

人工智能助力海外中文教学的新探索

——利用人工智能生成多语境高生字重复率的泛读文章

泓河中文学校人工智能辅助教学课题组

论文摘要：

本文探讨了利用人工智能（AI）生成多语境、高生字重复率的泛读文章，使泛读在扩展的语境中紧密配合精读。泓河中文学校在 17 年的教学实践中发现，三、四岁就开始学中文的海外华裔学生在 4-6 年级面临学习瓶颈，常因汉字遗忘速度加快，进步变缓而对中文学习失去信心甚至产生抵触情绪。AI 可以通过生成更优质的泛读文章来提高学生的学习兴趣和生字刷新频次，提高学习效率，帮助我们更加深入地实现“字不离词，词不离句，句不离段，段不离篇”的教学理念，并进一步延伸“篇反照字词”的方法。我校的“泓河开卷™”项目初步试验了使用 AI 工具（ChatGPT 和 Gemini、Kimi）辅助生成泛读文章，自 2023 年 10 月开始在我校四年级进行试验，效果得到学生和家长的认可。我校人工智能团队研发的线上阅读平台“泓河开卷™”已经于 2024 年 5 月向我校四年级老师和学生开放试用。

关键词：人工智能 紧密配合精读的泛读文章 多语境 高生字重复率

一、AI 工具的定位

在中文教学中使用人工智能（AI），我们并不是为了赶时髦或玩花活，而是为了解决实际存在的海外中文教学难题。当前市场上存在许多商业“人工智能”教学工

具，但我们的目标与他们当中的一些不同。我们并不打算用 AI 替代教师，而是希望通过 AI 的辅助，提升教师的教学能力，帮助教师更有效地开展教学工作。AI 可以为教师提供更多的资源和工具，使他们能够更好地应对教学中的各种挑战，从而提高教学效果和学生的学习体验。通过这种方式，AI 不仅不会取代教师，反而会成为教师的一双翅膀，助力他们飞得更高、更远。

二、海外中文教学阅读环节为什么要引入 AI?

泓河中文学校自 2007 年成立以来的 17 年中，已经为 10365 人-年的海外学生提供中文教学（注：一个学生参加 3 年的课程计入 3 人-年），这些学生的年龄涵盖三岁学前班至十二年级。在教学实践中我们观察到，从三、四岁开始学中文的海外华裔学生，大约在 4-6 年级，有不少学生要过一个“坎”：三年级之前，他们通常对学习中文还是颇有兴趣的，但从四年级开始对继续学习中文产生抵触，到六年级抵触情绪达到最高峰，这段时间，一部分学生放弃，另一部分学生的信心锐减。四年级后，学习中文的速度与此前大体一样，但学生忘记已学汉字的速度在不断加快。产生这种现象的原因很多，比如这个年龄段课外活动的增加造成对中文学习时间的挤压；长久以来学生对拼音的依赖造成的脱离拼音之后的文盲状态；海外缺乏中文阅读环境，学过的字词很少有机会在忘记之前通过反复见面来刷新巩固；适合海外学生的阅读资料非常稀缺；还有一个重要的原因是，中文教材中的课后泛读篇目对生字的重复率不够。

我们对中国暨南大学编写的《中文》四年级的几篇课后阅读文章做了测试，

数据见表 2.1:

表 2.1: 课后阅读中的生字重复率

| | |
|---------------|--|
| 课后阅读篇目 1 | 第一课: 给亮亮的回信 |
| 本课需要掌握的生字 | 信封城邮票京广安场全平提取进孙 (生字总数: 15) |
| 生字在阅读文章中的重复次数 | “封”:0,“城”:0,“邮”:0,“票”:0,“广”:0,“安”:0,“场”:0,“全”:0,“取”:0,“孙”:0,“京”:1,“平”:1,“提”:1,“进”:1,“信”:3 (出现的生字个数: 5/15) |
| 课后阅读篇目 2 | 第四课: 太阳和彩虹 |
| 本课需要掌握的生字 | 晒照处寻熟服暖才健康 (生字总数: 10) |
| 生字在阅读文章中的重复次数 | “晒”:0,“照”:0,“处”:0,“寻”:0,“熟”:0,“服”:0,“暖”:0,“才”:0,“健”:0,“康”:0 (出现的生字总数: 0/10) |
| 课后阅读篇目 3 | 第九课: 聪明的爱迪生 |
| 本课需要掌握的生字 | 数空靠仰颗久清连勺离努力著 (生字总数: 13) |
| 生字在阅读文章中的重复次数 | ”数”:0,“空”:0,“靠”:0,“仰”:0,“颗”:0,“久”:0,“清”:0,“连”:0,“勺”:0,“离”:0,“努”:0,“力”:0,“著”:0 |

| | |
|---------------|--|
| | (出现的生字个数: 0/13) |
| 课后阅读篇目 4 | 第十二课: 人造卫星 |
| 本课需要掌握的生字 | 宇航员次球实类梦号达泥非世界(生字总数: 14) |
| 生字在阅读文章中的重复次数 | "字":0,"航":0,"员":0,"次":0,"实":0,"梦":0,"号":0,"达":0,"泥":0,"非":0,"世":0,"界":0,"球":1,"类":1 (出现的生字个数: 2/14) |

海外中文学生可以接触到的适合自身阅读水平的中文资料比较有限，课后的阅读资料是一个非常重要的阅读来源，所以能让学生在有限的时间内读到生字高度重复率的泛读文章对海外学生学习中文非常重要。AI 可以帮助广大教师和学生生成更优质的泛读文章，使其在扩展的多种语境中紧密配合精读课文。AI 的使用不仅可以使生字在文中高频次出现，增加学生和生字见面机会，也可以通过多语境的设定，大幅度提高学生的学习兴趣。

三、引入 AI 工具之前尝试过的的解决方案

在尝试引入 AI 工具之前，我校曾经试验四个不同版本的方法帮助学生过这个 4-6 年級的“坎”，但总体效果有限。

第一个尝试是办“快速识字”班，试图通过迅速掌握一定数量的汉字使学生可以

开始自主阅读。类似这种明确以认字为目的的教学方法，在很多中文辅助材料和应用软件中也有体现。虽然能在短期内体现出较为明显的效果，但其弊端也比较明显：脱离了句、段、篇的支撑突击学习生字，学生难以真正了解汉字的用法和语境，很难达到长期掌握、结合语境使用自如的目的。而且在一段时间后，汉字掌握量进入瓶颈期，进步变缓，这也影响了学生对中文的深层次的理解和掌握。根据我们在海外教学的实践，对于不少学生来说，仅仅使用这个方法不足以规避中文学习不进反退，从而产生挫败感的情况。

第二、第三个尝试是办中文图书馆和编写阅读教材，虽然自己办的图书馆注意选择购买儿童读物，主观上注意编写贴近儿童的阅读材料，但由于内容与精读课本内容脱节，两者几乎没有配合，学过的字词和语言点被“稀释”在阅读材料中，难以起到巩固精读语言知识的目的。在英文社会环境中，这样的中文阅读材料难以深入到学生的阅读习惯中，起到巩固、积累语言知识的效果，在初级阶段，这种挑战显得尤其突出。单从词汇而言，无论是图书馆里的中文藏书，还是电子书或有声读物，都存在同样的一个问题：无法和华裔学生正在学习的精读课本内容环环相扣，达到举一反三的复习目的。

第四个尝试是引入课时内的泛读教学，并且参考借鉴了国内在中文泛读方面的经验。我们对照国内分级阅读的参考书目，和华裔学生的中文课本进行了比较，意识到两者之间存在着一定的差距。对于国内的孩子来说，由于从小生活在中文的语境之中，而且国内教育设置的语文课频次较高，所以学生在较小的年龄段就已经过了“识字”的关卡，在四年级的时候可以开始独立阅读较长的篇目，甚至中短篇的

小说。对于海外华裔学生来说，只有极少数的中文水平能够达到这一程度。此外，一些极具中国背景的阅读材料，华裔学生在理解方面也存在一定挑战。

我们曾把比较贴近海外学生生活的泛读材料引入课内。效果是有的，但除了疫情打断了有关试验之外，也遇到三个困难：1、安省政府支持的的国际语言课时目前是每周 2.5 小时，太少了，拿不出课时来做课内泛读；2、如果另外开设泛读课，会被教育局误解为与之竞争，产生敌意；3、普通的泛读材料，内容过于“稀释”，与精读内容结合太松散，积累效应不佳。

这一系列的尝试使我们意识到：1、海外中文教学需要泛读来配合精读教育，巩固精读知识；2、需要泛读内容更加“浓缩”，与精读教材配合更紧密；3、需要更深入更有渗透力地贯彻“字不离词、词不离句、句不离段、段不离篇”的教学理念。

四、AI 怎么用？

我们在阅读材料中有意识地高频次地使用刚刚学过的字词，让 AI 使用这些字词“造文”（类似于造句，但造出来的是适合学生阅读的整篇短文）。精读课本每学一课，创作组老师借助 AI 工具，让 AI 造出短文，通过“泓河开卷™”工具推送给学生阅读，在短文中包括刚学的生字，并同时尽量多地使用之前学过的字，尽量不用或少用还没学过的字（没学过的字一旦在文中出现会标注拼音），充分利用华人学生能听懂汉语的优势，通过字不离词、词不离句、句不离段、段不离篇的阅读过程，不断巩固并拓展中文认读能力。

为生成配合每课精读课文的高质量课外阅读文章，创作老师给 ChatGPT 按下面几点写一个详细的指示：

1. **明确目标：**说明你想生成的是一篇适合小学生阅读的文章。
2. **指定关键词：**列出文章需要使用的字，通常是当时所学课文后面的生字表所列生字。
3. **设定主题和情节：**给出文章的大概情节和主题。
 - a. 选择一个学生感兴趣的主体，如动物、魔法、冒险、友情等；通常与课文主题相近，也可以适当延伸；
 - b. 确定主要角色，或者设计几个有趣的角色，比如小动物、勇敢的小孩、魔法师等等；（这个部分也可以由 AI 自主完成）
 - c. 设计故事情节：开始介绍背景，中间设计一个问题去解决，结尾解决问题并给出一个温暖的结局；（这个部分也可以由 AI 自主完成）
4. **语言风格和长度：**
 - a. 使用简单易懂的语言，适合儿童阅读。
 - b. 故事长度要适中，不要太长，适合儿童的注意力跨度。

在 ChatGPT 生成故事之后，把文章导入泓河开卷™平台测试所包含的未学汉字总量比例和未学汉字用量比例。

创作老师需手动对文章进行打磨和润色。老师需要把未学汉字比例控制在一定范围内，以降低学生理解难度，并且不能影响文章的可读性和知识性。老师也需

要检验文章的语境是否能提高学生的阅读兴趣，尽量提高已学汉字的重复率，以增加学生和生字不断见面的机会。对于非虚构类文章，老师还要检查文章的真实性。老师也要检查文章的文学性，以期能够让学生体会到文章的文字美和文学美。

受限于目前 ChatGPT 的创作能力，这个人工干预过程一般要重复进行好几次直到文章达到创作老师满意的效果。希望今后 ChatGPT 的功能改进之后，或其他更强的 AI 工具出现之后，这个过程可以减少直至取消。

当完成 AI 泛读文章之后，再利用线上阅读平台泓河开卷™ 把文章推送给学生。图 4.1 为阅读平台学生端口显示的文章，学生们可以点开文章阅读。



图 4.1 阅读平台学生端口截图

下面的示例泛读文章以《中文》第四册第十二课《月亮上有什么》为例，包括无拼音版本和加注拼音版本（仅未学汉字加注拼音），学生在平台界面上阅读时，可通过图中箭头指向的开关随时自行选择阅读哪个版本。文中还配有主题插图，以此增加故事的趣味性，帮助学生们理解故事内容。故事插图也是由 ChatGPT 工具 DALL.E 生成。生成配图同样需要给指令，通常是简单概括故事内容，主要人物，以及主要事件。

AI 泛读文章示例

第十二课 月亮上有什么

生字：宇，航，员，次，数，球，实，类，梦，号，达，泥，非，世，界

未学汉字总量比 12% 未学汉字用字比 18%



泓河开卷™
Canada

Unread
Achieve

月亮上有什么

Need Pinyin?

同学们，今天我们来学习关于月球的一些小知识。



月球是地球的唯一一个天然卫星，它离我们的地球有38万公里远。如果用宇宙飞船过去的话，大概要三天的时间才能到达。当你看向夜空中明亮的月亮时，你能想象宇航员曾经走在上面吗？

月球上没有空气和水，所以那里没有声音和天气变化。月球表面布满了灰色的尘土和泥土，这些都是小行星撞到月球时形成的。这些泥土非常细，宇航员把它叫做“月尘”。如果你在月球上走一走，会留下非常清楚的脚印，就像在地球上走过一样。

第一次有人登上月球是在1969年，通过美国的“阿波罗11号”飞船实现的。那次任务中，宇航员在月球上留下了人类的第一步，还带回了月球上的石块和泥土进行研究。

在那以后，世界上的很多国家都对月球很感兴趣。科学家们想通过研究月球，更好地了解宇宙的形成和地球的起源。现在，探索月球仍然是宇航员们的梦想，他们要在将来建立月球基地，为进一步探索太空做工作。

月球是我们夜空中一个那么远又那么近的星球，是人类探索宇宙的第一步。希望有一天，你们中的一些人也能成为宇航员，亲自踏上月球，探索更多的未知世界！

图 4.2 无拼音版泛读文章

泓河开卷™
FUNDING FOR THE
GOVERNMENT OF ONTARIO
Ontario's Official Language Institute | Canada

Unread
Achieve

月亮上有什么

Need Pinyin?

同学们，今天我们来学习关于月球的一些小知识(shì)。



月球是地球的唯(wéi)一个天然卫(wèi)星，它离我们的地球有38万公里远。如果用宇宙(zhòu)飞船过去的话，大概(gài)要三天的时间才能到达。当你看向夜空中明亮的月亮时，你能想象(xiǎng)宇航员曾(céng)经走在上面吗？

月球上没有空气和水，所(suǒ)以那里没有声音(yīn)和天气变化。月球表(biǎo)面布满了灰(huī)色的尘(chén)土和泥土，这些都是小行(xíng)星撞(zhuàng)到月球时形(xíng)成的。这些泥土非常细，宇航员把它叫做“月尘(chén)”。如果你在月球上走一走，会留(liú)下非常清楚(chú)的脚印(jiǎoyìn)，就像在雪地上走过一样。

第一次有人登上月球是在1969年，通过美国的“阿波罗(Abóluó)11号”飞船实现的。那次任(rèn)务中，宇航员在月球上留(liú)下了人类的第一步，还带回了月球上的石块和泥土进行(xíng)研究(yánjiū)。

在那以后，世界上的很多国家都对月球很感兴趣。科学家们想通过研究(yánjiū)月球，更好地了解(jiě)宇宙的形(xíng)成和地球的起源(yuán)。现在，探索(tànsuǒ)月球仍然是宇航员们的梦想，他们要在将来建(jiàn)立月球基(jī)地，为进一步探(tàn)索太空做工作。

月球是我们夜空(kōng)中一个那么远又那么近(jìn)的星球，是人类探(tàn)索宇宙的第一步。希(xī)望有一天，你们中

图 4.3 加注拼音版本的泛读文章

每篇泛读文章后面还附有与文章内容相配的问题（图 4.4），检查学生对文章内容的理解程度，其结果汇总到老师界面加以显示。



图 4.4 泛读文章的问答题

总体来看，这篇 AI 泛读文章不仅包含了精读课文中刚学过的生字，大量使用了过去学过的汉字，还向学生们科普了关于月亮的小知识，非常适合认字量在这个程度的学生阅读，更重要的是能帮助他们巩固记忆所学生字，扩展了解生字在不同语境中的用法，使精读与泛读学习紧密配合。

五、效果如何？

自 2023 年 10 月份开始，我们为四年级三个班的学生提供了紧密配合《中

文》第四册每一篇课文的泛读短文。试验之初，刚刚脱离三年级的学生突然进入大篇的没有拼音辅助的阅读并不适应。

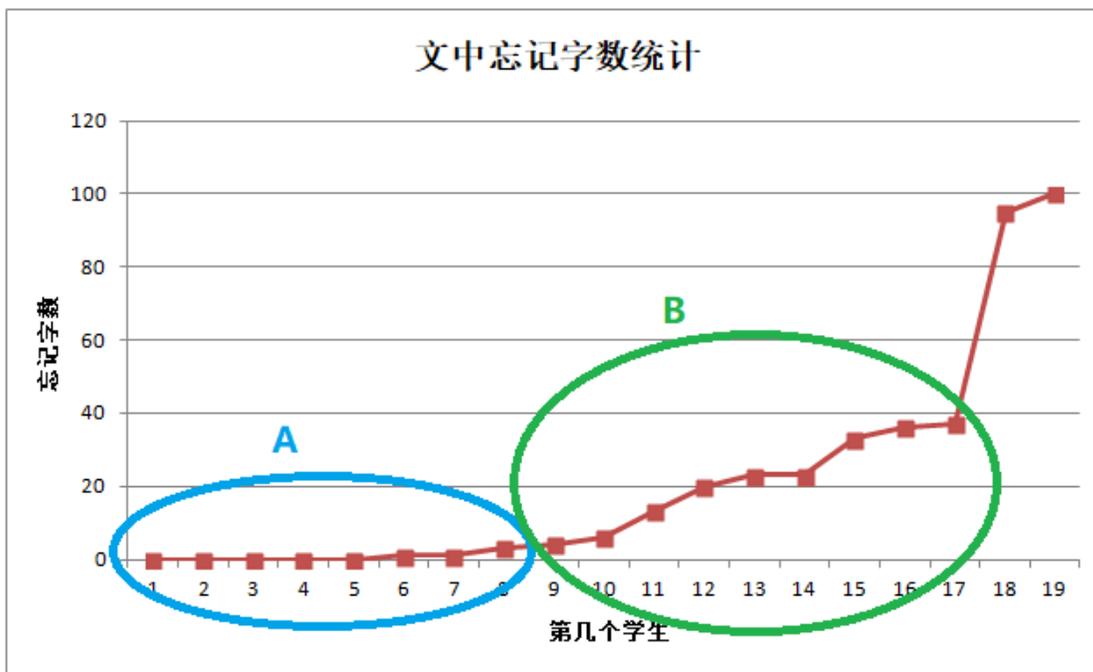


图 5.1 学生忘记字数统计

图 5.1 是开始试验之初，在一个班对当天出勤的 19 人做的随堂统计，在配合第二课的泛读短文中，有超过三分之一的学生（A），1-3 年级学过的字几乎全部还认识，这部分同学在阅读中并不感到吃力，他们在 AI 课外阅读中巩固学过的字，同时借助泛读拓展还没学过的字；另外近一半同学（B），前三册学过的字开始遗忘不认识了。阅读起来感到不同程度的吃力。这反而说明，这一阅读练习对于这部分同学尤为重要，通过它可以把忘记的字在阅读中逐步捡回来，是十分有益的。这些同学的确需要在课时内，在老师辅导下逐步进入泛读角色，尤其需要在老师带领下克服因教材中长期标注拼音，离了拼音就不认识字的弊端。剩下的个别学生几乎

所有的字都不认识，这些学生不是从小一路跟随我校进入四年级的，而是中途才加入这个班的。对于这些学生，他们需要补过去的课，而不是借助 AI 泛读文章来强化学习。

在做了整个学年的阅读练习之后，四年级的学生们开始逐渐适应了这些没有标注拼音的泛读文章，并从中获益。我们在学年结束前，对两个班级当天出勤的 31 位学生进行了问卷调查，了解他们的情况。还有六位同学参加了小组访谈，为部分问题提供了更详细的回答。

从调查结果来看，学生的阅读能力在这一年中得到了提高。表 5.1 显示，三分之二的学生已经能独立完成这些泛读文章的阅读。但仍有将近三分之一的学生，需要在同学或是老师家长的帮助下才能完成阅读。这一结果与学年初的随堂统计相比较，学生的阅读水平是有所提高的。

表 5.1: 学生能否独立完成阅读

| | 人数 | 百分比 |
|-------------------|----|------|
| 能独立完成 AI 泛读文章的阅读 | 21 | 68% |
| 觉得文章有点难，希望和同学一起完成 | 3 | 10% |
| 需要家长/老师的陪伴才能完成阅读 | 5 | 16% |
| 未回答 | 2 | 6% |
| 总人数 | 31 | 100% |

针对《中文》第四册的 12 篇课文，我们利用 AI 为每篇课文都生成了两到三篇

泛读文章。超过 80% 的学生每周至少能读完一篇文章，甚至有学生表示每天看一篇都可以（见表 5.2）。这显示了学生对课外阅读的接受度较高，阅读频率也较为理想。我们今后可以进一步优化内容，以保持和提高学生的阅读兴趣。

表 5.2: 学生每周能完成的阅读量

| | 人数 | 百分比 |
|----------------|----|------|
| 可以一天看完一篇泛读文章 | 6 | 19% |
| 每周可以看 3-4 篇 | 4 | 13% |
| 每周可以看 1-2 篇 | 16 | 52% |
| 一篇 AI 泛读文章都看不完 | 5 | 16% |
| 总人数 | 31 | 100% |

而从阅读习惯来看（表 5.3），目前只有将近 20% 的学生会主动寻找课外阅读材料，半数学生仍需家长提供文章才会阅读。由此可见，学生们虽然对 AI 课外阅读的接受度在逐渐提高，但主动性尚需加强。这一问题的存在与我们今年阅读材料的发布方式也有一定关系。从第一课到第十一课的泛读文章都是以邮件的形式发给家长的，因此学生的完成度对家长的依赖性很高。在与六位学生进行的小组访谈中，学生们都表示，以后能有自己的账号直接在电脑或者 iPad 上看文章的话，他们都会自己主动去看。对此，我们拭目以待，也会进一步鼓励学生养成自主阅读的习惯。

表 5.3: 学生能否主动完成阅读任务

| | 人数 | 百分比 |
|------------------------------|----|-----|
| 每学完一篇课文，都会主动问家长有没有收到 AI 泛读文章 | 6 | 19% |

| | | |
|------------------------|----|------|
| 家长给我，我就会看，不给我，我也不会提醒他们 | 17 | 55% |
| 经常忘记看 | 3 | 10% |
| 未回答 | 5 | 16% |
| 总人数 | 31 | 100% |

在调查问卷的最后，学生们对 AI 泛读文章能在多大程度上帮助他们记忆和巩固刚学的生词进行了评分（表 5.4）。超过一半的学生给出了 7 分（有些帮助）和 10 分（非常有帮助），对这些泛读文章的效果予以肯定。值得注意的是，觉得 AI 泛读文章对学生没有帮助的六位同学中，有四位在 2024 年 6 月参加了 Youth Chinese Test 中小学汉语水平四级的考试。其中两位成绩在 290/300 分以上，两位在 270/300 分以上。对他们来说，试验中推送的 AI 泛读文章更能帮助他们拓展课本中没有出现过的字词，而不是记忆和巩固他们已经学会的东西。

表 5.4: 学生认为 AI 文章有多大帮助

| | 人数 | 百分比 |
|-------------|----|------|
| 10 分（非常有帮助） | 6 | 19% |
| 7 分（有些帮助） | 10 | 32% |
| 3 分（不太有帮助） | 7 | 23% |
| 0 分（完全没有帮助） | 6 | 19% |
| 未回答 | 2 | 6% |
| 总人数 | 31 | 100% |

在六人的小组访谈中，学生们还提到在阅读时，除了会碰到没记住的生字之外，文章后的阅读理解题也有一定难度。有学生觉得自己文章看懂了，但一做题又不记得刚才看了什么，需要重新在文章中寻找答案。这也从另一个侧面说明他们需要更多的阅读训练。

学年结束前，我们对四年级的学生家长也进行了问卷调查，高达 88% 的答卷希望升入五年级后，他们会继续收到这样的 AI 泛读文章。

学中文是一个细水长流、水滴石穿的积累过程，海外中文教育尤其如此。企图在不到一年的时间，看到立竿见影的效果是不现实的。但是经过近一年的实验摸索，我们看到了 AI 是如何帮助学生提升阅读水平的。2024 年 6 月，我校现在八年级没有使用过 AI 泛读的学生参加了 HSK 考试。希望四年之后，经过四年 AI 泛读文章学习的这批四年级学生升到八年级的时候，我们会看到他们的 HSK 成绩与今年的八年级相比有一个较大的提高。

六、总结

提高海外华裔学生的中文水平，需要在教学中引入和加强泛读来巩固精读的学习成果，然而普通的泛读材料，内容过于稀疏，与精读内容脱节，对海外学生的中文学习帮助有限。“泓河开卷TM”项目初步试验了使用 AI 工具帮助创作与精读课程中所学字词紧密配合的“浓缩”的泛读短文，使学生及时巩固，不断积累所学精读知识。希望这个方法能够帮助海外华裔学生降低 4-6 年级学中文的“坎”，进而突破中文阅读关，使中文“过关”（注：指可以应付中文作为工作语言环境）的学生可以

批量产出。在使“字不离词、词不离句、句不离段、段不离篇”教学理念通过 AI 课外阅读进一步强化的同时，再加上“篇反照字词”的回馈增强效应。

“篇反照字词”的总结是否贴切，方向是否正确，也欢迎讨论。有关该试验的效果，有待进一步观察验证。

七、“泓河开卷™”AI 阅读项目近期计划

在四年级试验基础上，我校计划从 2024 年 9 月开始逐步完成与《中文》5-8 册紧密配套的浓缩泛读短文，并逐步扩展个性化功能和分级需求功能。我们希望“泓河开卷™”AI 泛读文章辅助中文教学项目，切实收到开卷有益的功效，为在海外中文学校学习的华裔学生能“批量”跨过中文阅读关淌出一条新路，“卷”出一大批中文过关的学生。

鸣谢：

感谢加拿大联邦政府通过 Community Services Recovery Fund 对“泓河开卷™”AI 辅助中文教学项目的资助。



参考文献：

[1] 暨南大学华文学院、华文教育研究所编《中文》第四册（修定版），暨南大学出版社出版，ISBN 7-81029-638-8，2006 年 12 月