

基于Coh-Metrix的译林版与外研社版高中英语教材阅读语篇可读性比较研究

孙燕茹

苏州科技大学外国语学院, 江苏 苏州

收稿日期: 2026年9月8日; 录用日期: 2026年10月9日; 发布日期: 2026年10月16日

摘 要

本研究基于Coh-Metrix 3.0比较译林版与外研社版高中英语教材阅读语篇的可读性。选取两套教材必修第一至第三册、选择性必修第一至第二册核心阅读语篇100篇, 其中译林版40篇、外研社版60篇。采用描述性统计、独立样本t检验与典型语篇细读, 考察整体可读性、叙事性、句法简约性、词汇具体度及衔接等指标。结果显示, 外研社版Flesch Reading Ease和叙事性显著高于译林版, 而两版在Coh-Metrix L2 Readability、句法简约性、词汇具体度及多数衔接指标上差异不显著。两套教材从必修到选择性必修阶段均呈可读性下降, 且译林版下降幅度更大。个案分析表明, 译林版高难度语篇的负荷主要来自长句、抽象概念与跨段整合, 外研社版高叙事性语篇则依靠稳定人物和时间-事件链降低理解阻力。研究提示, 两版教材差异主要体现在表层可读性、叙事组织与阶段递进方式, 教学中应据此配置差异化阅读支架。

关键词

Coh-Metrix, 译林版, 外研社版, 高中英语教材, 阅读语篇, 可读性比较

A Comparative Study of the Readability of Reading Texts in Yilin and FLTRP Senior High School English Textbooks Based on Coh-Metrix

Yanru Sun

School of Foreign Languages, Suzhou University of Science and Technology, Suzhou Jiangsu

Received: September 8, 2026; accepted: October 9, 2026; published: October 16, 2026

Abstract

This study compares the readability of reading texts in Yilin and Foreign Language Teaching and Research Press (FLTRP) senior high school English textbooks using Coh-Metrix 3.0. A total of 100 core reading texts were selected from Compulsory Books 1~3 and Selective Compulsory Books 1~2, including 40 Yilin texts and 60 FLTRP texts. Descriptive statistics, independent-samples t tests, and close reading of representative texts were used to examine overall readability, narrativity, syntactic simplicity, word concreteness, and cohesion. The FLTRP series scored significantly higher in Flesch Reading Ease and narrativity, while the two series did not differ significantly in Coh-Metrix L2 Readability, syntactic simplicity, word concreteness, or most cohesion indices. Readability declined from the compulsory to the selective compulsory stage in both series, with a larger decline in Yilin. Case analyses showed that the higher processing load of difficult Yilin texts mainly stemmed from longer sentences, abstract concepts, and cross-paragraph integration, whereas highly narrative FLTRP texts relied on stable characters and temporal-event chains to reduce comprehension barriers. The findings suggest that the two series differ mainly in surface readability, narrative organization, and progression patterns across stages, calling for differentiated reading scaffolds in classroom instruction.

Keywords

Coh-Metrix, Yilin Press, Foreign Language Teaching and Research Press, Senior High School English Textbooks, Reading Texts, Readability Comparison

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

普通高中英语课程强调发展学生的语言能力、文化意识、思维品质和学习能力。阅读语篇作为英语学习的重要输入材料,不仅承载词汇、语法和语篇知识,也承载主题意义、文化内容和思维训练任务。

《普通高中英语课程标准(2017年版 2020年修订)》指出,英语课程内容包括主题、语篇、语言知识、文化知识、语言技能和学习策略等要素,语篇是表达意义的基本单位,也是学生发展语言能力的重要依托。因此,对高中英语教材阅读语篇进行可读性比较分析,有助于教师更准确地把握不同版本教材的文本特征,进而改进阅读教学设计。

Coh-Metrix是目前较有代表性的自动化文本分析工具。Graesser等人开发的Coh-Metrix能够从词汇、句法、衔接、可读性等多个层面对文本进行分析[1]。McNamara等进一步指出,Coh-Metrix不仅关注文本表层语言特征,也关注语篇内部的衔接、连贯和认知加工特征,因而适合用于教材语篇和二语阅读文本分析[2]。Crossley等的研究也表明,基于认知的可读性指标能够比传统公式更有效地解释二语阅读文本难度[3]。

不过,现有研究留下的空间需要进一步精确界定。林颖已对人教版与外研版高中英语必修教材进行过Coh-Metrix比较[4],说明不同版本教材之间的横向比较并非空白;邢文骏则从译林版教材与高考阅读的匹配角度考察文本难度[5],肖金美进一步关注新课改背景下高中英语阅读文本复杂度及课程阶段递进问题[6]。相较而言,译林版与外研社版在同一课程标准框架下、覆盖必修与选择性必修多个册别的直接

比较仍缺少系统证据,已有量化研究对显著差异如何落实到具体句法、词汇和语篇组织上的解释也相对有限。因此,本研究不仅比较两套教材的总体可读性,还考察其阶段递进,并通过典型语篇细读解释量化结果,以期更清楚地呈现不同教材版本的语言输入特征及其教学意义。

2. 文献综述

2.1. 教材评价与阅读语篇比较研究

教材是课程理念转化为课堂教学实践的重要媒介。语言教材不仅提供语言输入,也通过主题选择、语篇类型、活动设计和任务安排影响学生语言能力的发展。教材评价应关注教材目标、学习者需求、语言内容、技能训练和教学可操作性等方面。Tomlinson 指出,教材开发和教材评价既具有理论属性,也具有实践属性,教材是否适合具体学习者和教学情境,是评价过程中必须考虑的问题[7]。

从比较研究角度看,不同版本教材虽然都依据课程标准编写,但在单元设计、文本选择、板块功能和语篇组织方式上可能存在不同取向。教材比较研究的价值不在于简单判断某一版本教材“更好”或“更难”,而在于通过比较揭示不同版本教材的语言输入特点和教学使用重点。对于高中英语阅读教学而言,比较不同版本教材阅读语篇的可读性,有助于教师理解教材文本的难度层次、体裁特征和语篇组织差异,从而更有针对性地使用教材。

2.2. 文本可读性研究

可读性通常指文本易于阅读和理解的程度。阅读理解不仅受文本本身影响,也受读者语言水平、背景知识和阅读策略影响。因此,可读性并不是单纯的文本属性,而是文本、读者和阅读任务共同作用的结果。尽管如此,文本自身的词汇、句法和衔接特征仍是影响阅读理解的重要因素。

早期可读性研究主要依靠公式计算。Flesch Reading Ease 是较常用的传统可读性指标之一,其核心依据是平均句长和平均音节数[8]。这类公式操作方便,适合快速判断文本表层难度。但其局限也比较明显:它无法判断语篇是否衔接连贯,无法反映因果、转折、指称关系等语篇内部结构,也不能充分解释二语学习者在阅读加工中面临的具体困难。

Bailin 和 Grafstein 认为,可读性研究应超越词长和句长,从文本与语境关系、语言结构、读者背景等方面综合考虑[9]。Crossley 也指出,传统可读性公式过于依赖文本表层特征,而 Coh-Metrix 等基于认知的指标能够更好地反映二语阅读文本的复杂性[3]。这说明,研究高中英语教材阅读语篇可读性时,仅使用传统公式是不够的,还需要进一步考察语篇层面的深层特征。

2.3. Coh-Metrix 在阅读语篇比较研究中的应用

Coh-Metrix 是一种基于自然语言处理的自动文本分析工具,能够计算多种语言和语篇指标。Graesser 等人最早介绍 Coh-Metrix 时指出,该工具可以分析文本在衔接、语言和可读性等方面的特征[1]。后来,Graesser 等进一步提出 Coh-Metrix 可以在多个语言和话语层面分析文本特征,其中包括叙事性、句法简约性、词汇具体度、指称衔接和深层衔接等重要维度[1]。McNamara 等的专著系统介绍了 Coh-Metrix 的理论基础、指标体系和应用场景,为阅读材料评价、教材分析和文本难度研究提供了方法依据[2]。

在国内研究中,Coh-Metrix 已逐渐应用于外语阅读语篇分析。江进林、韩宝成使用 Coh-Metrix 比较大学英语六级、托福和雅思阅读文本,说明不同类型考试文本在可读性和语篇特征上存在差异[10]。王艳伟比较专业英语教材与测试阅读文本的易读度,指出教材和测试文本的总体难度变化并不总是完全一致,需要结合多项指标综合解释[11]。邢文骏使用 Coh-Metrix 3.0 中的可读性指标和文本易读性主成分指标比较译林版高中英语教材与高考阅读语篇,发现教材与考试材料在整体可读性、句法、词汇和衔接等维度

上存在不同程度的差异[5]。这些研究共同提示：教材可读性研究不宜只回答“难不难”，还需要解释“难在哪里”以及这种难度是否与具体学习阶段和教学目标相匹配。

近几年，Coh-Metrix 及相关文本分析工具在高中英语教材研究中的应用进一步增多。林颖以人教版和外研版高中英语必修教材阅读语篇为对象，比较叙事性、句法简约性、词汇具体度和指称衔接等维度[8]，其研究将关注点从“教材 - 考试匹配”推进到“教材版本之间的横向差异”。肖金美则从词汇、句法和语篇维度讨论新课改背景下高中英语阅读文本复杂度[6]，强调课程阶段内部的难度递进值得单独考察。由此可见，当前高中英语教材可读性研究已经形成“多指标分析 - 版本比较 - 阶段递进”的发展趋势，但对译林版与外研社版跨多个课程阶段的直接比较，以及对量化差异的语篇层面解释仍有进一步深化空间。

从理论上讲，教材语篇难度的意义还需要放在二语习得和教材编写框架中理解。Tomlinson 强调教材评价应考虑学习者需求、材料吸引力和教学适切性[7]，因此统计意义上的“更易读”并不能直接等同于“更优”。Krashen 的可理解输入观点提示，语言输入应具有可理解性，同时保留适度超出学习者当前水平的挑战[12]；Carrell 和 Eisterhold 的图式理论则表明，阅读理解是文本信息与读者既有背景知识相互作用的过程[13]，不同体裁、主题和叙事组织可能通过不同方式激活或增加背景知识需求。基于这些理论，教材可读性比较应进一步讨论难度递进是否可被支架化、叙事性是否有助于激活图式，以及不同文本组织方式如何服务于课程阶段的学习目标。

2.4. 研究述评

现有研究仍存在四点可进一步推进之处。第一，横向比较对象仍相对有限：林颖比较了人教版与外研版必修教材[4]，但尚未覆盖译林版与外研社版在必修和选择性必修多个册别中的系统比较。第二，部分研究主要关注教材与高考等外部评价文本的匹配，对两套教材自身语言输入取向的直接比较不足。第三，已有研究已经注意到课程阶段递进问题，但版本差异与阶段递进往往被分开讨论，尚缺少在同一研究框架中同时考察“版本间横向差异”和“版本内纵向变化”的分析。第四，现有研究总体仍以量化结果呈现为主，对显著差异如何具体体现为长难句、抽象词汇、衔接线索和语篇组织方式，缺少可核查的典型文本分析。

基于此，本研究以译林版和外研社版高中英语教材为对象，选取必修第一册至必修第三册、选择性必修第一册至选择性必修第二册中的核心阅读语篇，借助 Coh-Metrix 3.0 比较两套教材的整体可读性与文本易读性主成分，并进一步考察两套教材从必修阶段到选择性必修阶段的可读性递进。在此基础上，本研究选取典型语篇进行解释性细读，将量化差异与具体语言特征相联系。相较于已有研究，本研究的增量主要体现在三个方面：扩大版本与课程阶段的比较范围；在同一框架中整合横向版本比较与纵向阶段比较；以质性语篇分析解释量化指标，从而增强结果对教材编写和课堂教学的可解释性。

3. 研究设计

3.1. 研究问题

本研究围绕以下四个问题展开：

第一，译林版与外研社版高中英语教材阅读语篇在整体可读性指标上是否存在显著差异？

第二，译林版与外研社版高中英语教材阅读语篇在文本易读性主成分上呈现出哪些异同？

第三，两套教材从必修阶段到选择性必修阶段的阅读语篇可读性递进是否存在差异？

第四，典型语篇中的句法、词汇和衔接组织如何解释两套教材在可读性与叙事性上的量化差异？

3.2. 研究对象

本研究选取译林版和外研社版普通高中英语教材各五册作为研究对象，具体包括必修第一册、必修

第二册、必修第三册、选择性必修第一册和选择性必修第二册，共计十册教材。为保证样本具有可比性，本研究仅选取两套教材中的核心阅读板块。译林版选取 Reading 和 Extended reading 板块语篇，共 40 篇；外研社版选取 Understanding ideas 和 Developing ideas 板块语篇，共 60 篇。由于两套教材单元数量不同，样本量并不完全一致。本研究采用“同类阅读板块全样本采集”的方式，不人为删减外研社版语篇，以保留教材本身的真实结构。

文本整理时，每篇语篇单独保存为 txt 文件，并进行统一清洗。文本清洗原则包括：删除题目、页码、题号、中文注释、练习指令、图片说明、词汇表、语法讲解和活动说明，只保留英文阅读正文；段落之间保留一个空行，段首不缩进，以减少格式差异对 Coh-Metrix 分析结果的影响。

3.3. 研究工具与指标

本研究使用 Coh-Metrix 3.0 作为文本分析工具。结合研究目的和已有研究，本研究主要选取两类指标。

第一类是整体可读性指标，包括 Flesch Reading Ease 和 Coh-Metrix L2 Readability。Flesch Reading Ease 主要基于词长和句长反映文本表层可读性，数值越高，文本越容易阅读。Coh-Metrix L2 Readability 则更适合二语阅读文本分析，其计算不仅关注词句特征，也考虑学习者阅读加工中的语篇因素。

第二类是文本易读性主成分指标，包括叙事性、句法简约性、词汇具体度、指称衔接、深层衔接、动词衔接、连接度和时序性。叙事性反映文本是否接近故事、人物经历和生活化表达；句法简约性反映句法结构是否简单；词汇具体度反映实义词的具体或抽象程度；指称衔接反映词汇、概念和论元的重叠程度；深层衔接反映因果、逻辑等深层关系；动词衔接反映事件和动作之间的连续性；连接度反映连接词使用情况；时序性反映时间顺序和时态关系。

3.4. 数据处理

本研究首先将所有语篇导入 Coh-Metrix 3.0 并导出指标数据。随后，对译林版和外研社版文本数据进行编码，编码信息包括教材版本、册别、课程阶段、单元、板块和语篇标题。整体研究采用量化分析为主、质性语篇细读为辅的解释性混合分析思路，即先通过量化结果识别稳定差异，再用典型个案解释差异在具体语言形式和语篇组织上的表现[14]。

数据分析分为四个步骤。第一，进行描述性统计，计算两套教材各项指标的平均值和标准差。第二，采用独立样本 t 检验比较译林版与外研社版阅读语篇在各项指标上的差异；由于两套教材样本量不同且并非一一配对关系，方差齐性不满足时参考 Welch 校正结果，并结合均值差和 Cohen's d 判断差异的实际幅度。第三，将样本进一步分为必修阶段和选择性必修阶段，分别考察两套教材内部的可读性递进情况。第四，在量化结果基础上进行目的性个案选择。译林版选取选择性必修第二册 Unit 4 Reading 中的 Artificial intelligence: friendly or frightening? (YL-选择性 B2-U4-R)：其 Coh-Metrix L2 Readability 为 8.32，在译林版 40 篇中最低；Flesch Reading Ease 为 51.65，为第 4 低；FKGL 为 11.30，为第 2 高；平均句长为 21.12 词，为第 3 长，能够集中呈现译林版选择性必修阶段的较高加工负荷。外研社版选取必修第一册 Unit 1 Developing ideas 中的 High School Hints (WY-B1-U1-DI)：其叙事性为 2.10、时序性为 1.04，均为外研社 60 篇中的最高值，同时 Flesch Reading Ease 为 80.71、L2 Readability 为 22.75，适合用于解释外研社版总体叙事性较高的统计结果。质性分析聚焦四个维度：句法层面关注长句、从属结构和信息嵌套；词汇层面关注抽象概念词与动作/经验表达；衔接层面关注指称复现、逻辑连接和时间线索；语篇层面关注人物-事件链、说明结构和观点组织。两个个案分别服务于不同的核心量化发现，并非用于一对一匹配比较；个案分析用于解释量化趋势，而不适用于替代或扩大总体统计推断。

4. 结果与讨论

4.1. 两套教材阅读语篇整体可读性比较

本研究首先比较两套教材在 Flesch Reading Ease 和 Coh-Metrix L2 Readability 两项整体可读性指标上的表现。结果见表 1。

Table 1. Comparison of overall readability of reading texts in the two textbook series
表 1. 两套教材阅读语篇整体可读性比较

指标	译林版 M ± SD	外研社版 M ± SD	p 值	结果
Flesch Reading Ease	66.34 ± 11.54	72.20 ± 8.20	0.007	差异显著
Coh-Metrix L2 Readability	16.85 ± 5.15	18.30 ± 3.50	0.127	差异不显著

从表 1 可以看出，外研社版教材在 Flesch Reading Ease 上的均值为 72.20，高于译林版的 66.34，且差异达到显著水平($p=0.007$)。按照两组均值和标准差估算，效应量 Cohen's d 约为 0.61，属于中等幅度。这说明外研社版阅读语篇在由平均句长和词形长度所反映的表层语言特征上整体更易读，差异不仅具有统计意义，也具有一定实际幅度。

然而，在 Coh-Metrix L2 Readability 指标上，两套教材差异并不显著($p=0.127$)，相应效应量约为 $d=0.34$ ，明显小于 Flesch 指标。传统公式与二语认知型可读性指标出现这种“方向一致但强度不同”的结果，说明版本差异主要集中在较表层的句词负荷，而未同步扩展为综合性的二语加工差异。该发现与 Crossley 等关于认知型可读性指标能够补充传统公式局限的观点相吻合[3]：单纯依赖 Flesch Reading Ease，容易把句子较短、词形较简单的文本直接解释为总体阅读负担更低，而忽略衔接、语义和学习者加工因素。

因此，更稳妥的解释不是“外研社版整体更容易”，而是“外研社版在表层可读性上具有中等幅度优势，但两套教材在综合二语可读性上未检测到显著差异”。这一结论也提示教材评价应避免将单一可读性指标转化为版本优劣判断，而应进一步观察差异来自何种语篇类型和语言组织方式。

4.2. 两套教材阅读语篇文本易读性主成分比较

为进一步考察两套教材在深层语篇特征上的差异，本研究比较了八项文本易读性主成分。结果见表 2。

Table 2. Comparison of text easability principal components in the two textbook series
表 2. 两套教材阅读语篇文本易读性主成分比较

指标	译林版 M ± SD	外研社版 M ± SD	p 值	结果
叙事性	0.11 ± 0.76	0.45 ± 0.74	0.028	差异显著
句法简约性	0.01 ± 0.56	0.04 ± 0.41	0.809	差异不显著
词汇具体度	0.42 ± 0.82	0.57 ± 0.76	0.337	差异不显著
指称衔接	-0.46 ± 0.63	-0.57 ± 0.59	0.379	差异不显著
深层衔接	0.29 ± 0.62	0.44 ± 0.69	0.251	差异不显著
动词衔接	0.15 ± 0.72	0.11 ± 0.68	0.786	差异不显著
连接度	-2.18 ± 0.99	-2.14 ± 0.97	0.830	差异不显著
时序性	-0.51 ± 0.90	-0.57 ± 0.95	0.754	差异不显著

从表 2 可以看出, 八项文本易读性主成分中, 只有叙事性指标存在显著差异。外研社版叙事性均值为 0.45, 显著高于译林版的 0.11 ($p = 0.028$), 效应量约为 $d = 0.45$, 属于小到中等幅度。叙事性较高意味着文本更常通过具体人物、事件、经历和情境组织信息, 因此读者更容易建立“谁 - 做了什么 - 发生了什么变化”的事件模型。结合图式理论看, 这类组织方式也更可能激活学习者已有的生活经验与事件图式, 从而为理解提供自上而下的支持[13]。

相比之下, 译林版叙事性均值较低, 说明其阅读语篇整体上较少持续依赖人物和事件链推进意义。这与其部分册别, 尤其是选择性必修阶段较多涉及文化阐释、传媒、科技、医学、艺术等信息性或论述性主题相吻合。此类文本往往要求学生在较短篇幅内处理概念界定、分类说明、观点关系和文化背景, 其理解难点不一定表现为句法显著更复杂, 而可能表现为抽象概念密度和知识结构要求更高。

在句法简约性方面, 两套教材差异不显著(译林版 0.01, 外研社版 0.04), 效应量也很小。这说明两套教材总体句法复杂度并不存在稳定的版本分化。值得注意的是, “总体不显著”并不意味着具体语篇中不存在长难句, 而是意味着长句、复合句和修饰结构的分布没有形成足以区分两个版本的系统性差异。因此, 教学中仍需要针对具体文本识别句子信息层级, 而不能依据教材版本预设句法难度。

在指称衔接、深层衔接、动词衔接、连接度和时序性方面, 两套教材均未呈现显著差异, 且效应量总体较小(绝对值约 0.04~0.23)。这说明版本间最稳定的差异并不在一般性的衔接密度, 而更集中在语篇如何“讲述”信息。换言之, 两套教材都能够维持相对规范的概念复现、逻辑连接和事件组织, 但外研社版更常把这些衔接资源嵌入人物与事件链中, 而译林版更多服务于说明、解释和观点组织。

总体而言, 表 2 呈现的是“局部差异而非全面分化”: 叙事性具有可重复解释价值, 其他主成分在本样本中未形成显著版本差异。这一结果既避免了把两个版本简单二分为“容易/困难”, 也为后续典型语篇分析提供了明确方向, 即重点考察叙事组织、抽象概念与句内信息结构如何共同塑造阅读体验。

4.3. 必修与选择性必修阶段阅读语篇可读性递进分析

教材阅读语篇的难度不仅需要在不同版本之间保持合理差异, 也应在课程阶段内部形成可接受的递进。选择性必修阶段承担更高层次的语言运用、文化理解与思维训练任务, 因此难度上升本身并非问题; 更值得关注的是上升幅度是否平滑、是否能够在适当教学支架下保持可理解性。基于这一考虑, 本研究进一步比较两套教材从必修阶段到选择性必修阶段的可读性变化。

Table 3. Comparison of readability between compulsory and selective compulsory stages in the Yilin series
表 3. 译林版必修与选择性必修阶段可读性比较

指标	必修阶段 $M \pm SD$	选择性必修阶段 $M \pm SD$	p 值	结果
Flesch Reading Ease	71.53 $\pm$ 10.71	58.56 $\pm$ 7.94	<0.001	显著下降
Coh-Metrix L2 Readability	19.19 $\pm$ 5.09	13.36 $\pm$ 2.72	<0.001	显著下降
叙事性	0.39 $\pm$ 0.77	-0.32 $\pm$ 0.50	0.001	显著下降
句法简约性	0.07 $\pm$ 0.64	-0.07 $\pm$ 0.44	0.393	差异不显著
连接度	-2.40 $\pm$ 1.09	-1.86 $\pm$ 0.76	0.075	差异不显著

从表 3 可以看出, 译林版教材从必修阶段到选择性必修阶段的可读性下降非常明显。Flesch Reading Ease 从 71.53 下降至 58.56, Coh-Metrix L2 Readability 从 19.19 下降至 13.36, 且均达到显著水平。这意味着难度上升同时发生在表层句词负荷与综合二语加工两个层面。相较于单一指标下降, 两项可读性指标同步下降更能说明学生进入选择性必修后可能面临较明显的加工跨度。

同时，译林版叙事性从 0.39 下降至-0.32，且差异显著。也就是说，难度提升与语篇组织方式的变化同时发生：文本一方面更难读，另一方面又减少了人物、事件和生活情境等可用于建立事件模型的线索。结合教材主题，选择性必修阶段涉及饮食文化、音乐、绘画、诗歌、传媒、体育赛事、医学进展和科技突破等内容，这些主题往往要求学生处理概念关系、文化解释和观点结构。因而，译林版的阶段递进更接近“语言负荷上升 + 叙事支持减弱”的复合变化，这也是其教学支架需求更突出的原因。

Table 4. Comparison of readability between compulsory and selective compulsory stages in the FLTRP series
表 4. 外研社版必修与选择性必修阶段可读性比较

指标	必修阶段 $M \pm SD$	选择性必修阶段 $M \pm SD$	p 值	结果
Flesch Reading Ease	75.36 ± 8.07	67.47 ± 5.88	<0.001	显著下降
Coh-Metrix L2 Readability	19.25 ± 3.42	16.86 ± 3.18	0.008	显著下降
叙事性	0.66 ± 0.71	0.14 ± 0.68	0.006	显著下降
句法简约性	0.09 ± 0.43	-0.03 ± 0.38	0.271	差异不显著
连接度	-0.32 ± 0.93	-0.94 ± 0.85	0.011	显著下降

从表 4 可以看出，外研社版同样呈现从必修阶段到选择性必修阶段的可读性下降。Flesch Reading Ease 从 75.36 下降至 67.47，Coh-Metrix L2 Readability 从 19.25 下降至 16.86，差异均达到显著水平。这表明外研社版也通过课程推进增加阅读挑战，但两项指标的下降幅度均小于译林版，整体递进相对平缓。

具体而言，译林版 Flesch Reading Ease 下降约 12.97，Coh-Metrix L2 Readability 下降约 5.83；外研社版分别下降约 7.89 和 2.39。两套教材因此呈现不同的输入梯度：译林版在阶段转换处提高难度的幅度更大，外研社版则更接近连续递增。这一发现与肖金美关于应关注必修与选择性必修之间复杂度递进的研究取向相呼应[10]，但本研究进一步显示，同样的课程阶段递进在不同教材版本中并非同幅度发生。

此外，外研社版叙事性也从 0.66 下降至 0.14，说明选择性必修阶段同样减少了故事化和生活化表达；但其选择性必修阶段的叙事性仍高于译林版对应阶段。这意味着外研社版即使提高概念与语言要求，仍较多保留人物、事件和情境作为理解支点。就二语输入而言，这种“难度上升但叙事支点部分保留”的组织方式可能有助于维持阅读过程的连续性，但并不意味着其认知要求更低，因为主题理解、价值判断和批判性思维仍可能在叙事表层之上提出更高要求。

4.4. 典型语篇的质性案例分析

为进一步解释量化结果，本研究选择两个具有明确量化定位的语篇进行定向细读。译林版个案 Artificial intelligence: friendly or frightening? 用于回答“选择性必修阶段的高阅读负荷在真实文本中如何表现”；外研社版个案 High School Hints 则用于回答“高叙事性在真实文本中由哪些语言和语篇机制实现”。因此，两篇文本并非作为同册、同体裁的配对样本进行优劣比较，而是分别作为两个核心统计发现的解释性案例，实现由“量化差异是什么”向“差异如何在文本中形成”的推进(表 5)。

译林版个案呈现出较典型的说明——论述型信息组织。该文围绕人工智能的定义、发展历史、现实应用及潜在风险逐层展开，语义重心不断在技术概念和观点判断之间转换。句法上，定义段中的长句同时容纳中心谓语、后置修饰以及多个并列非谓语成分；风险讨论中又出现 while 引导的从属结构与 where 引导的定语从句，要求读者在一个句子内维持多个信息单元。词汇上，artificial intelligence、computer intelligence、deep learning、capacity、technology、turning point、existence 等抽象或专业概念连续出现，增加了概念建构负荷。值得注意的是，文本并非缺少显性衔接：In 1997、Then in 2017、Nowadays、In short、

While、Apparently 等标记能够提示局部关系；但其宏观结构主要是“定义 - 历史 - 应用 - 风险”的说明链，而非由稳定人物主体贯穿的持续事件链。这与该篇较低的叙事性(-0.48)、指称衔接(-1.43)和时序性(-1.01)相呼应。结合 L2 Readability 最低、平均句长较长等指标可以看出，该篇的阅读负荷不仅来自句长，还来自抽象概念、句内嵌套与跨段信息整合的叠加。

Table 5. Qualitative comparison of two representative reading texts
表 5. 两篇典型语篇的质性比较

个案语篇	量化定位	文本细读证据	对总体结果的解释
译林版: Artificial intelligence: friendly or frightening? (选择性必修二 Unit 4 Reading)	L2 Readability = 8.32 (40 篇最低); Flesch = 51.65 (第 4 低); FKGL = 11.30 (第 2 高); 平均句长 = 21.12 词(第 3 长); 叙事性 = -0.48; 指称衔接 = -1.43; 时序性 = -1.01。	技术与抽象概念密集; 长句中包含后置修饰、并列非谓语及从属/定语从句。宏观按“定义→历史→应用→风险”推进, 虽有年份、Nowadays、In short 等显性标记, 但缺少持续稳定的人物 - 事件链。	译林版后期的阅读负荷不仅与句长有关, 也与抽象概念建构和跨段信息整合有关; “有连接词”并不等同于指称/时序支点充分或叙事性高。
外研社版: High School Hints (必修一 Unit 1 Developing ideas)	叙事性 = 2.10 (60 篇最高); 时序性 = 1.04 (60 篇最高); Flesch = 80.71; L2 Readability = 22.75; 平均句长 = 12.41 词; 词汇具体度 = -1.23。	以 Lisa 为稳定人物中心, 采访问答将信息分块; I/my 持续维持第一人称视角, first day、looking back、when、next time 等形成时间推进, joined、tried、failed、learned 等动作/经历词构成事件链。	高叙事性主要由人物、经验和时间组织产生, 而非词汇具体度普遍更高。这解释了外研社版叙事性显著较高而两版词汇具体度总体差异不显著的并存结果。

外研社版个案则显示, 高叙事性主要由人物、经验和时间组织共同形成。全文以 Lisa 这一稳定人物主体为中心, 通过采访者提问和 Lisa 第一人称回答将信息分解为若干局部单元; I、my 等第一人称形式持续维持话语中心。文本中的 first day、looking back、when、next time 等时间表达把“进入高中 - 面对挑战 - 经历失望 - 给出建议”组织为可追踪的经验序列, joined、helped、tried、failed、learned 等动作和经历动词进一步强化事件模型。其平均句长为 12.41 词, Flesch Reading Ease 为 80.71, 局部加工入口相对清晰; 叙事性(2.10)和时序性(1.04)均为外研社 60 篇中的最高值。尤其值得注意的是, 该篇词汇具体度主成分为-1.23, 句法简约性为-0.33, 说明“高叙事性”并不等同于所有词汇更具体或所有句法更简单。它更多体现为稳定的人际角色、第一人称经验、问答分块以及时间 - 事件链所提供的情境支点。这一现象也有助于解释为什么外研社版总体叙事性显著较高, 而两版在词汇具体度和句法简约性上的总体差异却并不显著。

4.5. 结果的理论解释与已有研究对话

综合量化结果与典型语篇分析可以发现, 两套教材的差异并非“全面难度差异”, 而是不同语言输入组织方式的差异: 外研社版在表层可读性和叙事性上更具优势, 译林版则在选择性必修阶段表现出更明显的难度跃升; 与此同时, 多数深层衔接和句法、词汇主成分并未形成稳定的版本分化。

第一, 本研究中 Flesch Reading Ease 显著不同而 Coh-Metrix L2 Readability 不显著, 进一步印证了传统可读性公式与认知型指标不能相互替代。Crossley 等指出, 二语可读性需要纳入词汇、句法、衔接和认知加工等因素[3]。本研究的结果说明, 两套教材在句长和词形层面存在可检测差异, 但这种差异并未扩展为所有深层加工维度的同步分化, 因此仅凭传统公式判断版本整体难度容易过度简化。

第二, 阶段递进可以从可理解输入的角度重新解释。Krashen 强调学习者需要接触能够理解、同时略具挑战性的语言输入[12]。因此, 译林版选择性必修阶段难度提升更明显并不能直接视为缺点, 关键在于

这种跨度是否有足够的背景知识、词汇预处理、句法支架和任务分解来支持学生完成加工。外研社版递进相对平缓,则体现出另一种材料组织逻辑,即在增加主题和思维要求的同时,更多保留叙事与情境支点。两种路径都可能合理,但对课堂支架的需求不同。

第三,叙事性差异与图式激活具有直接联系。Carrell 和 Eisterhold 认为,阅读理解依赖文本信息与读者既有图式之间的互动[13]。外研社版较高的叙事性能够通过稳定人物、个人经历、动作和时间序列为学生提供较容易调用的经验图式;译林版较多的信息性、科技性或论述性语篇则更依赖内容图式和概念知识。如果背景知识不足,学生即使面对句法主成分并未显著更复杂的文本,也可能因缺少概念支点而感到困难。High School Hints 的个案进一步说明,叙事优势并不必然来自词汇更具体:该篇词汇具体度主成分并不高,但凭借人物-事件-时间结构仍取得最高叙事性和时序性。这说明叙事性应主要从语篇情境模型的建构方式理解,而不能与“简单词汇”或“简单句法”等同。

第四,本研究与国内已有高中教材研究形成了较清晰的递进关系。林颖的研究已证明不同版本高中英语必修教材可以通过 Coh-Metrix 进行横向比较[8],本研究在此基础上加入译林版并将范围扩展到选择性必修阶段;邢文骏关注译林版教材与高考文本的外部匹配[11],本研究则转向教材版本之间的内部比较;肖金美提出课程阶段复杂度递进值得关注[10],本研究进一步发现这种递进具有版本特异性。由此,本研究的理论贡献不在于判定哪一版本“更好”,而在于说明教材可读性应被理解为“文本特征-课程阶段-学习者加工-教学支架”之间的动态适配关系,这也与 Tomlinson 强调教材适切性的观点相一致[6]。

4.6. 教学启示

基于上述结果,教学启示不宜停留在一般性的“增加支架”或“关注衔接”,而应与不同版本的具体数据特征相对应。

第一,针对译林版,尤其是选择性必修阶段,应重点处理“难度跃升+叙事支持减弱”的组合问题。读前可建立“主题-核心概念-背景知识”微型概念图,并提前处理少量决定文本理解的抽象词汇,而不是一次性讲解全部生词;读中可要求学生先圈出谓语动词,再按主句、从句和修饰成分切分长句,同时用流程图、论证图或段落功能标签呈现信息结构;读后逐步撤除支架,让学生用自己的语言重构文本逻辑、完成摘要、评价或迁移任务。这样既降低初始加工负荷,又保留选择性必修阶段应有的认知挑战。

第二,针对外研社版,应把较高叙事性从“理解优势”转化为“高阶思维入口”。对于 High School Hints 这类访谈和人物经历型文本,教师可以先让学生沿“人物-经历/问题-应对-结果/建议”建立事件线,利用第一人称经验和时间线索快速形成整体理解;随后通过“某条建议建立在什么经历或证据上”“如果换成另一位学生的经历,建议是否仍然成立”“文本如何从个人经验上升为一般性观点”等问题,把课堂从内容复述推进到证据评价和观点迁移。还可设计建议排序、情境改写、立场辩论或 Claim-Evidence-Reasoning 任务,使学生在较易进入的叙事情境上开展更深层的主题意义建构和批判性思考。

第三,两套教材都需要显性开展语篇衔接教学。由于指称衔接、深层衔接、动词衔接和连接度等指标总体差异不显著,说明衔接能力更适合作为跨版本的共同教学目标。教师可让学生标注代词指代、关键词复现、因果与转折连接词、时间线索,并进一步讨论“删除某个衔接词后意义如何变化”“同一关系还能怎样表达”,使学生从识别语言形式走向理解语篇关系。

第四,教材使用应从“依据版本预设难度”转向“依据文本证据调节任务”。对于可读性较低且抽象度较高的文本,应增加背景和结构支架;对于叙事性较强、表层较易进入的文本,则应相应提高分析、评价和迁移任务的认知要求。换言之,阅读材料的语言难度与课堂任务的思维难度可以形成补偿关系:文本越难,支架越应清晰;文本越易进入,越需要通过高阶任务防止学习停留在表层理解。

5. 结语

本研究基于 Coh-Metrix 3.0, 并结合典型语篇细读, 对译林版和外研社版高中英语教材阅读语篇的可读性进行了比较。量化结果显示, 外研社版在 Flesch Reading Ease 和叙事性上显著高于译林版, 说明其在表层句词负荷和叙事情境组织方面更有利于学生进入文本; 然而, 两套教材在 Coh-Metrix L2 Readability、句法简约性、词汇具体度以及多数衔接指标上差异不显著, 表明版本间并不存在全面、稳定的深层难度分化。

阶段比较进一步发现, 两套教材均体现出从必修到选择性必修的难度递进, 但递进方式不同。译林版的 Flesch Reading Ease 和 L2 Readability 下降幅度更大, 同时叙事性明显降低, 呈现出较为集中的阶段跃升; 外研社版也提高阅读挑战, 但整体递进较平缓。典型语篇分析为这种总体趋势提供了可观察的语言证据: 译林版 Artificial intelligence: friendly or frightening? 的较高加工负荷同时体现为较长句子、抽象技术概念和说明-论述型信息整合; 外研社版 High School Hints 的高叙事性则主要由稳定人物主体、第一人称经历、问答分块和时间-事件链实现。两个个案同时提示, 可读性与叙事性不应被简单还原为单一词汇或句法特征, 而应从多层次语篇组织及其对学习者的支持方式加以理解。

本研究的意义主要体现在三个方面: 其一, 在同一框架中整合译林版与外研社版的横向比较以及必修-选择性必修的纵向比较; 其二, 将传统可读性指标、Coh-Metrix 多维指标与质性语篇分析结合, 避免仅凭 p 值或单一公式解释教材难度; 其三, 从可理解输入、图式激活和教材适切性的角度解释版本差异, 将统计结果进一步转化为差异化教学策略。由此, 教材可读性更适合被理解为文本特征、学习阶段、学习者背景与教学支架之间的适配问题, 而非版本优劣的单一排序。

当然, 本研究仍存在一定局限。首先, 研究对象仅包括两套教材各五册中的核心阅读语篇, 不能完全代表教材的全部语言输入。其次, 质性分析仅选取两个典型个案, 主要承担解释量化趋势的功能, 尚不足以概括全部体裁内部差异。再次, Coh-Metrix 与文本细读均主要关注文本本身, 无法直接反映不同语言水平学生的真实阅读时间、理解正确率、认知负荷和课堂互动。后续研究可扩大典型语篇样本, 结合学生阅读测试、眼动或阅读时长数据、教师访谈和课堂观察, 并进一步检验不同支架策略是否能够缓冲版本和阶段之间的难度差异。

参考文献

- [1] Graesser, A.C., McNamara, D.S., Louwerse, M.M. and Cai, Z. (2004) Coh-Metrix: Analysis of Text on Cohesion and Language. *Behavior Research Methods, Instruments, & Computers*, **36**, 193-202. <https://doi.org/10.3758/bf03195564>
- [2] McNamara, D.S., Graesser, A.C., McCarthy, P.M. and Cai, Z. (2014). Automated Evaluation of Text and Discourse with Coh-Metrix. Cambridge University Press. <https://doi.org/10.1017/cbo9780511894664>
- [3] Crossley, S.A., Greenfield, J. and McNamara, D.S. (2008) Assessing Text Readability Using Cognitively Based Indices. *TESOL Quarterly*, **42**, 475-493. <https://doi.org/10.1002/j.1545-7249.2008.tb00142.x>
- [4] 林颖. 高中英语教材阅读语篇可读性研究: 以人教版和外研版为例[D]: [硕士学位论文]. 重庆: 四川外国语大学, 2025.
- [5] 邢文骏. 基于 Coh-Metrix 的高中英语教材与高考阅读语篇难度比较研究[J]. 中小学英语教学与研究, 2025(2): 71-76.
- [6] 肖金美. 新课改下中国高中英语教材的阅读文本复杂度研究[J]. 现代语言学, 2025, 13(6): 588-597.
- [7] Tomlinson, B. (2012) Materials Development for Language Learning and Teaching. *Language Teaching*, **45**, 143-179. <https://doi.org/10.1017/s0261444811000528>
- [8] Flesch, R. (1948) A New Readability Yardstick. *Journal of Applied Psychology*, **32**, 221-233. <https://doi.org/10.1037/h0057532>
- [9] Bailin, A. and Grafstein, A. (2001) The Linguistic Assumptions Underlying Readability Formulae: A Critique. *Language & Communication*, **21**, 285-301. [https://doi.org/10.1016/s0271-5309\(01\)00005-2](https://doi.org/10.1016/s0271-5309(01)00005-2)

- [10] 江进林, 韩宝成. 基于 Coh-Metrix 的大学英语六级与托福、雅思阅读语篇难度研究[J]. 中国外语, 2018, 15(3): 86-95.
- [11] 王艳伟. 专业英语教材与测试阅读文本易读度比较研究[J]. 外语研究, 2021, 38(2): 70-75.
- [12] Krashen, S.D. (1985) *The Input Hypothesis: Issues and Implications*. Longman.
- [13] Carrell, P.L. and Eisterhold, J.C. (1983) Schema Theory and ESL Reading Pedagogy. *TESOL Quarterly*, **17**, 553-573.
<https://doi.org/10.2307/3586613>
- [14] Creswell, J.W. and Plano Clark, V.L. (2018) *Designing and Conducting Mixed Methods Research*. 3rd Edition, Sage.