赋能自动障碍语音识别：阿坎语障碍语音数据集

Isaac Wiafe、*Akon Obu Ekpezu*、*Sumaya Ahmed Salihs*、*Elikem Doe Atsakpo*、*Fiifi Baffoe Payin Winful*、*Jamal-Deen Abdulai*

加纳大学计算机科学系

摘要

障碍语音数据的缺乏阻碍了包容性语音技术的发展，尤其是在阿坎语等低资源语言领域。为弥补这一空缺，本研究整理并发布了一个语料库，其中包含母语为阿坎语且有言语障碍的说话人的语音样本。该数据集包含 50.01 小时的音频录音，涵盖四类障碍语音：口吃、脑瘫、腭裂和中风诱发的言语障碍。录音在受控且有监督的环境中完成，参与者用自己的话描述预先选定的图像。最终形成的数据集包含音频录音、转录文本，以及有关说话人人口统计信息、障碍类别、录音环境和设备的相关元数据。该数据集旨在支持低资源语言自动障碍语音识别系统和辅助语音技术的研究。

关键词：障碍语音数据；自动语音识别、自动障碍语音识别、低资源语言；语言包容性；机器学习

1　背景

言语障碍会影响沟通、教育和社会参与。然而，现有的大多数语音数据集和辅助语音技术（如 Alexa 和 Siri）都聚焦于临床标准化语音。这些系统并非为准确识别言语障碍者的言语模式而设计。高资源语言与低资源语言之间的技术鸿沟进一步加剧了这一挑战。阿坎语是加纳一种重要的本土语言，但在障碍语音语料库中的代表性仍尤其不足。截至本研究开展时，尚无关于阿坎语自动障碍语音识别（ADSR）的公开数据或研究。此前对阿坎语进行数字化的尝试主要聚焦于标准言语 [1–3,5]。

为填补这一空白，本研究提出了一个阿坎语言语障碍数据集。该数据集包含约50小时的阿坎语音频录音，涵盖四种障碍病因，即腭裂、脑瘫、口吃和中风所致言语障碍。因此，本研究为包容性语音建模以及开发能够在低资源语言情境中识别言语障碍的技术奠定了基础。该数据集部分内容在ADSR中的潜力已在Salihs等人[4]中得到证明。本文其余部分将该数据集称为UGAkan-ImpairedSpeechData[6]。

通信邮箱：aoekpezu@st.ug.edu.gh

2 方法与材料

2.1 材料

本研究使用了两种主要工具进行数据收集和转录：UGSpeechData移动应用和定制的桌面转录界面。每种工具将在随后的小节中详细说明。

2.1.1 UGSpeechData 移动应用

本研究采用了 UGSpeechData 移动应用程序。该应用最初设计用于采集五种加纳语的标准语音数据，此前也曾用于此目的 [5]。这是一款基于安卓的工具，可用于录制音频、跟踪录音会话、进行验证和转录，并将转录文本导出到云服务器进行安全存储。为确保 UGSpeechData 移动应用能够有效支持言语障碍数据的采集，研究人员进行了多项改进，以满足具有不同障碍病因的个体在独特言语模式和录音需求方面的要求。这些调整对于保留言语障碍的真实性、完善录音过程中使用的描述性提示语，以及准确分类参与者至关重要。表1总结了用于采集标准言语的应用原始版本与本研究采用的修改版本之间的主要差异。

表1：为收集言语障碍数据而对 UGSpeechData 应用进行的修改

组件	标准言语版本	言语障碍版本	依据
暂停／中断处理	受限 频繁停顿；包含超过 3 秒静音或过多的填充音未保存。	已移除 all 停顿限制，允许自然的录音停顿，填充词，言语不流畅，以及含混的语段。	言语障碍通常包括频繁停顿、言语不流畅，以及作为自然语音一部分的含混发音模式。限制这些行为将导致失去 of 真实的 语音言语模式。
提示图像	包含 1,000 张图像用于引出自然的描述。	新增 200 张图像（总计 = 1,200），包括更简单且更容易识别的图像。	扩大图像集，尤其是更简单且更容易识别的图像 增加 词汇覆盖率 and 引出 更容易描述 or 讲故事，尤其是对于言语障碍人士，他们可能还存在一定程度的认知障碍。
最长录音时长	每段录音允许时长为 15–30 秒每个音频文件的限制。	将时长延长至最多 to 60 秒 per 录制。	言语障碍通常需要更多时间，因为发音较慢，并伴有重复、犹豫和停顿。较长的时长确保了完整的话语均被采集且未受干扰。
User 个人资料字段	已收集 基本人口统计信息仅。	添加了 病因用于分类的信息每位用户的语音类型障碍。	病因标注实现了准确的分类并强化了数据集结构。
验证/Transcription	验证 and 转录 were 进行 在……之内 the 移动应用中使用内置功能 on 安卓设备。	这些 功能 were 已禁用。	为提高效率，请支持更长的音频录音，提供更大的界面，并实现准确的渲染 of 阿坎语特有字符，使用专用的阿坎语键盘。

通讯作者：aoekpezu@st.ug.edu.gh

2.1.2 转录应用程序

如表1所示，言语障碍数据集所采用的转录工作流程与标准语音采集集所采用的工作流程存在显著差异。UGSpeechData 移动应用原本支持在安卓设备上进行应用内验证和转录，但在言语障碍项目中，这些功能已被禁用。将转录流程与移动应用分开的决定，是根据负责转录的语言专家的需求作出的。他们需要一款基于桌面端的工具，具备更大的界面、更便捷的导航功能，并支持较长且更复杂的语音录音所需的长时间转录会话。

为满足这些要求，我们开发了一款适用于 Windows 的定制桌面转录应用程序（https://github.com/HCI-LAB-UGSPEECHDATA/Transcription-App）。该转录工具（见图2）使用 Python 构建，采用 Tkinter 实现图形界面，使用 Pygame 播放音频，并通过 Pillow 渲染图像。它会加载每个音频文件及其对应的图像提示，并以直观的界面呈现，使转录员能够实时收听音频并输入转录文本。

为确保准确呈现语言特点，我们还开发了一款定制的阿坎语键盘，并在 [5] 中使用。该键盘（https://github.com/HCI-LAB-~~UGSPEECHDATA/SPEECH-DATA-KEYBOARDS/blob/main/akan.kmp~~）支持准确输入“ɔ”等阿坎语特有字符；在标准系统键盘中，这些字符往往会被误用右括号“)”等符号代替。

该转录应用程序提供标准的播放控制功能（播放、暂停、继续、重播），并包含可视化进度条、转录状态指示器以及“下一项”/“上一项”导航按钮，方便用户在数据集中高效切换。转录文本输入集成的文本字段后，会自动保存到可编辑的 Excel 电子表格中，并持续更新以防止数据丢失。该工具使用 PyInstaller 打包，便于在多台计算机上分发和安装。

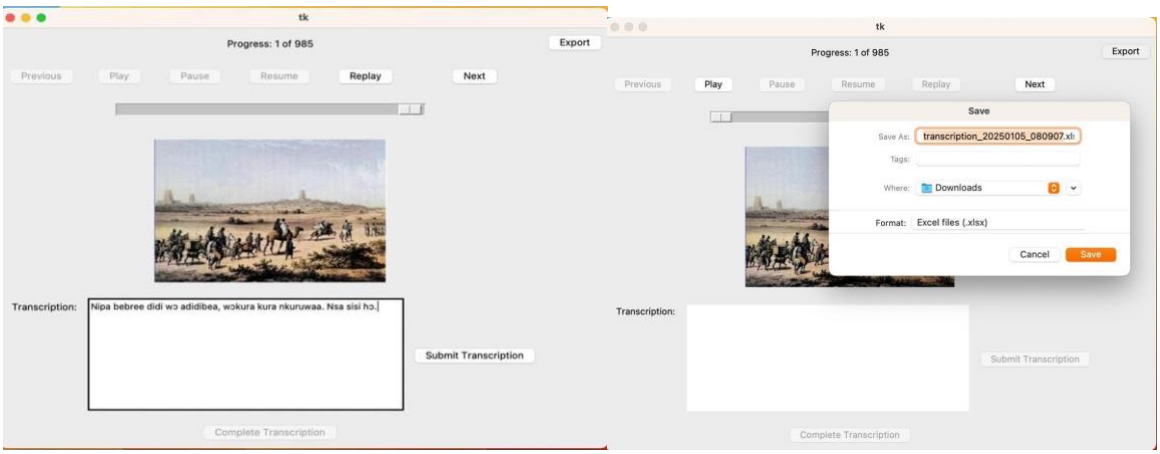


图 1. 转录应用程序界面（左：应用程序运行中；右：导出）

2.2 方法

本研究采用多阶段方法，包括参与者筛选与招募、使用经过修改的 UGSpeechData 移动应用程序进行音频录制、验证收集到的录音，以及通过定制的桌面界面进行转录。这些流程旨在确保可靠地收集言语障碍

通信作者：aoekpezu@st.ug.edu.gh

数据，同时遵守伦理规范，确保语言准确性和方法严谨性。下文各小节将详细介绍每个阶段。

2.2.1 参与者筛选、选择与知情同意

参与者通过便利抽样招募。常驻阿克拉和库马西的言语治疗师协助寻找脑瘫和口吃患者，主要来自库马西的机构。为确保录音反映阿散蒂语使用者的言语模式，从库马西招募参与者至关重要。这是因为该项目主要关注阿坎语的阿散蒂方言。此后，为更全面地呈现阿坎语内部的语言变异，研究团队又从阿克拉招募了其他参与者。这些参与者包括中风康复者，以及脑瘫、腭裂和口吃患者。

研究团队成员创建了用户账户，并为每个账户分配了专用应用电子邮件，格式为 akanXX@ugspeechdata.com，其中 XX 表示分配给每位参与者的唯一编号，用于身份识别和数据管理。参与者使用提供的电子邮件和密码登录移动应用并录制音频。研究团队鼓励用户在登录后立即确认并更新个人资料（姓名、年龄、性别和病因类别），并在必要时提供协助。在个人资料页面中，参与者必须阅读并接受知情同意书。

通过应用直接获取了知情同意。知情同意书清楚说明了项目目的、收集数据的预期用途、参与者随时退出的权利，以及补偿的具体安排。知情同意页面的截图请参见图3。此外，所有参与者都必须确认自己年满18岁。

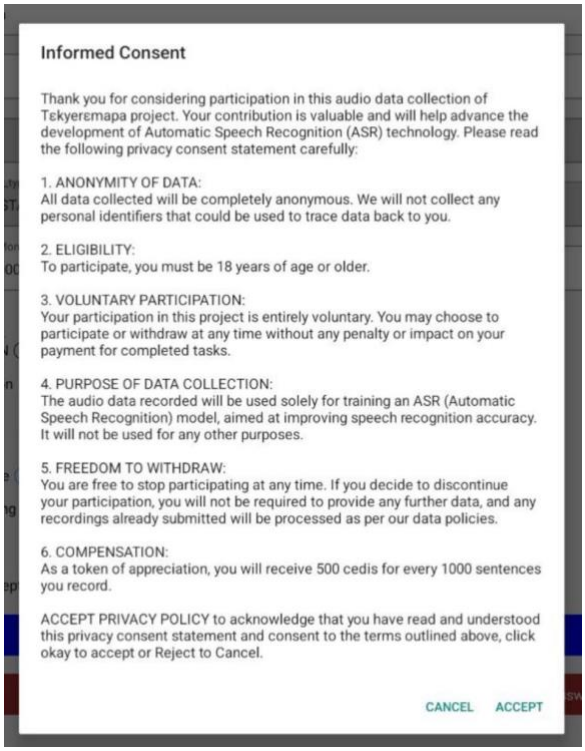


图2. 知情同意截图

通信：aoekpezu@st.ug.edu.gh

2.2.2 音频录制

为了 eliciting 自然、自发的语音，研究人员请参与者用自己的话描述一组提供的图像。采用图像提示法，是因为研究表明，该方法能够获得自然的语音录音，并能围绕同一视觉刺激图像产生多样的表达 [5]。移动应用还允许用户追踪自己已录制并上传的音频片段数量。

由于认知和行动方面的困难，尤其是脑瘫参与者面临的挑战，一些参与者在录音过程中需要协助。言语与语言治疗师和研究团队成员为这些参与者提供支持，帮助他们拿稳移动设备，并协助引导语音表达过程。有时，这包括讲述简短的故事，或提供有关图片的上下文线索，帮助参与者在录音前生成描述。

研究团队向参与者提供了以下录音指南，随后验证人员和转录员也依据这些指南标记不符合要求的音频文件：

i. 录音应在安静、没有背景噪音的环境中进行。 ii. 所有录音都必须使用阿坎语；仅当阿坎语中没有对应词语时，才允许使用英语。 iii. 每段录音的时长应为 5 至 60 秒。 iv. 鼓励参与者自由、自然地说话，尤其是口吃者，以避免他们自我审查。 v. 录音不应以 “ 我看到 ”“ 在这张图片中 ” 或 “ 看起来像是 ” 等填充语开头。 vi. 录音不得包含亵渎或冒犯性语言。

只有在环境噪声足够低时，应用程序中的录音功能才会启用；否则，应用程序会显示“噪声过大”通知，并阻止录音。参与者可以跳过任何他们无法描述的提示。每次录音结束后，应用程序允许用户播放、保存或删除录音，用户也可以随时退出应用程序休息。

2.2.3 音频验证与转录

音频录音收集完成后，由八位语言学专家进行验证和转录。他们均为阿散蒂契维语母语者，并接受过大学层次的阿坎语培训。在正式开展工作之前，这些专家接受了培训，学习如何使用桌面转录应用程序、如何导出转录文本，以及如何应用验证和转录指南。

验证和转录按照结构化的两阶段工作流程同步开展，旨在提高准确性、可靠性和质量控制水平。八位专家被秘密分成两组，每组分配到相同的一批音频文件，但他们并不知道文件重复分配的情况。这一设计引入了评估者间信度机制，使研究团队能够发现差异、核实独立转录员之间的一致性，并尽量减少对言语障碍进行解读时的主观偏见。

音频录音被分成多个批次，每批约 500 个文件，并尽可能按说话人分组。这种做法证明颇有优势，因为不同个体的言语障碍特征差异很大；因此，对某位说话人的

通信邮箱：aoekpezu@st.ug.edu.gh

模式越熟悉，专家就越能理解、解读并准确转录后续录音。

每位语言学专家以每批500个的方式依次接收并处理音频文件，在完成一批的验证和转录后，才会被分配下一批。这一迭代过程持续进行，直至全部14,312条音频录音均由整个团队处理完毕。平均而言，每位专家验证并转录了约1,800个音频文件，不过具体数量会因各位专家完成先前批次的速度而略有差异。

在被称为转录验证的第二阶段，研究团队对每个音频文件生成的两份独立转录文本进行了比较。若两位语言学家产出的结果完全相同或近乎相同，则接受该转录文本。对于存在冲突的转录文本的文件，音频会被重新分配给此前未处理过该文件的一对新专家，以进行独立审核和冲突解决。这种多层次方法增强了数据集可靠性，最大限度地减少了转录错误，并确保了整个语料库的一致性。

在音频验证方面，语言学家须识别并标记不可接受的内容，包括冒犯性或亵渎的语言、仇恨或辱骂性表达，以及个人身份标识符（PI），如姓名或电话号码。这些内容通过在转录栏中键入关键词“reject”进行标记。

鉴于言语障碍的复杂性，以及确保多名转录员之间一致性的需要，制定清晰、详细的转录规则至关重要。这些与语言学专家共同设计的指南规定了对重复、言语不流畅、言语错误、改造英语词、歧义话语及其他转录挑战的处理方式。

双方商定了一般指南，用于处理非语言声音，例如将“ɛɛɛɛɛɛɛ、ɛɛ、ɛɛɛ”记作“ɛɛɛ”，将“aaaaaaa”记作“aaa”，并将“mmmmm”记作“mmm”。阿坎语中的某些词可以有不同拼写，但含义相同。例如，Tɛkyerɛma 等同于 Tɛkerɛma，soronko 等同于 sononko，而 nneɛma 等同于 nnoɔma。不过，需要注意的是，词语的口语形式可能与其标准书写形式不同。例如，Ɔmo 在阿坎语标准书写形式中写作 wɔn。转录员应遵循正确的规范。主要指南见表2。

表2：转录指南

可接受的做法	不可接受的做法
如实转录所听到的语音，确保 主要说话者的所有话语都被 并且保留每个单词的原始拼写 。	不要纠正任何语法错误，并避免 转录包含背景噪音或 次要说话者重叠语音的音频。
纳入自我重复；如果说话人重复了某个单词， 短语或句子，都要按说话内容准确转录。	不要省略重复，即使它们看起来 没有必要。
通过转录所有内容来记录自我纠正， 包括说话人更改某个词或短语，或 在句中进行自我纠正的情况。	不要省略自我纠正；一个词的两次尝试 或短语的两次尝试都应保留。不要修改词语 顺序；保留语音的自然流畅性。
每次听3到4个词，并分小段进行转录 以确保准确性。	不要试图记住听到的内容，然后再回想 之后在输入时。
使用标准的阿坎语标点符号和大小写 规则。For 例如，写作 Tɛkyerɛma，而不是 of T3kyer3ma，或使用 Nnoɔma，而不是 Nnocma。	不要以违反标准阿坎语规则的方式更改标点符号、大小写或字符 以一种违背标准阿坎语规则的方式。

通讯作者：aoekpezu@st.ug.edu.gh

将数字拼写出来，并将电话号码和身份证号码标记为个人身份信息（个人身份信息）。	不要使用数字，也不要转录电话号码或身份证号码。
将英语单词音译成阿散蒂语-特定形式（例如，写成“kɔmputa”而不是“computer”，即使说话人说的是 Computer）。	请勿自动写成英语单词，如果你不知道其阿散蒂语特有形式。请将其标记用关键词“language”。这将在转录验证阶段中处理。

言语障碍常呈现出一些独特模式，例如拉长词、拆分词、未完成的词、与标准不流畅现象相关的重复，以及口吃相关阻塞和重复。为确保在整个数据集中一致处理这些模式，转录员与研究团队经过多轮讨论，就每种模式的适当转录方法达成共识。这一协作过程形成了一套针对言语障碍的专门指南，确保所有转录文本保持一致且准确。表3列出了这些言语模式及相应的建议操作。

表3：言语障碍的特定模式及其转录指南

特定模式	建议操作
拉长词	为完整呈现话语的声学形式，应尽可能保留拉长的声音模式。因此，应按照词语被拉长的实际形式进行转录，并发音。例如，如果说话人说“baaaanku”，转录应保留拉长的形式（例如“baaaanku”），而不是将其缩短为“banku”。
拆分词	按说话人实际发出的方式准确转录单词或短语。例如，“Anidie”（意为尊重）是正确且预期的发音，应当在作为单个单位说出时转录为“Anidie”。但是，如果说话人由于其语音模式会将词语分成两部分，例如“Ani die”（字面意思是“眼睛吃”），那么转录应保留这种拆分形式。这样可以确保转录文本准确反映说话者的发音，而不是符合语言规范的形式。
未说完的词语（不完整的发音）	根据说话人的意图转录完整的单词，并利用以下内容中的上下文线索：音频或关联图像。例如，如果说话人说的是“ased”，但描述的上下文表明其意图是说“aseda”，则转录应体现完整的目标单词（“aseda”）。
重复 for 标准不流畅现象	将所有重复都按说话内容准确转录，确保每个重复的单词、短语，或片段都完整呈现在转录文本中。例如，如果说话人说：“I want I want banku”，转录应保留重复内容，写作：“I want I want banku。”这种方法旨在确保自然发生的言语不流畅现象得到准确应予以呈现，不得简化或省略。
基于语流不流畅的口吃 and 阻塞	按原样转录所有重复，保留每次口吃或阻塞。例如，如果说话人说“Yɛ yɛ yɛ yɛfrɛ”，转录应完整保留实际生成的整个序列。这种方法确保所有不流畅的模式都能被准确记录，不遗漏、不平滑处理，从而反映说话人言语障碍语音的真实声学特征。

鉴于部分言语障碍录音难以辨识，专家和研究团队认识到，有必要制定明确且一致的指南来处理不确定情况。经过讨论，双方商定了一套框架，列明转录员可能无法确定音频文件内容的各种情形，以及遇到这些情况时应采取的适当操作。这些指南旨在确保只有转录员有把握的录音才会纳入数据集，从而避免转录文本质量低劣或不准确。这种系统化的方法还有助于高效识别并筛除被标记为存在问题的文件，这些文件随后在转录后阶段被排除。表4列出了主要的不确定情形、相应的建议操作及标记规范。

联系邮箱：aoekpezu@st.ug.edu.gh

表4：转录过程中处理不确定情况和问题的指南

场景	操作	标记方法 （**输入关键词 “问题：用自己的话描述问题 尽可能多地转录词语”**
无法听清的音频或不清晰的语音	多次聆听该片段并调整速度。如果仍然不清晰，请标记该音频。	“问题：无法听清”或“问题：语无伦次”
音频所用语言 a 超出指定范围	语言 如果音频所用语言不是 than 阿坎语, do not 尝试转录。	如果已知另一种语言，例如加语，则输入“Issue: Language [Ga]”或“Issue: Language[N/A]”（如果语言未知）
不确定性 关于转录准确性	标注不确定程度 即对转录内容的不确定程度。	“问题：比较确定，但仍有一些疑虑”或“问题：非常不确定”或“问题：无法理解"
脏话 or 冒犯性语言	请勿转录包含以下内容的音频亵渎、粗俗或仇恨言论。	“问题：亵渎的/粗俗的”或“问题：仇恨言论/冒犯性语言”
存在 of 电话号码 or 个人标识符（PIID）	标记音频中的电话号码和其他个人身份标识符。	“问题：个人身份信息”

3 数据描述

3.1 数据集摘要

UGAkan-ImpairedSpeechData是一个数据集，包含14,312个音频文件及其对应的转录文本，总时长为50.01小时。这些音频文件收录了来自四种不同障碍病因的自发图像描述，分别是口吃、脑瘫、腭裂和中风所致言语障碍；每种病因都呈现出独特的声学和语言特征。录音在不同环境中完成，包括户外（7,706个音频文件）、未注明环境（3,075个）、室内（2,254个）、录音室（982个）和车内（295个）。

UGAkan-ImpairedSpeechData 分为五个子文件夹：四个文件夹分别对应不同的病因，另有一个图像文件夹，其中包含用于引导描述的提示。每个病因文件夹都包含：

i. 一个“audios”文件夹，用于存放 MP3 音频文件；以及ii. 一个元数据 CSV 文件，其命名格式为 *metadata_[aetiology].csv*（例如，中风类别的文件名为 *metadata_stroke.csv*），其中提供了每个音频文件的结构化信息。

每个元数据文件包含十三项属性，用于描述说话人、录音环境和音频特征。这些属性包括原始文件名（orig_file_name）、障碍类型（condition）、性别、说话人标识符（speaker_id）、图像提示（images）、录音环境、年龄、所用设备、转录文本（text）、话语标识符（utterance_id）、数据收集来源（例如阿克拉或库马西）以及时长（秒）。图1展示了 UGAkan-ImpairedSpeechData 语料库的可视化结构。

联系邮箱：aoekpezu@st.ug.edu.gh

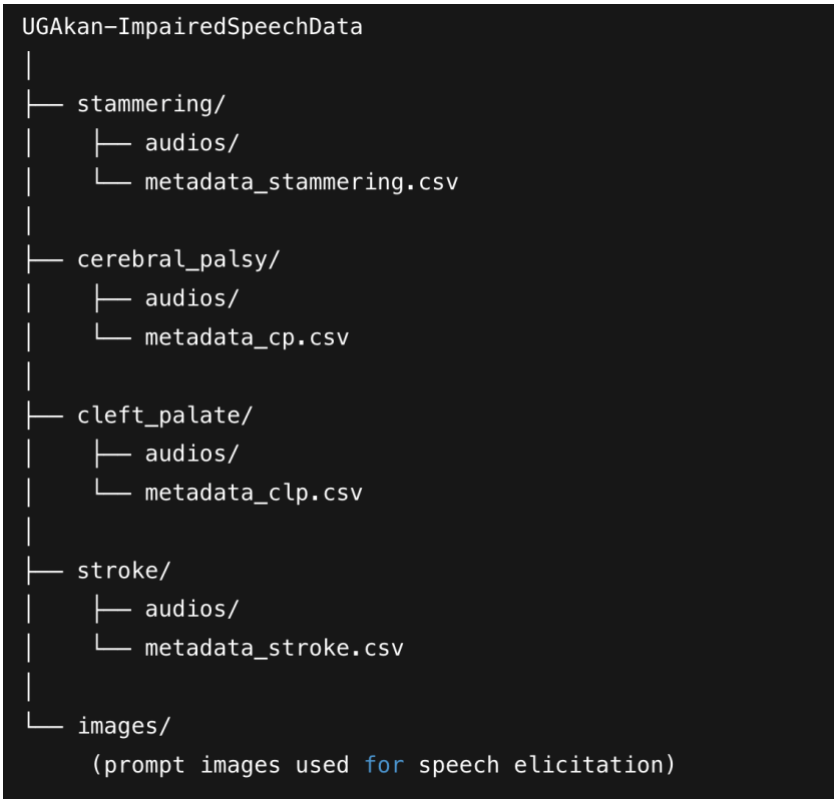


图3. 言语障碍语料库的可视化结构

表1按病因、性别、音频文件数量和总时长（小时）汇总了该数据集的分布。脑瘫子集规模最大（251 MB），其次是口吃（198 MB）、腭裂（55.2 MB）和中风（7.53 MB）。平均录音时长为 12.46 秒，大多数样本时长介于 6.59 秒至 16 秒之间。

表5：按病因和性别划分的音频文件分布及总时长（小时）

	病因	音频文件总数	折合小时数
Male	脑瘫	2,881	8.84
	腭裂	1,677	4.93
	中风所致言语障碍	295	0.74
	口吃	1,901	5.51
女性	脑瘫	4,835	15.66
	腭裂	149	0.45
	卒中后	0	0
	口吃	2,574	13.88
唯一文件总数		14,312	50.01

3.2 数据价值

i. 该数据集同时涵盖低资源语言（阿坎语）和低资源语音领域（言语障碍）。尤其是，它提供了首个公开发布的阿坎语障碍语音数据集。这些代表性不足的领域相结合，有助于推动加纳及其周边阿坎语国家的包容性语音技术发展。

ii. 该数据集支持自动障碍语音识别（ADSR）研究，可用于将现有自动语音识别（ASR）模型适配于言语障碍。它尤其有助于开发能够

通讯：aoekpezu@st.ug.edu.gh

准确识别阿坎语中脑瘫、腭裂、口吃和卒中后言语障碍患者言语模式的系统。

iii. 该数据集可用于为言语障碍人士设计辅助交流工具，也可用于开发辅助技术，以解读或增强障碍人士的语音，从而改善他们在临床、教育或日常社交场景中的交流。

iv. 该数据集涵盖四种不同的言语障碍病因，因此可支持多种研究应用。研究人员、临床医生和技术开发者可利用该数据集，在低资源环境中构建并评估包容且扎根于当地文化的语音识别与合成 AI 模型。该数据集包含脑瘫、腭裂、口吃和中风患者的数据，可用于比较不同障碍类型的言语模式。它也支持对多种病症下自动语音识别模型泛化能力的研究。研究人员可重复利用这些数据，为自动语音识别、社会语音学和临床语言学领域的模型进行训练、基准测试或数据扩充。

v. 该数据集包含 50.01 小时的音频录音及相应的转录文本。语音与文本的对齐使其可用于自动语音识别（ASR）、文本转语音（TTS）合成及其他自然语言处理（NLP）任务等语言技术应用。

4 局限性

本研究的一项主要局限是，由于招募严重言语障碍人士面临社会和后勤限制，因此采用了便利抽样。这限制了我们在所有严重程度和障碍类型中纳入均衡且多样化代表样本的能力。招募某些群体的参与者较为困难，例如患有腭裂的人士，尤其是正在中风后康复的人士。中风病例的代表性最不足，仅收集到296条记录，且全部来自同一性别。在涉及脑瘫等运动障碍的情况下，无法采用无人监督或远程录音等可扩展的数据收集方法。理想情况下，在受控录音室环境中录制严重障碍人士的语音，有助于更好地捕捉其言语模式。然而，这并不实际，因为会大幅延长数据收集周期。此外，该数据集不包含极重度障碍人士的音频录音。这进一步缩小了数据集中所涵盖的障碍严重程度范围。这些局限可能会导致抽样偏差，降低数据集的可推广性，尤其会影响针对障碍最严重且代表性不足的言语障碍病例开发模型。

5 伦理声明

本研究已获得加纳大学基础与应用科学伦理委员会（ECBAS）的伦理批准。所有参与者均已充分了解音频数据收集的目的，以及开发语音数据集可能带来的益处。尽管收集了姓名和电话号码等个人信息，但这些信息仅用于行政管理和发放补偿。

数据可访问性 存储库名称：UGAkan-ImpairedSpeechData：阿肯语言语障碍数据集
数据标识号：10.17632/vc84vdw8tb.3

通讯作者：aoekpezu@st.ug.edu.gh

数据直接 URL：https://data.mendeley.com/datasets/vc84vdw8tb/4 **相关研究文章** 数据集的一半用于评估和调整 ASR 模型 [4]。

6 致谢

本研究由 Google Research Ghana 资助。我们谨向所有为本项目成功作出贡献、但不符合作者署名标准的个人表示感谢。特别感谢言语与语言治疗师、语言专家以及提供技术支持的人员，他们投入时间和精力，确保数据集的质量与完整性。我们也感谢在参与者招募过程中给予支持、并在数据收集期间提供后勤协助的社区领袖、机构和照护者。

7 参考文献

[1] P. Azunre, S. Osei, S. Addo, L.A. Adu-Gyamfi, S. Moore, B. Adabankah, B. Opoku, C. Asare-Nyarko, S. Nyarko, C. Amoaba, E.D. Appiah, F. Akwerh, R.N.L. Lawson, J. Budu, E. Debrah, N. Boateng, W. Ofori, E. Buabeng-Munkoh, F. Adjei, I.K.E. Ampomah, J. Otoo, R. Borkor, S.B. Mensah, L. Mensah, M.A. Marcel, A.A. Amponsah, J. Ben Hayfron-Acquah, 《英语—契维语平行语料库：用于机器翻译》，载于：非洲自然语言处理研讨会 @EACL 2021，2021：第9页。[2] A. Owusu, K. D., Q. A., N. B., S. S., D. D., Adjepon-Yamoah, O. Boateng，GitHub - Ashesi-Org/Financial-Inclusion-Speech-Dataset：由阿散蒂大学和 Nokwary Technologies 创建、旨在支持金融普惠的语音数据集，资金由 Lacuna Fund 提供。[在线] GitHub，（2022）。[3] P. Resnik, M.B. Olsen, M. Diab, 《圣经》作为平行语料库：为《两千种语言之书》添加标注，《计算与人文》33（1999）129–153。[4] S.A. Salihs, I. Wiafe, J.-D. Abdulai, E. Doe Atsakpo, G. Ayoka, R. Cave, A.O. Ekpezu, C. Holloway, K. Tomanek, F. Baffoe, P. Winful, 《低资源语言障碍语音社区驱动数据收集指南》，载于：August（编），《InterSpeech 2025 论文集》，2025年8月17日至21日，鹿特丹，荷兰，鹿特丹，荷兰，2025：第4623–4627页。[5] I. Wiafe, J.-D. Abdulai, A.O. Ekpezu, R.D. Helegah, E.D. Atsakpo, C. Nutrokpor, F.B. Payin Winful, K.K. Solaga，推进加纳低资源语言的自动语音识别：阿坎语、埃维语、达格巴尼语、达加雷语和伊克波索语音频数据集，《数据简报》61（2025）111880。[6] I. Wiafe, J.-D. Abdulai, S.A. Salihs, A.O. Ekpezu, E.D. Atsakpo, F.B.P. Winful, G. Ayoka, K. Tomanek, R. Cave, C. Holloway, UGAkan-ImpairedSpeechData：阿坎语言语障碍数据集，3（2025）。

通信邮箱：aoekpezu@st.ug.edu.gh