即享朗读评分——国际中文朗读语料评分模型实践平台项目报告

作者 叶晓良

目录

[一、项目概述 3](#_Toc185961416)

[二、项目整体架构 5](#_Toc185961417)

[三、项目模块详述 6](#_Toc185961418)

[1. 前端模块 6](#_Toc185961419)

[2. 后端模块 7](#_Toc185961420)

[3. AI模块 8](#_Toc185961420)

[4. 数据库模块 9](#_Toc185961420)

[四、项目主要页面贴图 9](#_Toc185961418)

[总结 12](#_Toc185961421)

# 项目概述

**（一）本项目旨在构建一个科学、易用、高效的中文朗读评分与智能学习平台，专门解决国际中文学习者在朗读评分与发音改进方面的需求**。通过改进评分标准、接入AI大模型以及扩展功能，探索更科学的中文朗读评分体系，并为师生提供实用的学习与教学工具。

项目的核心目标之一是简化评分标准，改为以音节为单位的综合评分，减少细分内容，以更贴近人类评估习惯。此外，评分体系引入了合理的分布调整分数计算方式，确保能够准确反映学生的真实水平并满足不同学习背景的需求。为此，项目已构建了高质量的语料数据库，收集并标注了1000余条真实学习语料，涵盖发音错误类型，并为AI大模型训练提供精准数据支持。

**平台功能涵盖学生端和教师端**，学生端支持朗读练习、评分、发音分析和学习进度跟踪等功能，教师端则提供作业布置、批改、录音管理以及发音问题分析等服务，并逐步实现自动化教学建议。**结合智谱AI平台的GLM-4-9B模型和LoRA微调技术**，项目成功提升了评分准确性和F1分数。平台将通过学院语言课堂内测，不断改进评分模型与功能，并逐步扩大用户规模，覆盖多学段、多层次的中文学习者。创新方面，项目采用了先进的AI模型和精细化语音数据标注技术，同时简化了评分方式，优化了评分宽严度的调节机制，确保评分体系既科学又具有实际适用性。通过学生激励与教师辅助功能，项目有效激发了学生的学习兴趣，提高了教师的教学效率。未来，平台将持续优化评分模型与用户体验，推动中文学习技术的普及与发展。**目前只完成了学生端的开发。**

**（二）功能介绍**

**1.朗读评分练习**

朗读评分练习是平台的核心功能之一，旨在帮助留学生通过朗读和录音来提升中文口语能力。留学生可以从平台提供的中文朗读材料库中选择文本进行朗读，并通过麦克风录制自己的朗读音频。上传录音后，系统会调用训练好的AI模型进行评分，并提供详细的评分报告和发音建议。交互设计方面，文本选择采用卡片式布局，提供多种难度和主题的朗读材料，每个素材卡片都有树状结构，方便学生根据课本、单元等选择合适的练习内容，确保能够根据个人的水平和兴趣进行选择。录音界面设计简洁，主要包含“开始录音”、“停止录音”、“上传”和“取消”按钮，录音过程中会显示录音时长和剩余时间，上传时则显示进度条和状态提示，确保学生清楚了解上传进度。上传完成后，系统会自动生成评分报告，详细列出发音错误、语调问题等，并提供改进建议，评分报告采用图文结合的形式，便于学生理解和改进。通过这些设计和技术支持，朗读评分练习功能能够为留学生提供高效、便捷的中文朗读学习环境，帮助他们提升口语能力，并通过个性化的评分报告和改进建议，确保针对性学习和持续进步。

**2. 自定义朗读评分**

自定义朗读评分是平台的重要功能模块之一，旨在为学生提供更大的自由度，让他们根据个人学习需求自行输入文本进行朗读，并通过上传录音获得评分。该功能特别适用于留学生在练习特定词汇或句子时，帮助他们更好地掌握中文发音。交互设计方面，文本输入界面提供了一个简洁的输入框，学生可以在其中输入自定义文本，方便他们根据需要进行练习。录音界面设计简洁明了，主要包括“开始录音”、“停止录音”、“上传”和“取消”按钮，录音过程中会显示录音时长和剩余时间，提供清晰的时间提示；上传时则会显示进度条和状态提示，确保学生清楚了解上传进度。上传完成后，系统会自动生成评分报告，详细列出发音错误、语调问题等，并提供改进建议，评分报告通过图文结合的方式呈现，便于学生理解。通过这些设计和技术支持，自定义朗读评分功能能够为留学生提供高效、便捷的中文朗读学习环境，让他们通过自主输入文本进行练习，提升口语能力。详细的评分报告和个性化的改进建议，可以帮助学生针对性地提高发音和语调，实现更有效的学习进步。

**3. 个人主页**

个人主页是平台的重要功能模块之一，旨在为留学生提供一个集中管理自己学习记录和个性化设置的中心。该模块主要包括查看口语评分记录、获取专业模型针对用户学习记录的“发音特点”和“学习建议”、以及修改个人信息等功能。通过这些功能，留学生可以更好地了解自己的学习进度和发音特点，并根据个性化建议进行针对性的改进。交互设计方面，评分记录部分采用时间轴式布局，清晰展示每次朗读练习的次数、文本内容、评分和发音建议。发音特点与学习建议部分使用Markdown语法呈现，确保信息结构清晰且易于阅读。个人信息修改界面简洁明了，包含头像上传、姓名、手机号、密码、国籍等选项，方便用户更新个人资料。通过这些设计和技术支持，个人主页功能模块能够为留学生提供高效、便捷的学习管理环境，帮助他们更好地掌握自己的学习进度、发音特点和个性化建议，实现更有效的学习进步。同时，个人信息管理和数据安全措施的优化，确保了留学生的隐私和数据安全，进一步提升了用户体验。

# 项目整体架构

**在本项目的开发过程中，我们采用了前端、后端及数据库等多层次的技术栈并结合了AI大模型，以确保平台的稳定性、性能和用户体验。**

本项目的技术架构分为前端层、后端层、AI模型层、数据库层和运维与部署层，构建了一个高效、可扩展的系统。以下是项目中使用的主要技术的大致介绍：

**（一）前端层通过网页和移动端提供用户交互，主要技术包括HTML5、CSS3、JavaScript（ES6+）**结合Axios处理API请求，使用Webpack进行资源打包和优化，确保页面美观和动态交互功能。

（二）后端层采用Spring Boot框架，结合Spring MVC实现RESTful API接口，使用MyBatis进行数据库交互，Spring Security保障系统安全，Spring Boot Actuator用于系统监控，，Git用于代码和数据版本管理，JWT确保API安全传输，Logback用于日志管理，JUnit用于单元测试。

**（三）AI模型层使用GLM-4-9B语言模型，并通过LoRA微调技术进行优化**，结合讯飞语音分析提升发音评分准确性，支持通过Java API调用，计划结合Docker实现容器化部署。

（四）数据库层采用MySQL存储用户信息、任务数据和评分结果，使用Redis进行缓存，提高系统响应速度。

（五）运维与部署层使用阿里云ESC服务器支持云端部署，确保系统的高可用性和扩展性。

# 三、项目模块详述

## 1. 前端

本项目的前端模块主要作为用户交互层，提供学生和教师使用的图形化界面，包含录音上传、评分查看、学习记录管理、教师作业布置等功能。

**学生端功能**包括朗读评分练习，学生可以选择平台提供的文本素材进行朗读，系统自动评分并生成分析报告；自定义朗读功能允许学生输入自定义文本并录音上传，平台则根据发音特点生成个性化评分和改进建议；进度查看功能展示学生的学习记录，包括评分历史、发音特点和改进建议；激励功能通过积分奖励和动态排行榜等方式，激励学生积极参与学习。

**教师端**则支持布置朗读作业、批量检查学生提交的录音，并通过系统自动评分辅助教师完成评分任务，还提供录音文件的分类管理功能，便于教师高效查找和分析学生的学习记录。

**前端技术栈主要包括HTML5、CSS3和JavaScript（ES6+），使用React.js或Vue.js构建模块化的用户界面**，通过Axios与后端API交互，处理数据请求和响应；WebSocket技术用于实时更新用户排行榜和进度图等动态信息。设计方面，前端采用响应式设计，支持移动端优先的页面布局，方便学生在零碎时间进行练习；**同时，UI设计注重友好性**，增加了吉祥物“小吉祥”与互动元素，提升学习趣味性；页面模块化设计确保学生端和教师端功能界面分离、各自功能明确，操作便捷。**目前只完成了学生端的开发。**

## 2. 后端

本项目的后端模块作为系统的核心逻辑层，负责处理用户请求、任务分发、数据管理及与AI模块的交互，是前端与AI模型之间的桥梁。

**后端功能包括API服务模块、用户管理模块、任务分发模块和数据管理模块**。**API服务模块**通过实现RESTful API，为前端提供标准化接口，支持用户注册与登录、音频上传与管理、评分结果获取等功能。**用户管理模块**通过session实现身份认证和权限管理，支持学生与教师两类用户，并采用JWT进行安全的身份验证。数据管理模块负责管理用户学习记录、评分结果及上传的音频文件，数据库结构设计合理，确保数据的快速检索和统计分析。

**后端技术栈主要包括Spring Boot**，支持MySQL关系型数据库，使用Redis作为缓存层，优化评分结果的读取速度，后端使用多线程处理任务，提高对高并发请求的处理能力。

## 3. AI模块

AI模块是系统的智能核心，负责处理用户上传的音频数据，进行发音分析和评分，并生成改进建议。该模块包括**数据预处理、模型训练和模型部署**三个功能模块。

1. 数据预处理模块首先对用户上传的音频进行格式转换和归一化处理，调整采样率至16000Hz，并使用讯飞语音评测模型提取音素信息（包括声母、韵母、声调等）。同时，结合Praat工具生成文本网格（textgrid）文件，用于标注发音错误类型。
2. 模型训练模块使用智谱AI平台的GLM-4-9B模型，并结合LoRA微调技术，使其能够适应中小规模数据集任务。训练数据来源于自建的韩国留学生语料库，包括5000条音频，时长超过3000分钟，数据标注涵盖声母、韵母、声调、语速、停顿等多个维度，确保评分的准确性。评分逻辑通过综合分析语速、停顿、声母、韵母、声调等特征进行细粒度评分，输出具体分数、错误分析和改进建议，验证集上的评分准确率达85%，F1得分为82%。
3. 模型部署模块则通过Java后端API集成模型推理功能，实现在线实时评分，支持批量处理和实时推理两种模式。

**技术实现方面**，AI模块采用Python进行模型加载和训练。数据标注与分析则结合讯飞非开源语音评分模型提取错误类型，并通过人工标注补充一些讯飞未覆盖的细节。

## 4. 数据库模块

数据库模块主要负责**存储用户信息、音频数据、评分记录及日志**。其设计特点包括用户数据表、任务记录表和日志表。用户数据表存储用户的基本信息，如用户名和角色类型。任务记录表用于记录用户上传的音频任务及其评分结果，便于后续的追踪和分析。日志表则存储系统操作日志，包括用户登录、任务请求、模型调用等记录，以便于系统维护和监控。

**技术实现方面，数据库采用关系型数据库（MySQL）**，支持复杂查询和事务处理，确保数据一致性和完整性。同时，使用Redis作为非关系型数据库，缓存热点数据（如session），提高系统的访问速度和性能。

# 项目主要界面贴图

为便利学生获取朗读评分学练资源，我们设计了一个界面简洁、易于使用的网页平台，即“即享朗读评分”。该平台旨在为用户提供中文朗读语料及配套的朗读评分工具，以支持学生的练习需求。考虑到学生在零碎时间进行口语练习的需求，我们最终选择移动端（手机）界面作为主要展示界面（图1），实现即时录音和上传功能。此外，我们设计的吉祥物“小吉祥”会在学生上传朗读录音后为用户评分，并提供相应的评语和建议，寓教于乐，帮助学生更有针对性地提升朗读水平。

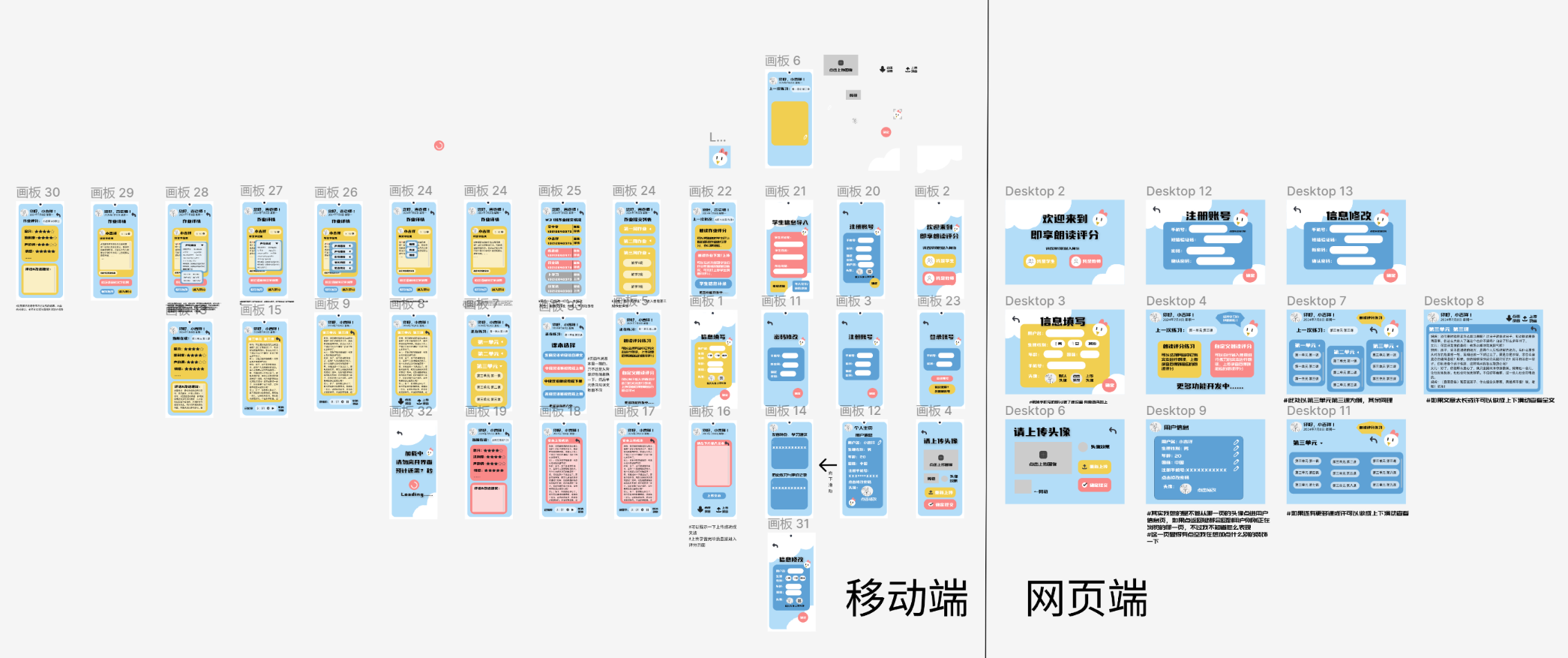


图1“即享朗读评分”移动端与网页端UI设计（部分）

目前，平台主要分为学生端（图2）和教师端（图3）两大板块，分别面向不同的用户群体。学生在填写基本个人信息后，可以进入学生端学习界面。该界面主要包括两大功能：一是朗读评分练习，即选择网站中已有的文本进行朗读并上传录音以获得相应的评分；二是自定义朗读评分，即自行输入文本进行朗读并上传录音以获得评分。在学生个人页面中，用户可以修改个人信息、进行个性化设置，并查看历史练习与评价记录，还可看到系统分析出的发音特点和个性化学习建议。教师端的设计则更为复杂，可用于布置朗读作业、辅助评分以及管理学生上传的录音等。鉴于开发时间有限，**我们目前主要集中在学生功能界面的开发上**，未来有望根据已设计的教师端UI界面进一步完善其功能。



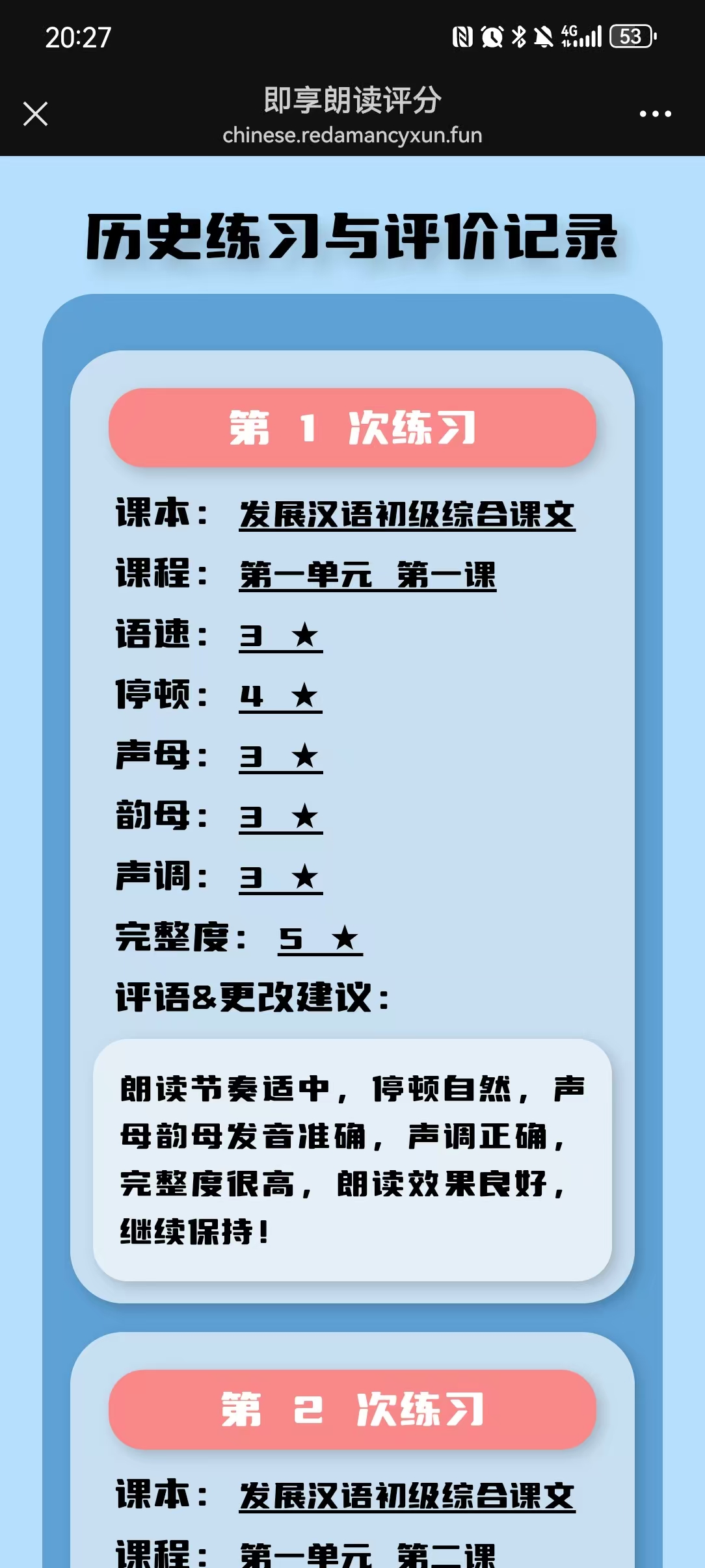
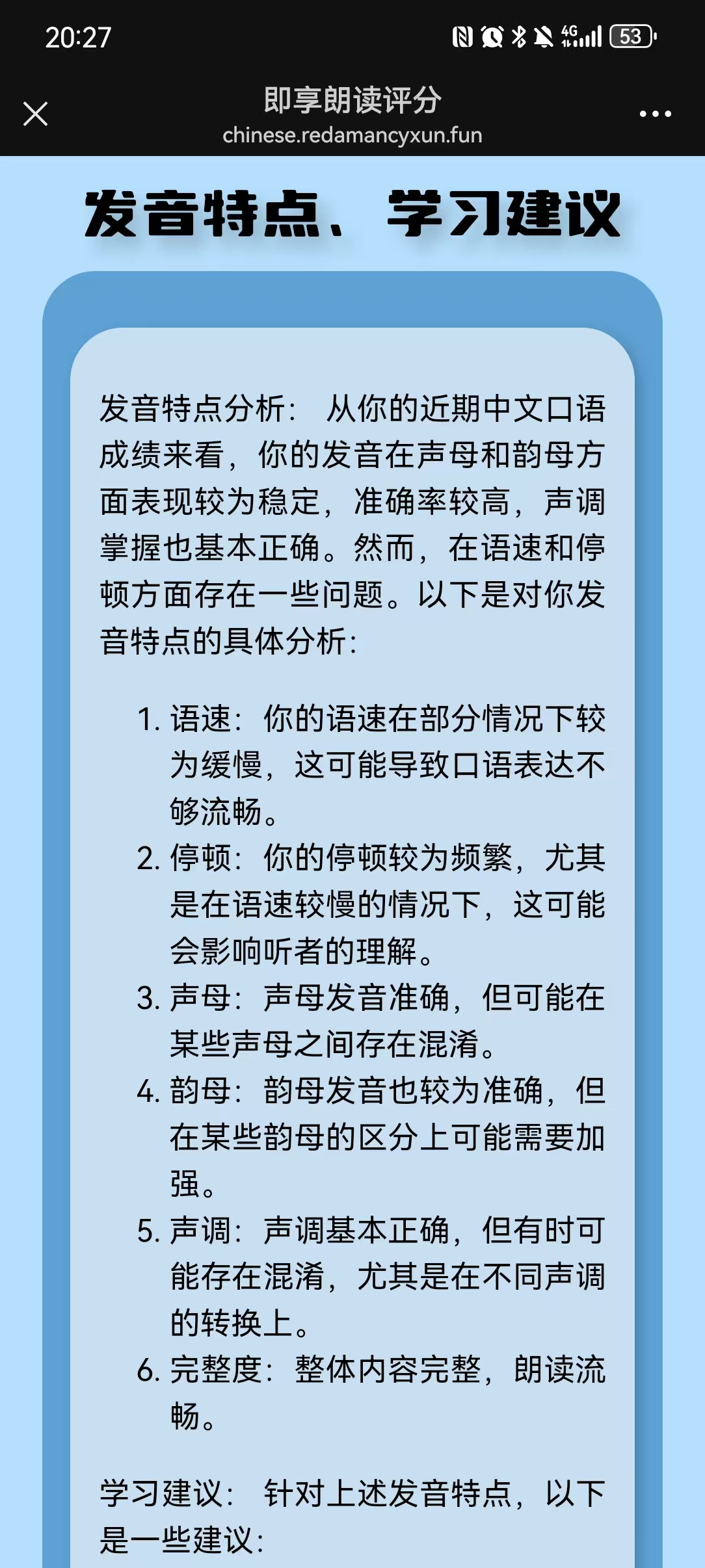
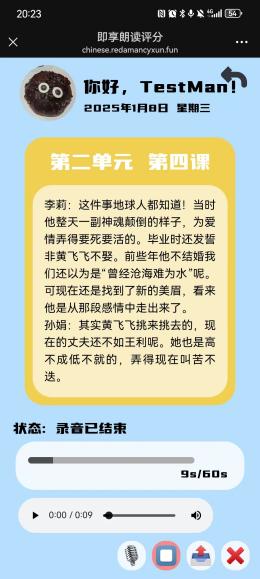
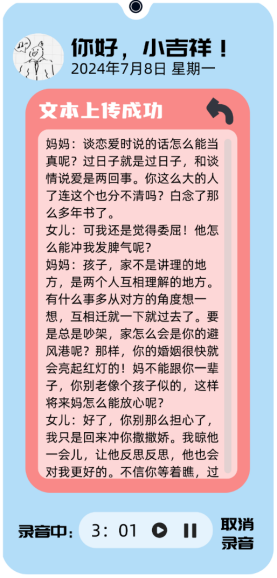
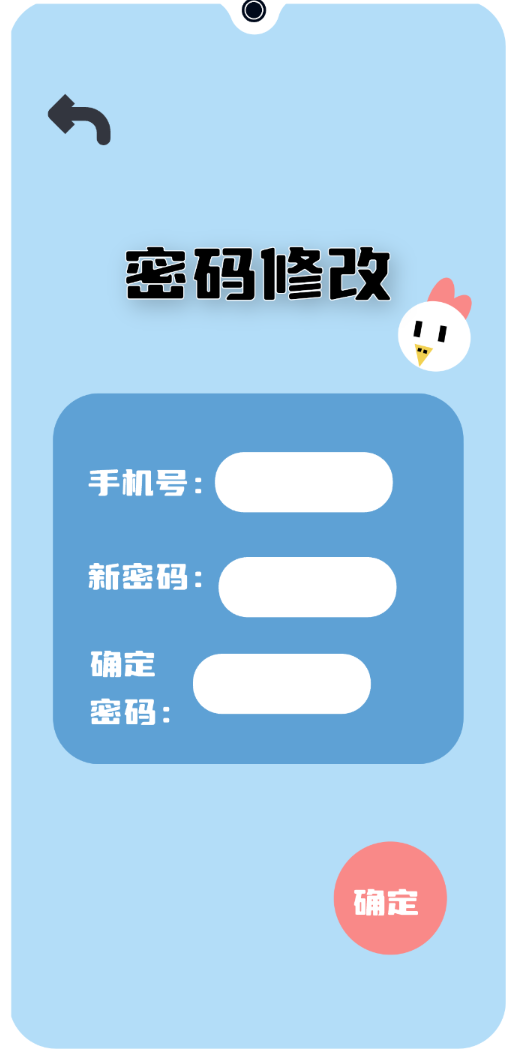
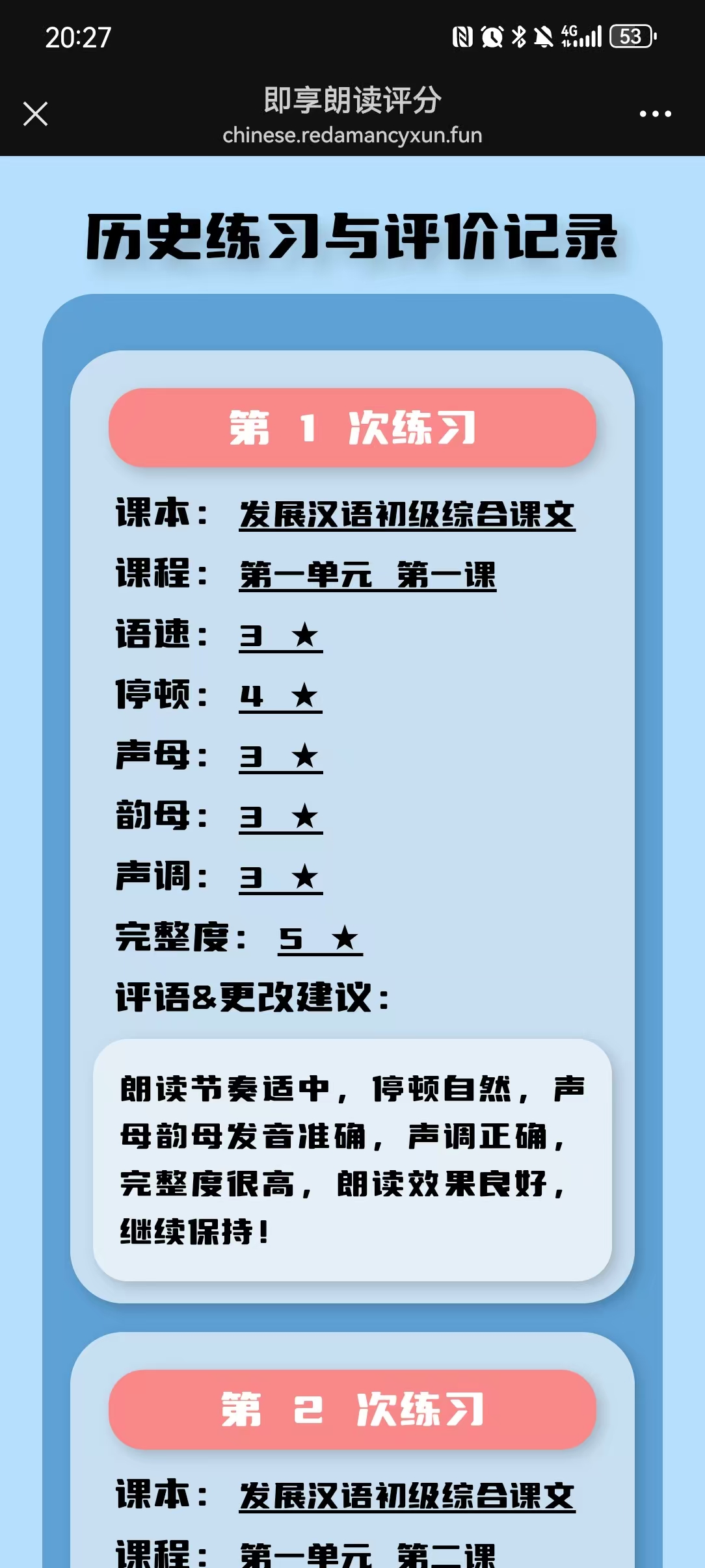
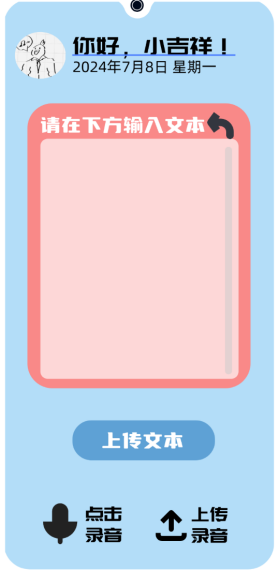
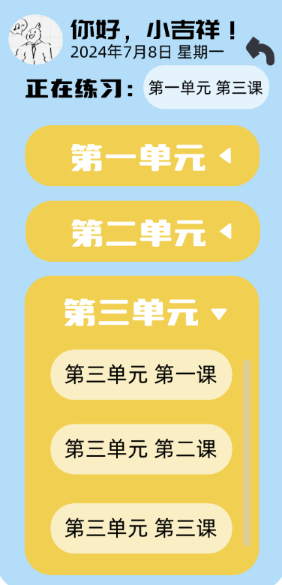
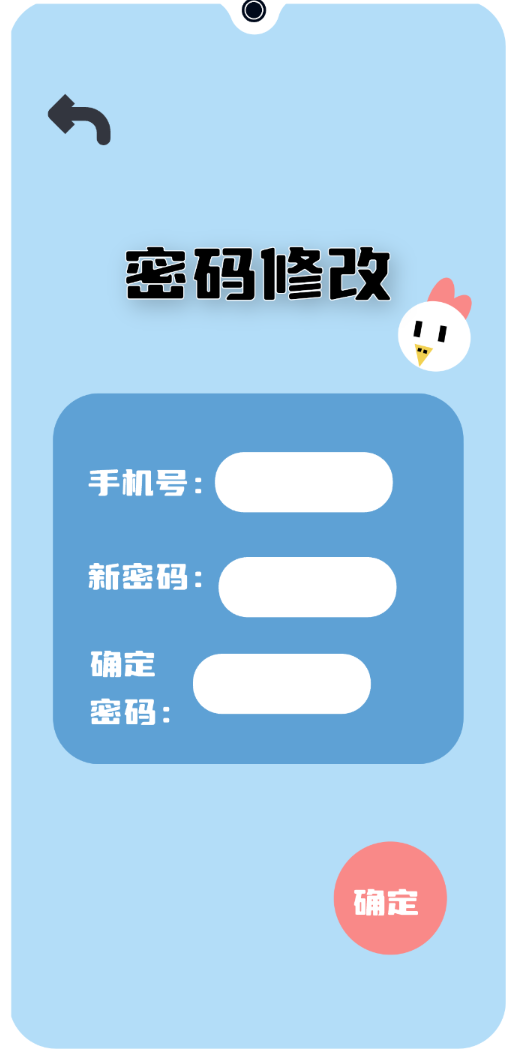
   

图2 “即享朗读评分”移动端学生界面截图（部分）



图3 “即享朗读评分”移动端教师界面截图（部分）



图4 “即享朗读评分”登录界面截图（部分）

# 五、总结

**（一）项目创新点**

该项目在AI大模型训练、语料处理、数据库建设、标注精细化、评分体系多维创新、以及用户界面设计等六大环节体现了显著的创新点，具有较强的实践性和推广价值。

**首先，本项目采用了先进的AI技术。**我们应用智谱AI平台的GLM-4-9B模型，并采用LoRA（Low-Rank Adaptation）微调技术，对多项模型训练参数进行调整和优化，辅助以Java编写后端API和讯飞语音评测平台功能，初步实现了朗读语料评分和评语建议的智能生成。这使得机器在朗读语料评分合理度上以及语音评分效率上有所提升，同时节省了学生的学习成本和教师的人力劳动。

**其次，在文本与音频数据的处理环节，本项目实现了精准的数据切分和标注，并建立了高质量的语音评分数据库**。通过学院自建的韩国留学生语料库结合北大版《高级汉语听说教程》，项目团队实现了数据的多重细致分段与音频归一化处理，确保了数据的高效使用与精度提升。在音频分析中，利用讯飞非开源评分模型与Python接口进行自动化分析，并结合Praat进一步细化标注的创新性工作，弥补了声母和韵母具体偏误的标注空白。这一结合不仅提高了分析的精准度，还为未来大模型的细致评分提供了坚实的数据基础。此外，项目团队通过人工复检与格式化标注，如“/p/->/b/”和“[ɕ]->[ʃ]”，实现了系统的语内与语际偏误区分，进一步推动了语音研究的细化和系统化。项目构建的数据库中包含不少于1000条可供模型学习的高质量音频，可以很好地应用于AI大模型的训练。

此外，在评分标准的制定方面，**我们的评分方式更加科学客观**。项目在机器评分的基础上，采用Z-Score和T值法实现了数据的量化和归一化处理，生成百分制评分，确保了评分结果的科学性和公平性。通过结合人工评分与机器评分的反馈，项目逐步完善了评分体系，创新地涵盖了准确度、流利度和完整度等多维度的分析，使评估更加全面和具有指导性。人工补充的声调标注尤其创新，将声调错误区分为“声调混读”和“怪调”类别，首次细致呈现了学生声调偏误的多样性。

**最后，在用户体验方面，本项目设计的平台简单易用、功能强大。**该项目设计了“即享朗读评分”平台，采用移动端优先的策略，充分考虑了用户碎片化时间的使用场景，使学生能够随时随地进行口语练习。此外，项目设计的吉祥物“小吉祥”所带来的人机互动设计提升了用户的学习参与度和使用体验。项目还综合了用户反馈并持续迭代优化，彰显了以用户为中心的设计理念。

综上，我们的项目在多个方面具有创新性，对既往的口语评分标准和市面上现有的朗读语料评分模型进行了完善与突破。

**（二）文本分段**

在文本分段环节，我们深刻体会到合理的语段划分对于模型评分精确性的关键作用。通过实验验证，我们意识到文本的结构直接影响到朗读时的流畅度和自然性，从而影响评分的准确性。得益于此，我们在文本分段的过程中逐渐形成了一套系统化的方法，确保语段的逻辑性和易读性。该阶段的成功经验在于团队成员充分发挥各自的语言学和计算机技术专长，通过多次研讨与验证，最终确定了合理的分段策略，这为后续工作奠定了良好的基础。

**（三）音频剪切与分析**

音频剪切与分析的实施则是项目进展中最具挑战性的部分之一。在多次尝试与实验后，我们发现音频剪切的时长和方式直接影响模型的学习效果。过短的音频无法提供足够的语音特征，导致评分不稳定；而过长的音频又会增加计算负担，降低模型响应速度。经过细致的研究，我们为音频剪切制定了严格的规范，确保模型能够有效进行标准语音与学生朗读的对比。在此过程中，我们使用讯飞、praat等语音分析工具，对音频的声学信号进行深度处理，提取发音的声母、韵母、声调、语速、停顿、完整度等多维度特征。该环节的技术挑战在于如何保证音频分析结果的稳定性和准确性，同时实现对语言学特征的精准捕捉。通过团队的共同努力和技术攻关，我们逐步克服了这些困难，提升了模型的评分效率和可靠性。

**（四）评分标准的制定**

评分标准的制定是项目实施的重要节点之一。初期，我们尝试构建了综合评价体系，将机器评分与人工评分相结合，以确保评分的全面性和细致性。讯飞的评分结果提供了多维度的初始反馈，我们基于这些反馈信息进一步细化了评分算法，开发了涵盖不同发音特征的计算模块。然而，我们在此过程中也遇到了一些瓶颈，例如如何在人工评分中有效补充机器评分的不足，特别是在情感表达和语调细微变化的评估上。为了应对这些问题，我们在后期引入了逐项评估机制，使声母、韵母和声调的评分更加精确，并在情感维度上进行了初步尝试。虽然这一探索尚未完全落地，但它为今后的研究提供了有价值的思路。

**（五）UI设计和网页开发**

UI设计和网页开发是整个项目中将技术成果转化为用户友好体验的关键步骤。为了满足用户需求，我们设计了简洁、直观的界面，使学生能够轻松进行朗读测试并及时获得反馈。这一过程中，我们充分考虑了使用者的便捷性和功能的直观性，通过多轮的设计和用户测试，优化了界面布局和交互细节。此外，网页平台的开发为教师提供了便捷的教学辅助工具，使评分流程更加透明和高效，提升了教学效果和师生互动。