

基于语料库的《文城》分析研究

徐冰榆

杭州师范大学人文学院, 浙江 杭州

收稿日期: 2026年7月31日; 录用日期: 2026年8月28日; 发布日期: 2026年9月14日

摘要

《文城》是余华创作生涯中的一部重要作品。本文通过使用AntConc软件对余华的《文城》进行深入分析, 旨在利用该软件对《文城》的主题、人物形象以及文本写作风格做出量化分析。通过分析数据, 验证小说所要传达的对人性善的关注以及对苦难抗争的主题, 详细分析林祥福善良、慈爱和乐观的人物形象, 最后总结得出《文城》具有用词简单质朴、重视人物描写和小说的写作风格。

关键词

语料库, AntConc软件, 《文城》

A Corpus-Based Analysis and Study of *Wencheng*

Bingyu Xu

College of Humanities, Hangzhou Normal University, Hangzhou Zhejiang

Received: July 31, 2026; accepted: August 28, 2026; published: September 14, 2026

Abstract

Wencheng is an important work in Yu Hua's literary career. This paper conducts an in-depth study of Yu Hua's *Wencheng* with the aid of AntConc software. It aims to make a quantitative analysis of the novel's themes, character images and textual writing style. Based on the analyzed data, this paper verifies the novel's core themes concerning the concern for human goodness and resistance against suffering. It also offers a detailed analysis of Lin Xiangfu's character traits: kindness, affection and optimism. Finally, it concludes that *Wencheng* features plain and simple diction, with emphasis on character portrayal and plot construction.

文章引用: 徐冰榆. 基于语料库的《文城》分析研究[J]. 现代语言学, 2026, 14(9): 316-322.

DOI: 10.12677/ml.2026.149928

Keywords

Corpus, AntConc Software, Wencheng

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

作为中国当代先锋文学的代表性作家,余华自 20 世纪 80 年代步入文坛以来,从第一部短篇小说《第一宿舍》到长篇小说《文城》,始终保持稳定且持续的创作活力。余华创作主要可以分为以下阶段:20 世纪 80 年代的先锋写作以零度叙事、暴力与荒诞为核心,以此拷问人性,呼吁读者对人道主义的关注,代表作品有《现实一种》《十八岁出门远行》等;90 年代以后,余华开始回归传统现实主义,写作逐渐转向民间苦难叙事,并在苦难中凸显人性的力量,以《活着》《许三观卖血记》为代表,洪治刚(2004)指出“以往的暴力快感不见了,代之而起的却是受难的主题”[1]。此后,余华在书写人性困境的同时,也开始对人性善进行思考,代表作有《兄弟》《第七天》《文城》等。

《文城》是余华睽违八年之后的又一部长篇小说,小说背景设定在清末民初,北方青年林祥福遇见了女子纪小美,二人坠入爱河。然而小美生下女儿后便悄然离去,为寻妻子,也为找寻小美口中的“文城”,林祥福背着年幼的女儿,踏上了南下寻人的漫漫长路。作为余华后期重要长篇作品,《文城》也引发学界众多讨论与争议,洪治刚(2021)指出《文城》以写实、抒情、诙谐、魔幻等诸多叙事手法,凸显了人性的温暖与晦暗、谦卑与暴烈[2],体现出余华对传统道德与人间温情的诗意回望。屠舒一(2024)从精神母题与象征维度出发,指出“文城”是一种精神乌托邦,这延续了余华一贯的“远行与追寻”母题[3]。“文城”的虚构性注定了主人公的终身无法抵达终点,但其一生的坚守、漂泊与救赎构成了真正的生命价值。与此同时,学界也对《文城》展开了激烈的讨论。李彦姝(2021)指出,小说人物存在过度理想化、伦理纯化的问题,林祥福近乎完美的形象弱化了人性张力,使得人物心理层次相对单薄[4]。在叙事层面,孟觉之、胡小兰(2021)认为《文城》“空间叙事是模糊的,历史叙事是断裂的,且传奇性的情节、景观式的场面、类型化的人物无疑在批判力度和思想深度上有所削减”[5]。总体而言,当前学界对《文城》的研究多集中于主题象征、叙事转型等宏观层面,结合文本微观语料的精细化分析的相关研究仍较为薄弱,有进一步拓展的研究空间。

自 20 世纪 80 年代起,语料库语言学开始逐渐与文学作品文本结合,通过运用计算机统计技术对文学语言、结构进行量化分析,以此为文学研究提供更可靠的量化数据,对文学作品进行更加科学、客观的分析。近几年,国内学者开始尝试运用语料库检索技术对部分文学作品展开研究,国外作品如《傲慢与偏见》(王雁, 2007) [6]《哈姆雷特》(胡开宝、朱一凡, 2008) [7]等,国内作品有《红楼梦》(刘惠丽, 2009) [8]《蛙》(陈蝉、程乐, 2014) [9]《白鹿原》(朱洪, 2021) [10]等。

本文以 AntConc 检索工具为基础,对《文城》全文文本进行量化统计与数据梳理,在检索结果的基础上结合现有研究成果分析小说的主题、人物形象特征与文本写作风格,实现数据量化分析与文本细读的有机结合。

2. 研究概述

本文以《文城》文本为研究对象,利用 CorpusWordParser 分词软件对语料进行分词,得到结果后对

新词、命名实体以及专业名词等未登录词进行勘误以确保后续语料库分析的准确性和严谨性。其次，运用语料库分析软件 AntConc 进行进一步的检索分析。在主题和情节分析中，主要运用 Keyword 功能来检索和梳理小说主题词，并对其进行细致地分析，得出文本的主题；在人物形象分析部分中，运用 Collocate (搭配)功能，以“林祥福”为检索词，详细分析与之搭配的共现词，并对林祥福的人物形象进行描摹；最后分析文本写作风格，主要利用 TTR、Concordance Plot (索引定位)以及 Keyword 功能，通过对比《现实一种》文本语料库，分析得到《文城》的写作风格，同时试图找寻余华写作风格转变的证明。

3. 《文城》的语料库分析

3.1. 主题和情节分析

通过提取和分析文学作品中具有超常频率的词或词组，可以确定文本的主题词和表达该主题的词集 [11]。AntConc 软件中的主题词功能主要是对文本中的词或词组进行统计分析，依据词或词组在文本中出现的频率、分布情况等因素，识别出能够代表文本主题或核心内容的词或词组。因此，获得主题词之前首先需要建立两个语料库，一是目标语料库，二是参照语料库。《文城》为本文目标语料库，其总字数约 19.4 万字；为准确获得小说主题词，此处所采用的参照语料库为《平凡的世界》，总计 104 万字。《平凡的世界》为茅盾文学奖获奖作品，基本可以反映当代中国文学作品的价值取向和总体写作风格 [9]，通过与其对比可以找到《文城》在主题等方面的特点。

通过使用 AntConc 软件中的 Keyword 功能检索两个语料库，对所获得的主题词进行归纳和分类，得到如下结果：

Table 1. Thematic word list of *Wencheng*
表 1. 《文城》主题词表

序号	类别	主题词
1	人物角色	小美、林祥福、陈永良、顾益民、陈耀武、林百家、张一斧、李美莲、女佣、婆婆、婴儿、士兵、女儿、朱伯崇、仆人、陈耀文
2	地点	溪镇、沈家、沈店、齐家村、铺子、码头、文城、城隍
3	与社会活动相关的名词	包袱、棺材、板车、田地、枪支、船家、木器、

通过对表 1 分析可知，第一类“人物角色”的主题词中出现频率最高的是“小美”，她是串联小说整体内容的关键，正是“小美”的存在才有了林祥福寻找文城的下文；“陈永良”“顾益民”“陈耀武”“林百家”都是正义、善良、温情的代表，也是林祥福生命中非常重要的人，这些人物的存在使得小说充满了温情和情义。主题词中也有恶势力的代表“张一斧”等等，这些人物都与小说情节的发展息息相关。

第二类“地点”表明故事发生地主要是在“溪镇”，值得注意的是小说题目是《文城》，而“文城”的出现频率却比“溪镇”的出现频率低。故事都与溪镇有关，却取名“文城”，余华曾评价“它是一个开放的书名，但所有故事都跟它有关。”可见，“文城”具有浓厚的象征意味，从表意上看它是一个并不真实存在的南方城镇；但是从深层意义上来看，它是林祥福对梦想执着地追求，从而真实地展现出余华小说对生命意识与本真的思考 [12]。

第三类“与社会活动相关的名词”中的“棺材”“枪支”表明死亡和斗争，该类词的高频出现说明小说中的矛盾冲突十分激烈。“枪支”可以对准百姓，也可以对向敌人，因此这既是灾难和苦难的印证，也是抗争和超越苦难的证明。

通过对这三类词的分析和归纳,基本可以得出《文城》故事情节和主题。《文城》以清末民初的社会历史为叙事背景,围绕主人公林祥福南下寻妻、奔赴“文城”展开叙事。首先,“人物角色”和“社会相关”类主题词印证了作者对人性的关注,在动荡的社会中以林祥福为代表的底层民众善良、正义,以民间温情对抗苦难,凸显人性的力量。其次,“地点”类主题词印证了故事发生的地点在“溪镇”,而“文城”则具有精神乌托邦的象征意义,象征了林祥福一生渴望的精神归宿。

3.2. 人物形象分析

Antconc 软件中的 Collocate (搭配)功能,可以系统地罗列出与检索词高频出现的其他共现词。利用这一功能,可以精准检索出与该词相关的所有内容,了解检索词与其他词汇之间的关联关系。在 Collocate 功能中 likelihood 通常是指对数似然值。对数似然值是一种反映真实性的指标,一般数值越高,相关性越高,它可以用于衡量两个词之间搭配强度。在分析人物形象时使用 Collocate 功能能够深入剖析角色性格。Likelihood 指标可以帮助确定与该检索词具有较强搭配关系的词汇,归纳整理与小说人物搭配频率较高的词汇有助于进一步分析人物形象。本文通过 Antconc 软件中的 Collocate (搭配)功能,以“林祥福”为检索词并将 Likelihood 作为排序指标,检索与“林祥福”有关的高频共现词,结果如下:

Table 2. Collocation search results for “Lin Xiangfu”
表 2. “林祥福”的 Collocate 检索结果

	Collocate	Rank	FreqLR	FreqL	FreqR	Range	Likelihood	Effect
1	女儿	1	66	21	45	1	48.047	1.417
2	告诉	2	38	30	8	1	40.224	1.768
3	来到	3	52	30	22	1	37.696	1.415
4	怀抱	4	18	1	17	1	30.928	2.376
5	尖刀	5	17	10	7	1	26.740	2.250
6	个	6	29	20	9	1	25.519	-1.164
7	毛驴	7	20	5	15	1	24.144	1.914
8	问	8	46	20	26	1	23.551	1.160
9	对	9	64	39	25	1	22.494	0.940
10	晚上	10	14	12	2	1	21.340	2.206
11	笑	11	35	16	19	1	21.306	1.280
12	看见	12	65	25	40	1	19.377	0.858
13	感到	13	24	8	16	1	18.940	1.488

由表 2 可知,检索词“林祥福”关联度最高的词汇为“女儿”,这意味着在“林祥福”的语境序列中,二者共现搭配最为密切,直观体现出林祥福对女儿的看重。不同于传统封建家长制里普遍存在的专制家长形象,林祥福摆脱了重男轻女的传统观念束缚,始终悉心照料、疼爱女儿林百家。小说主线虽围绕林祥福寻访纪小美展开,但对林百家而言,这段漫长岁月亦是一场长久的守候[13]。综上,林祥福是一位慈爱宽厚、思想超前,具有超越时代特质的父亲形象。

在检索结果中,可以发现与“林祥福”高频搭配的动词有很多,如:“告诉”“来到”“怀抱”(此处通过核查原文发现该词大多以动词形式出现)“问”“笑”等,这些动词以中性词或褒义词为主,侧面印证林祥福的性格底色善良慈爱、心怀乐观,并非凶狠暴戾之辈。

其中“毛驴”一词与“林祥福”的搭配频次同样突出。《文城》中的“毛驴”和《活着》中的“老黄牛”都具有相同的象征意蕴。在小说诸多情节中，自他拜师学艺、下地耕耘起，这头毛驴便朝夕相伴，成为他困顿境遇中重要的精神依托。即便命运坎坷、生计艰难，毛驴的陪伴反衬了林祥福始终未曾丧失对生活与未来的盼望。

3.3. 写作风格分析

为了对小说《文城》的写作风格进行深入的分析，此处试图通过与余华早期代表作品《现实一种》的比较分析得出《文城》的语言使用特点，同时印证余华前后期写作风格的变化。以《文城》和《现实一种》的文本为研究对象，用 TTR 值比较两个语料库在用词方面的特征，再通过 Concordance Plot (索引定位)功能分析两个语料库中小说人物在文本中的分布状况，得出小说写作的侧重点。

首先，在语料库语言学中，TTR (Type-Token Ratio)指的是类符(type)与形符(token)的比率。类符指语料库中出现的所有不同的词语，即文本中去除重复词后的词汇数量，形符指出现的所有词形，即总字数。TTR 在一定程度上反映了词汇的丰富度，TTR 值越大，文本蕴含的词汇量就越丰富。通过 AntConc 软件分析两个语料库，得到如下结果：

Table 3. TTR analysis results of *Wencheng* and *One Kind of Reality*
表 3. 《文城》《现实一种》TTR 分析结果

语料库	类符(type)	形符(token)	TTR
《文城》	10,132	107,632	0.094
《现实一种》	2434	16,060	0.152

通过表 3 可知，《文城》的 TTR 为 0.094，而《现实一种》的 TTR 为 0.152，《文城》的 TTR 远小于《现实一种》，由此可见相比于《现实一种》，《文城》的用词相对简单朴实，语言更加精炼准确。通过余华前后期作品在用词方面的对比，可以发现余华在后期创作时文字更加简洁，倾向于用质朴的文字呈现事件和人物。李佳贤(2024)指出在语言上，余华始终保持着先锋时期简洁、精准又充满张力的语言风格，并且有重叙事的倾向，几乎过滤掉一切不必要的描写和修饰[14]。通过检索可以更细节地发现《文城》中的语言不仅继承了先锋时期的语言特点，并在此基础上对文本用词加以锤炼，词汇表达更为精炼、准确。

其次，采用 Concordance Plot (索引定位)功能检索两个语料库中小说人物在文本中的分布状况。Concordance Plot 是 AntConc 软件用于检索的另一功能，用此功能可以直观地看到检索词在文本中出现的位置，条形图的蓝色部分越密集，表示该检索词出现的次数越多且频数也越高[15]。在 Plot 中有一个 Dispersion (分散性)指标，该指标通常表示聚类的紧密程度，Dispersion 越小，表示聚类程度越紧密，反之则越分散。本文选取两个小说文本中两位主人公作为检索词，得到其在小说中的基本分布情况。《文城》选取检索词“林祥福”和“小美”，《现实一种》选取检索词“山岗”和“山峰”(小说主人公名字)，检索结果如下：

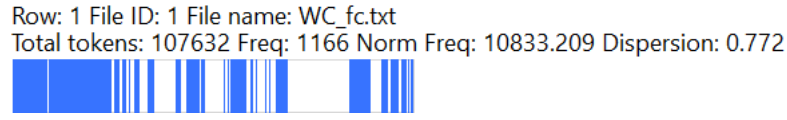


Figure 1. Plot search results for “Lin Xiangfu” in *Wencheng*
图 1. 《文城》“林祥福”Plot 检索结果

Row: 1 File ID: 1 File name: WC_fc.txt
Total tokens: 107632 Freq: 1271 Norm Freq: 11808.756 Dispersion: 0.679



Figure 2. Plot search results for “Xiao Mei” in *Wencheng*

图 2. 《文城》“小美” Plot 检索结果

Row: 1 File ID: 1 File name: XSYZ_fc.txt
Total tokens: 16060 Freq: 276 Norm Freq: 17185.554 Dispersion: 0.823



Figure 3. Plot search results for “Shan Gang” in *One Kind of Reality*

图 3. 《现实》“山岗” Plot 检索结果

Row: 1 File ID: 1 File name: XSYZ_fc.txt
Total tokens: 16060 Freq: 210 Norm Freq: 13075.965 Dispersion: 0.733



Figure 4. Plot search results for “Shan Feng” in *One Kind of Reality*

图 4. 《现实》“山峰” Plot 检索结果

通过图 1~4 可知,从条形图来看,《文城》“林祥福”和“小美”在文本中分布的情况较为集中,《现实一种》“山岗”和“山峰”在文本中分布的情况较为分散;从 Dispersion 来看,“林祥福”和“小美”的 Dispersion 分别为 0.772 和 0.679,而“山岗”和“山峰”分别为 0.823 和 0.733,《文城》的 Dispersion 结果整体低于《现实一种》。

由此可见,《文城》核心人物的叙事排布相对集中,作品更侧重于人物的刻画与情节的发展。反观《现实一种》,小说主要人物分散于各个叙事段落,人物出场密度较低,整体布局较为松散疏离。这一特征反映了在该时期余华的创作更倾向于对叙事细节的描述和把握。高玉(2023)指出余华早期的中短篇小说基本不刻画人物形象,那时的余华甚至是否定人物性格的[16]。同时,余华也在早期访谈中表明更倾向于关注人物的欲望,欲望比性格更能代表一个人的存在价值。通过检索数据发现,从《现实一种》到《文城》,余华开始大量着墨于人物形象的刻画。此外,通过比较两部作品主人公名称发现,《现实一种》中“山峰”“山岗”人物名称简单随意,类似于一种扁平符号;而《文城》中主要人物“林祥福”“小美”以及上文提及的“林百家”,甚至反面人物“张一斧”等名称都是作者精心安排的。可见,相比过去将人物简化为抽象符号,余华在《文城》中赋予了姓名丰富的象征内涵,其不再是指代人物的空洞标记,而与人物性格、命运息息相关。“祥福”暗含祥和善良的人物底色,反派“张一斧”则以直白粗暴的姓名隐喻其凶狠暴戾的本性。命名方式的转变直观印证了余华创作观念的转型:早年写作刻意剥离人物个体特质,以符号的形式印证人性恶;后期则重视人物的立体塑造。同时,现实主义最主要的特征是在典型环境中刻画典型人物,《文城》对人物形象细致地塑造印证了余华从“先锋”向“现实主义”的转型。

为了进一步分析余华小说写作风格的变化,此处使用 keyword 功能,对《文城》和《现实一种》文本进行关键词检索,整理归纳高频词中的名词(除人名、地名等)。检索《文城》主题词时,将《文城》文本作为目标语料库,将《现实一种》作为参照语料库;检索《现实一种》时,则以《文城》为参照语料库。得到如下结果。

由表 4 可知,《文城》多偏向日常生活,语言平实自然;而《现实一种》名词多为“争斗”类名词,场面暴力血腥,这一点与余华先锋时期文学创作中“暴力写作”不谋而合。周广学、姜耀东(2006)指出余

华前期的创作用语言建构了一个充满暴力、血腥和死亡的阴暗世界。余华后期的创作开始有所转变，作者开始以冷静的笔触关注着苦难中的人的存在[17]。这一点在《文城》中也有所体现，然而余华的转型并非完全放弃了过去的写作风格，在《文城》中也存在部分暴力描写，如“棺材”“枪支”等，但小说整体更偏向于人的书写，如与人生活相关的词汇“衣裳”“包袱”等。

Table 4. Keyword list of *Wencheng* and *One Kind of Reality*
表 4. 《文城》《现实一种》关键词表

作品名称	关键词
《文城》	婴儿、衣裳、包袱、兄弟、棺材、枪支
《现实一种》	血、拳头、骨头、血迹、腐烂

4. 结语

本文借助 CorpusWordParser、AntConc 等语料分析工具，从小说主题、人物形象、小说写作风格三个维度对《文城》展开系统性的文本分析与研究。研究表明，《文城》深刻诠释了敬畏生命、抗争超越苦难并完成生命救赎的核心主旨，生动印证了主人公林祥福温润善良、心怀赤诚的温情人物形象；同时也佐证了余华后期创作的鲜明特征——注重完整的情节架构与立体的人物塑造，语言质朴平实，其创作也从揭露人性恶转向情义建构，整体呈现出鲜明的现实主义创作特质。综上可见，语料库分析软件为文学文本研究提供了科学有效的研究方法与工具支撑，为文学文本的量化分析提供全新的研究路径。

参考文献

[1] 洪治纲. 悲悯的力量——论余华的三部长篇小说及其精神走向[J]. 当代作家评论, 2004(6): 20-37.

[2] 洪治纲. 寻找诗性的正义——论余华的《文城》[J]. 中国现代文学研究丛刊, 2021(7): 66-78.

[3] 屠舒一. 无法寻得的乌托邦——以余华的《文城》为例[J]. 余华研究, 2024(2): 122-127.

[4] 李彦姝. 《文城》的纯粹与简薄[J]. 文艺理论与批评, 2021(6): 86-92.

[5] 孟觉之, 胡小兰. 先锋作家的“出城”记——从《文城》看余华创作的再转型[J]. 南方文坛, 2021(6): 182-186.

[6] 王雁. 《傲慢与偏见》的语料库检索分析[J]. 江苏教育学院学报(社会科学版), 2007(2): 100-102.

[7] 胡开宝, 朱一凡. 基于语料库的莎剧《哈姆雷特》汉译文本中显化现象及其动因研究[J]. 外语研究, 2008(2): 72-80+112.

[8] 刘慧丽. 基于语料库的《红楼梦》“就是”探究[D]: [硕士学位论文]. 济宁: 曲阜师范大学, 2009.

[9] 陈婵, 程乐. 莫言作品《蛙》的语料库检索分析[J]. 浙江工商大学学报, 2014(5): 26-34.

[10] 朱洪. 基于语料库的《白鹿原》语言风格研究[J]. 今古文创, 2021(24): 26-27.

[11] 杨惠中. 语料库语言学导论[M]. 上海: 上海外语教育出版社, 2002.

[12] 曹洋. 文城无止, 缘入溪镇——基于象征叙事视角下的《文城》研究[J]. 名作欣赏, 2022(15): 122-124+166.

[13] 赵长志. 论《文城》中的林祥福形象[J]. 名作欣赏, 2023(12): 110-113.

[14] 李佳贤. “单纯”的辩证法——从《文城》反观余华的小说创作[J]. 当代文坛, 2024(2): 104-110.

[15] 金想琴. 基于 AntConc 软件对余华前后期作品写作风格的对比研究——以《世事如烟》和《活着》为例[J]. 海外英语, 2016(22): 172-173.

[16] 高玉. 余华长篇小说新作《文城》的承续与创新[J]. 河南大学学报(社会科学版), 2023, 63(1): 70-79+154.

[17] 周广学, 姜耀东. 试析余华写作风格的转变[J]. 辽宁行政学院学报, 2006(5): 224+226.