

Trinka与Grammarly和 LanguageTool在学术写 作上的比较

● Enhancement

It was a declaration by
President Nixon four decades
~~previously~~ → ago.



概要：

- 1 有一些英语语法检查器可用于日常英语写作，但很少有专门针对学术的英语写作、术语、科学惯例和出版准备等相关问题的检查工具。
- 2 我们在覆盖率五个学术领域的258个英文句子中测试了Trinka，一个由人工智能驱动的学术写作助手。我们还将Trinka与Grammarly Premium和LanguageTool在相同的句子上进行比较，用以测试纠正的准确性和对各种语言错误类别的覆盖率。
- 3 我们的结果显示，Trinka在深层语言和特定学科的用法和术语问题上的表现远远好于Grammarly和LanguageTool。
- 4 Trinka在语法和风格问题上的表现与Grammarly相当，而LanguageTool的表现并不理想。
- 5 在三个语法检查器中，Trinka是最适合学术写作的，因为它在准确性和覆盖面之间有很好的平衡。以及它独有的出版准备检查的特点。



版权声明：www.trinka.ai 的所有内容都受到国际版权法的保护。本网站的任何部分都不得以任何形式或手段进行复制、存储在检索系统中或传输。

背景

英语是国际交流的语言，也是无可争议的科学交流的通用语言。在学术界、商界和教育界的一系列以职业为导向的活动中，使用成语和标准英语语法是一项必不可少的技能。良好的写作技巧尤其与研究人员的职业前景息息相关。为了获得成功，使用英语作为第二语言（ESL）或英语作为附加语言（EAL）的科学家和研究人员面临着在国际英语期刊上发表作品的压力。一方面，他们必须努力解决阻碍其研究成功发表的语言和文化问题¹。另一方面，学术出版业也在不断寻找解决方案，以管理不断增加的投稿数量和ESL/EAL学者的出版需求，以及对他们写作的认知偏见^{2,3}。最常见的情况是，这些研究人员与学术同行、以英语为母语的同事或语言服务公司联系，使他们的论文准备好发表。

虽然有大量的语法和拼写检查器供普通作家使用，但几乎没有任何针对于研究人员的产物。学术写作有别于一般写作。这种写作主要是客观的、精确的和结构化的；使用正式的语气的；包含特定领域的术语和引文的；并且是为特定的学术团体准备的。大多数语法检查器至少是要检查拼写，但最多也就是检查超微观的语法。然而，简单的语法和拼写检查对于提交给期刊发表的稿件是不够的。我们非常需要一个写作助手，它不仅可以提供特定学科的语言纠正和改进，还可以帮助研究人员处理出版方面的具体要求，如遵守风格指南、一致性检查和字数限制等。



英语语法检查器

目前，市场上出现了一些针对一般英语文本的语法检查器。例如，Grammarly Premium⁴和LanguageTool⁵。Grammarly是一个世界领先的语法纠正应用程序，使其自然成为这次比较的对象。LanguageTool是一个被广泛采用的开源校对工具，提供语法和拼写纠正。尽管这些工具在不同程度上是成功的，但它们并不是为学术内容设计的。

为了解决这一问题，Trinka被开发出来，以改善学术英语的使用和技术词汇，并为出版做准备。

Trinka自动纠正语法、拼写、用法和领域特定的错误，旨在帮助学者以清晰、简明和吸引人的方式用英语交流他们的研究。



分析

为了分析Trinka的修正能力以及它与其他语法检查器的区别，我们比较了Trinka、LanguageTool和Grammarly在五个学术领域中随机抽取的258个句子（5778个单词）的表现（表1）。






领域	句子 (n = 258)	字数 (n = 5778)
 生命科学	56	1170
 工程学	47	1172
 经济学	41	923
 社会科学和人文科学	57	1540
 医学	57	973

表1. 按领域计算的句子和单词数

另外，我们将检测并修改的内容分为以下几类。领域、语法、用法、拼写、标点符号和风格。每个类别的描述见表²（如下）。

类别	描述
专业领域	技术用词，科学惯例，符号使用等。
语法	名词数量、动词形式、主谓一致、冠词等。
词语用法	词语/短语的选择、语法、冗余、习惯用语等。
拼写	打字误，混淆的地方
标点符号	逗号、分号、冒号、连字符等。
学术风格	美英风格的拼写和标点符号，大写字母，间距，等等。

表2. 错误类别

分析

这258个句子是用Grammarly、Trinka和LanguageTool的在线编辑器编辑的。编辑工作由一位拥有12年经验的专业学术编辑进行评估。258个句子中总共有437个错误。表3(如下)显示了三个工具提供的正确和不正确的修改建议与结果，以及被它们遗漏的存在的错误数量。

工具	正确	不正确	遗漏
Grammarly	203	52	234
Language Tool	37	7	400
Trinka	223	44	214

表3. 各系统的编辑计数

结果

我们的分析结果显示，总体而言，Trinka和Grammarly涵盖了表2中所示的所有类别。Trinka在领域、用法和风格方面修正的最多，而Grammarly则在语法和标点符号方面做了最多的修正。LanguageTool则远远落后。每个系统的总体准确性得分为：Trinka 51%；Grammarly 46.4%；LanguageTool 8.4%。

类别	Trinka	Grammarly	LanguageTool
专业领域	15	3	0
语法	77	89	6
语词用法	12	4	0
拼写	9	9	9
标点符号	22	44	4
学术风格	58	51	13

表4. 按类别划分的修正数

对TRINKA、GRAMMARLY和LANGUAGETOOL的修改结果的分析

我们的结果表明，Grammarly和Trinka不但在正确和错误的筛查量方面取得了相当的结果。在正确和错误的修改编辑量方面也取得了几乎相等的结果。因此，我们需要再深入调查一下，以检查这些修改的特性。

表4显示，虽然Grammarly比Trinka纠正了更多的语法、拼写和标点符号错误，但Trinka纠正了更多与学术领域、风格和语言使用方面相关的错误。LanguageTool在所有类别中都只做了很少的修正。下面是一些由Trinka修改的句子的例子。这些是没有被Grammarly或LanguageTool正确纠正的。

学术领域

Geology:

Along with giving an overview of oil and gas exploration theories, oil accumulation laws in the **Ordos** Mesozoic **Ordos** Basin are discussed. [Technical word choice]

The domain term was incorrectly ordered.

Medicine:

The pre-pathogenic course of **novel Covid COVID-19** as well as the risk factors and medical intervention has not yet been determined. [Redundancy]

The term 'novel Covid-19' has been corrected to *its official term COVID-19*.

语词用法

Medicine:

Observing the nutritional status of pregnant women and **keeping** **maintaining** their health is important for the success of this venture.

[Word/phrase choice]

The word 'keeping' is not idiomatic in the context, so it has been changed to 'maintaining'.

Law:

The present proceeding is an appeal from a judgement of the Federal Magistrates Court ('FMC') of 17th August 2005 dismissing the appellants' applications for review **on the basis of** the Anshun estoppel. [Word/phrase choice]

The use of 'basis' in place of 'on the basis of' or 'based on' is usually seen in marketing content. Here Trinka replaces it, keeping intact the formal tone of the sentence.

对TRINKA、GRAMMARLY和LANGUAGETOOL的修改结果的分析

语法

Physics:

Four KB mirrors were pressed on two conical reference cones by ball plungers and then **be** aligned in the tangential and sagittal directions. [Verb form]

This sentence has one subject and two predicates—"were pressed..." and "be aligned...". They are not grammatically parallel. The revision corrects the predicate.

Engineering:

In modern times it demands more efficient methods for the oil/water and **emulsions** **emulsion** separation. [Noun number]

When plural nouns are used to describe other nouns, the singular form is generally used. There are a few exceptions. However, in this sentence, the noun 'emulsions' should have been 'emulsion'.

Physics:

We study metal ionophores for **the property of** anticancer **properties** in both in vivo and in vitro cancer. [Syntax]

The phrasing is ungrammatical as the prepositional object is an adjective. The object is usually a noun, pronoun, or an adjective that works as a noun (e.g., the affluent).



接下来，我们检查了Trinka和Grammarly错误地标记或完全遗漏的错误（表3）。我们在这里没有考虑LanguageTool，因为它的最终改修完成的数量很少。我们发现，正如被修改的错误的特质不同一样 这两个系统中，错误的类型也不同。Grammarly的错误编辑的前3个类别是文章（11），单词/短语选择（11）和语法（8）。Trinka的前3个类别是冠词（13）。逗号（4）和名词数量（3）。关于遗漏的项目，Grammarly的前3个类别是文章（38）、单词/短语选择（23）和逗号（14），而Trinka的分类是冠词（53），逗号（28）和名词数量(16)。

超越单纯的语法改错，Trinka的学术写作特点

除了语法检查，Trinka和Grammarly都提供了一些增强功能 诸如对简洁性、一致性和正式语体的修改。这样的 增强功能，虽然不是更正，但可以提高写作的质量。在这方面，我们发现Trinka提供了更多的增强功能，不仅针对上述三个类别，而且还针对包容性/精准正确的的语言和风格指南的遵守。我们在下面展示几个例子

Psychiatry:

~~The study by~~ Menzo et al. found that 15 of 83 patients with Parkinson's disease fulfilled the criteria for anxiety disorder.
[Conciseness]

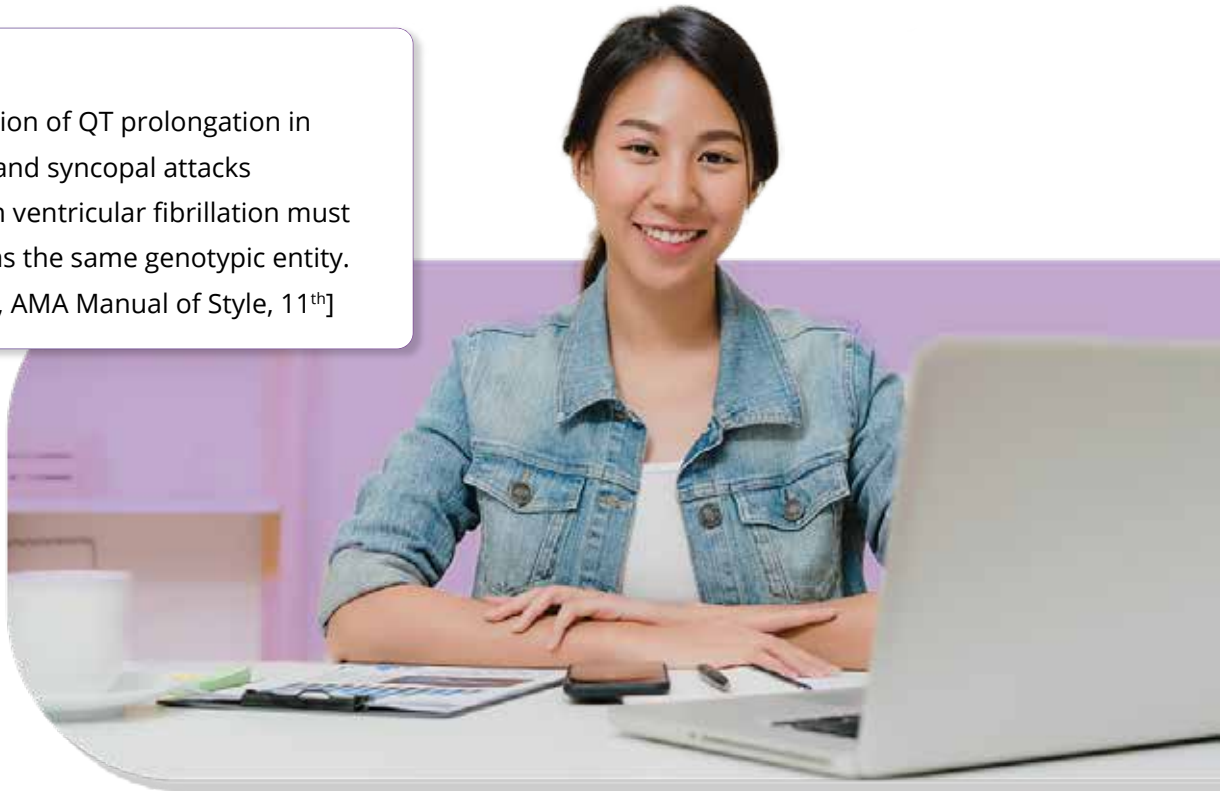
Medicine:

Individuals ~~confined to~~ using a wheelchair cannot board a plane on their own; hence, special procedures and infrastructure are required to enable their travel. [Inclusive/Unbiased language]

除了技术术语外，许多学术出版物还遵循特定的风格指南。这种风格指南对书面文本的各个方面都有严格要求。例如，数学符号前后是否应该有空格 (例如， $2+2=4$ vs. $2+2=4$)，数字及其相关单位作为修饰语时是否应该用连字符 (4mm条 vs. 4-mm条)，等等。下面，我们展示一个基于《美国医学会（AMA）风格手册》第11版的修订实例。

Medicine:

The combination of QT prolongation in the ~~EKG ECG~~ and syncopal attacks resulting from ventricular fibrillation must be regarded as the same genotypic entity.
[Abbreviation, AMA Manual of Style, 11th]



结论

我们在Trinka、Grammarly Premium和LanguageTool上测试了五个领域的258个句子。我们发现Trinka是最适合学术写作的英文写作助手。因为它是最准确的，并且涵盖了广泛的写作问题。它也最适合处理期刊或书籍的出版要求，这一点在其他语法检查工具中是看不到的。这在其他语法检查器中是看不到的，因此它完美完成了作为转学及学术写作的助手的角色。

另外，句子和词语可以被不同专业的编辑以不同方式解释。来自不同学科领域的读者和普通读者对句子和单词的解释可能也有所不同。因此，更进一步，需要由更多来自不同领域、具有不同交际目标的评估者进行进一步的评估。

交际目标，以减少评价中的任何主观性。

在任何机器学习系统中，都有一个已知的召回率和精确度之间的权衡。召回率是指在所有编辑中预测到的正确编辑的百分比。精度是指正确的编辑数量占总编辑数量的百分比。为提高召回率而进行的任何调整往往会导致对精度的负面影响。

反之亦然。一个自动写作助手在多大程度上掌握这种微妙的平衡，决定了该助手的价值。我们的结果清楚地表明Trinka提供了最佳的平衡，无论是在其覆盖率和准确性之间，还是在语言和领域的错误纠正之间。总之，Trinka在三个语法检查器中表现最好，并且很好地适应了跨领域的学术和专业写作。



参考资料

- 1 Flowerdew, J. (2019). The linguistic disadvantage of scholars who write in English as an additional language: Myth or reality. *Chest*, 52, 2, 1728–1730. doi:10.1378/chest.129.6.1728
- 2 Englander, K. (2014). Writing and publishing science research papers in English. *SpringerBriefs in Education*. doi:10.1007/978-94-007-7714-9
- 3 Walker R, Barros B, Conejo R., Neumann, K., & Telefont, M. (2015) Personal attributes of authors and reviewers, social bias and the outcomes of peer review: a case study [version 2; peer review: 2 approved]. *F1000Research*, 4, 21. doi: 10.12688/f1000research.6012.2
- 4 Trink AI. (2020, May 17). <https://www.trinka.ai>
- 5 Grammarly Premium. (2020, May 17). <https://www.grammarly.com/premium>
- 6 LanguageTool. (2020, May 17). <https://languagetool.org>
- 7 Machine Learning Crash Course. (2020, May 20). <https://developers.google.com/machine-learning/crash-course/classification/precision-and-recall>

