面上的意义就被消除了,可以在不同的媒介之间流动。正是这样一种观点,也使得许多人相信文本在不同媒介中的转换,就像水从一个容器转移到另一个容器当中,其本质不会发生变化。

但是,当我们在与电子文本遭遇的过程中,我

们又会发现,电子媒介的特性不可避免的会对电子文本生成产生影响。电子文学的文本以不同于印刷的方式存在,有数据文件和程序访问和处理文件、硬件功能性。这些功能性负责解释或者搜集程序等等。所有这些共同生产出电子文本,略掉其中任何一个,文学意义上的文本都不能生产出来。因此,更准确的说,应当称呼一个电子文本为一个过程,而不是一个对象。当然,它不能被认为是一个卡式录音机或者光盘只读处理器,因为它们绝不可能独立的生产出一个文本,除非它们是由在专门的硬件上的专门的软件来执行的。

与此同时,文本作为过程还体现在文本的存在方式上。在印刷书籍中,文本是先于阅读存在的。电子文本没有这样的先在,它不存在与电脑或者网络系统的任何地方。当它显示在屏幕上以后,阅读的过程就开始了。这是一个只在创造文本的程序被击活的时候才发生的过程——需要与读者认知过程同样的操作。在这个意义上,电子文本比印刷文本更过程(Processual)一些;它从它的本性来说是操作性的——不管那些使用者带来什么样的图片和过程,也可以无视版本和副本之间的不同。

兰道对于印刷文本的超文本化过程的论述从超文本的角度对“文本作为过程”的观念进行了说明。他指出至少可以采用四种不同的方式去创造一个印刷术的超文本版本:1) 仅仅需要把印刷书册中的文本放在一个电子环境中,例如 Windows 或者导航扩展书 坚持原始的页面^①;2) 将书页的再现更适于显示器,比如采用将所有的注释都转换成文片的方式调整印刷版面;3) 最低要求的超文本转换,即保留了印刷书的原始章节作为基本单位,然后将它们之间以及它们与目录之间用一种线性序列链接起来;4) 最具有挑战性、最有意思的翻译,就需要利用超文本的可能性,需要涵盖一个超文本版本所要求的材料(其中包括其他学者的著作,评论者的评语以及其它创作的材料……),最大程度的利用超文本作为一种书写形式所具有的能力^[17](P.15)。因此,我们必须意识到,媒介的特性在作为过程的文本中起到了非常重要的作用。

事实上,作为信息和过程的文本观念也不是对立的。这两种观念可以互为补充。确实,文本可以以信息的方式在不同的媒介之间流动,它可以被看做一种模式,但与此同时,在不同的个别媒

介中,它又与媒介的特性相结合,产生出新的文本性。而这就要求我们必须要注意到媒介的特性在电子文本(性)中所具有的作用。

四、从文本到数据:被处理的文本

然而,当我们将文本作为一种可用计算机程序处理的对象时,我们实际上并不那么在乎文本的媒介特性。也就是说,文本最重要的是如何将特定文档格式中的内容转化为计算机可处理和计算的数据。最典型的就是数字化的印刷文学文本。印刷文学文本的数字化则是利用不同的软件以不同的方式实现。比如,在当下的出版流程中,印刷书籍在出版之前,都会以数字化的方式进行编辑、排版、设计等。事实上,论文、报告、新闻等印刷文本在打印或者印制之前,都是以数字化的方式储存在计算机硬件和软件中。这种数字化的储存和展示方式一方面是利用了计算机的数字化的信息处理方式,另一方面则是利用了相应的软件再造印刷页面的视觉效果。在这样的情况下,数字化的印刷文本比手稿更接近印刷文本。因为在计算机的文字处理软件中,进行编辑、修改和设计都不会留下任何的痕迹,而且,由于计算机提供的可复制性,使得一个文本可以复制多个同样的文本出现,就像书籍的印刷传播一样。另一种比较典型的数字化的印刷文本形式,是一个从物理形态到数字形态的转化过程。无论是具化为何种介质的文本在经过数字化后都会成为数字格式。但在这个过程中会有大量的数据损失。以书籍数字化为例,物理形态的书页在被数字设备扫描或者翻拍后成为位图的图片。在这个过程中纸张的物理属性首先被放弃,而成为了二维平面的图片。书页脱离了独立的物质实体书籍而成为具有独立文件名的图片文档。其色彩、大小、肌理感也发生了变化,并随着硬件配置和软件环境,比如显示器、显卡、操作软件甚至网速的影响。但作为位图的文本数据是无法作为数据被处理的,因此还要经过OCR(光学字符识别)的处理,然后附加上一层经过字符集转码的文本图层,从而可以被提取出来转存为文本格式的文本数据。在这个过程中,文本的排版、布局、注释等版本目录信息往往会被忽略或者作为噪音被处理掉。与此同时,这个过程中还存在另外一个风险,即字符集对应的问题。如最常用的Unicode字符集是以单字节来表示一个英文字符,但表示中文字符就需要3-4

个字节,如果运行环境中没有相应的字符集,那特殊中文字符无法被正常识别^{[18](P.228-246)}。在之后的结构化过程中,线性书写的文本被彻底打散并重新按照数据库或者数据分析的要求进行结构化和清理,文本就转换成了可处理的数据。

与此同时,从电子文本的生成层面上来说,数据化结构构成了电子文本的深层生成结构。电子文本生成最基础的层面就在于数据库的构成形式上。数据结构决定了我们通过什么样的结构来描述我们需要描述的现实事务。同样的,电子文本需要通过什么样的方式建立一种文本的内部关系就需要通过设计数据库来达到。而界面则提供了一种将这种数据组织结构的隐喻。传统的文学作品中数据库和界面是一致的,或者说界面是不存在的,我们直接面对的就是被组织起来的数据,也就是由文字构成的叙事。而在电子文本这里中,界面则成为了深层数据组织结构的直观呈现。比如,叙事的“用户”正是跟随由数据库创造者建立的数据库记录之间的链接来穿越数据库。而一个交互叙事就可以理解为通过一个数据库的多重轨迹的总和。而一个传统的线性叙事就是多个可能的轨迹中的一个^{[20](P.227)}。由此,我们也就可以理解为何在网络小说中,再阅读会具有重要的作用。这是因为网络小说的作者创建一部作品的时候也就是在组织一个数据库,其所提供的数据(在此可以认为是文片)是通过一定的组织方式建构起来的,其本身就提供了多种选择的叙事路径。因此,读者或者“用户”在选择叙事路径的时候,每次只能选择一条线性叙事的路径,而只有当重复地去点击链接的时候,才可能发现或者体验多重叙事的效果。

结语

计算机技术对于当代社会的影响是巨大的。文学所具有的电子烙印如此之深,以至于口头、手抄和印刷时代的很多媒介特点在当下的电子文本中“重新流行”(refashioned),比如现在的文本标记、文本语音化等等。数字技术以及相应的数字人文、数字艺术的发展确实使得批评家们无法置身事外。我们必须要去了解技术,也必须去认识到基于数字技术的创作方式已经改变了我们的批评方式乃至教育方式。事实上,国内外一些大学已经开设了有关数字人文、电子文本、人工智能的创作和理论课程。厄尔默所说的“电子能力”也已

经在逐步取代“识字能力”。在中国也是如此,当计算机的基本技能成为学生们的基本素质、计算机文字处理及网络已经成为人们主要的书写工具和交流方式的时候,我们是否还能在批评的领域中坚持我们基于印刷技术之上的文学观?

注释:

①这也是我们最常看到的利用软件制作的电子书的形式。

参考文献:

- [1][法]德里达. 书写与差异[M]. 北京: 三联书店, 2000.
- [2][美]米勒. 全球化时代文学研究还会继续吗? [J]. 文学评论, 2001(1).
- [3][美]马克·波斯特. 信息方式——后结构主义与社会语境[M]. 北京: 商务印书馆, 2000.
- [4][美]史蒂芬·德罗斯, 戴维德·杜兰德, 艾利·米洛纳斯, 艾兰·瑞尼尔. 文本到底是什么? [J]. 王晓光, 凌宇翔, 王俊芳译. 出版科学, 2016(3).
- [5]Jerome McGann. “Information Technology and the Troubled Humanities” *Text Technology*, 2005(2).
- [6]Alan Liu. The State of the Digital Humanities: A Report and a Critique. *Arts and Humanities in Higher Education* 2012(1-2).
- [7]Roland Barthes. *Image Music Text*. London: Fontana Press, 1977.
- [8]Roland Barthes, “Theory of The Text” ,in *Untying The Text: A Post-Structuralist Reader*. edited by Robert Young ,Boston: Routledge & Kegan Paul, 1981.
- [9]George P. Landow ,*Hypertext 3. 0: Critical Theory and New Media in an Era of Globalization*. Baltimore: The John Hopkins University Press 2006.
- [10][美]布鲁姆. 影响的焦虑[M]. 北京: 三联书店, 1989.

[11]Espen J. Aarseth, “Nonlinearity and literary Theory” ,in *Hyper/Text/Theory* ,edited by George P. Landow ,Baltimore: The Johns Hopkins University Press ,1994.

[12]Anna Gunder ,*Froming the Text ,Performing the Work* ,http://etjanst. hb. se/bhs/ith//23-01/ag. htm.

[13]N. Katherine Hayles ,*My Mother Was a Computer: Digital Subjects and Literary Texts* ,Chicago: University of Chicago Press 2005.

[14]Bolter ,J. David. *Writing Space: Computers ,Hypertext ,and the Remediation of Print* ,Mahwah: Lawrence Erlbaum Associates , Inc. 2001.

[15]John Cayley, “Time Code Language: New Media Poetics and Programmed Signification” ,in *New Media Poetics: Context ,Techno-texts ,and Theories* ,edited by Adalaide Morris&Thomas Swiss ,Cambridge: The MIT Press 2006.

[16][美]维纳. 人有人的用处——控制论和社会[M]. 北京: 商务印书馆, 1989.

[17]George P. Landow, “What’s a Critic to Do?: Critical Theory in the Age of Hypertext” ,in *Hyper/Text/Theory* ,edited by George P. Landow ,Baltimore: The Johns Hopkins University Press ,1994.

[18]Lik Hang Tsui and Hongsu Wang, “Semi - Automating the Transformation of Chinese Historical Records into Structured Biographical Data” ,in *Digital Humanities and Scholarly Research Trends in the Asia - Pacific* ,edited by Rebekah Wong ,Haipeng Li ,and Min Chou , Hershey ,PA: IGI Global 2019.

[19]Marie - Laure Ryan, “Introduction” ,in *Cyberspace Textuality: Computer Technology and Literary Theory* ,edited by Marie - Laurp Ryan ,Bloomington and Indianapolis: Indiana University Press ,1999.

[20]Lev Manovich ,*The Language of the New Media* ,Cambridge: The MIT Press 2001.

收稿日期 2020-06-05 责任编辑 申燕