Bridging Literacy Gaps in African Informal Business Management with Low-Resource Conversational Agents

Maïmouna Ouattara, Abdoul Kader Kaboré, Jacques Klein and Tegawendé F. Bissyandé
University of Luxembourg
{maimouna.ouattara, abdoulkader.kabore, jacques.klein, tegawende.bissyande}@uni.lu

Abstract

Position paper: In many African countries, the informal business sector represents the backbone of the economy, providing essential livelihoods and opportunities where formal employment is limited. Despite, however, the growing adoption of digital tools, entrepreneurs in this sector often face significant challenges due to lack of literacy and language barriers. These barriers not only limit accessibility but also increase the risk of fraud and financial insecurity. This position paper explores the potential of conversational agents (CAs) adapted to low-resource languages (LRLs), focusing specifically on Mooré, a language widely spoken in Burkina Faso. By enabling natural language interactions in local languages, AI-driven conversational agents offer a promising solution to enable informal traders to manage their financial transactions independently, thus promoting greater autonomy and security in business, while providing a step towards formalization of their business. Our study examines the main challenges in developing AI for African languages, including data scarcity and linguistic diversity, and reviews viable strategies for addressing them, such as cross-lingual transfer learning and data augmentation techniques.

1 Introduction

Commerce, particularly in the informal sector, plays an essential role in the economy of most African countries. A significant portion of the population is involved in informal trading activities, which include small-scale retail, street vending, and local artisan markets (Martínez and Short, 2022). In urban areas, vibrant markets abound, where vendors offer a wide array of goods, from fresh produce to handmade crafts, providing a livelihood for many families. The informal sector is especially vital for job creation in contexts where formal employment opportunities remain scarce (Martínez and Short, 2022). Recently, technological advancements have significantly transformed informal commerce.

Problem Statement. Consider, for example, an entrepreneur in the bustling retail sector who operates a successful shop specializing in local crafts and essential goods. Leveraging digital tools, they use mobile payment[1] platforms to streamline operations and improve customer convenience, enabling fast and accessible transactions via smartphone technology. However, despite the operational advantages of these digital platforms, the businessman may face challenges in exploiting the technology due to limited literacy skills and the fact that most of the existing solutions are not inclusive to enable easy use in case of linguistic barriers. This can potentially lead them to seek assistance from other people. As a result, they often may rely on others for instance to check their balance or make payments on their behalf, a practice which, while necessary, presents potential security risks. This dependence on outside help unintentionally increases their vulnerability to fraud, as sharing sensitive information with third parties can compromise their financial security (Anthony et al., 2024). The potential of technology to empower is then undermined when these tools remain inaccessible to a large portion of the population. With recent advances, Artificial Intelligence (AI) has the potential to act as a powerful lever, enabling accessible technology use through conversational agents that communicate in users' native languages, making digital tools universally approachable.

AI-based Conversational Agent as a Solution. Conversational agents (CAs) are software systems designed to simulate interactions with real people (Khatri et al., 2018). They interact with users using written or spoken natural language, as well as

[1]With a large portion of the population lacking access to traditional banking services, mobile payment platforms have emerged as a vital tool for facilitating transactions (Osabutey and Jackson, 2024).

Proceedings of the First Workshop on Language Models for Low-Resource Languages (LoResLM 2025), pages 219–229
January 20, 2025.

gestures and other non-verbal expressions (Mariani et al., 2023). Recent AI-powered agents, such as Amazon's Alexa, Apple's Siri, and Google Assistant, have become popular around the world due to their ability to help users with everyday tasks. Indeed, AI-powered CAs can perform tasks such as setting reminders, checking balances, and answering questions through simple voice commands or text input. Then, they reduce the need for extensive technological knowledge, making digital interactions more accessible to users across varying literacy and technological levels. However, despite the rapid development of conversational AI, most agents are designed to work effectively in high-resource languages such as English, Spanish and Mandarin. African languages are significantly under-represented in technology, despite the fact that Africa is home to around a third of the world's languages. This under-representation is largely due to the fact that these languages are Low-Resource Languages (LRLs). So, millions of native speakers in Africa are therefore unable to use technology tools effectively in their daily or professional interactions because they speak LRLs[2]. The development of CAs in LRLs, particularly African languages, would enable greater inclusion in communities as they would enable individuals to use technology in their native language, creating more personalized and accessible interactions that promote financial independence and business autonomy (Magueresse et al., 2020). For example, using a conversational agent, the entrepreneur could verbally request his account balance in his own words. The agent would then respond by providing the requested information via voice output, eliminating the need for external assistance and protecting the entrepreneur from the vulnerabilities that this entails.

The main challenge in developing CAs for African languages is data scarcity. These languages often lack the datasets needed to effectively train AI models, and they often lack the resources to create and collect sufficient data for language processing models. A another challenge is the diversity of these languages. African languages generally have a wide range of accents and dialects. Even within the same language, pronunciation, vocabulary and grammatical structures can vary considerably from one region to another and from one social group to another within the same region (way). These variations can result in misunderstandings and misinterpretations by Natural Language Processing (NLP) models.

Contribution This position paper lays the groundwork for developing an AI-based conversational agent for low-resource African languages (LRLs). We focus on Mooré (also known as Moré), the most widely spoken national language in Burkina Faso, spoken by 52.9% of the country's 20.5 million people (INSD, 2029[3]). Mooré is the native language of the Mossi people and belongs to the Niger-Congo language family's Gur (Voltaic) subgroup. While prevalent in Burkina Faso, Mooré is also spoken in neighboring Benin, Côte d'Ivoire, Ghana, Togo, and Mali. Despite its widespread use, Mooré remains a low-resource language due to its primarily oral tradition, with limited written resources available. We explore in particular, the potential solutions for developing conversational agents for LRLs like Mooré, particularly those designed to assist with informal business management as a sweet spot for adoption of these agents. Indeed, the informal sector is keen for adopting innovations that could add value to their business. Yet, it is also the place where innovation is hardest to implement due to the high literacy rates. Conversational agents then constitute a formidable bridge if we can overcome the challenges related with LRLs. By analyzing the unique challenges posed by LRLs, we investigate how state-of-the-art NLP techniques can be adapted and applied to overcome these limitations. Our methodology involves identifying key challenges, such as data scarcity, language model adaptation, and cultural nuances, and then proposing tailored solutions based on relevant NLP techniques. Our work makes the following contributions.

- We highlight the need for more inclusive solutions and discuss how AI-based conversational agents for low-resource languages can serve as a bridge to closing literacy gaps in Africa, empowering marginalized communities, and promoting digital inclusion.
- We establish a foundational framework for developing AI-based conversational agents tailored for low-resource languages, with a focus on addressing the unique challenges posed by linguistic scarcity and complexity.

[2]Africa and India collectively host approximately 2,000 low-resource languages and are home to over 2.5 billion inhabitants (Magueresse et al., 2020).

[3]https://www.insd.bf/fr/resultats

- We propose adapted solutions to overcome the challenges of low-resource languages by leveraging state-of-the-art techniques in natural language processing (NLP), including data augmentation and multilingual model integration.

2 Background

2.1 Low-Resource Languages

According to UNESCO's World Atlas of Languages[4], there are 8,324 languages (spoken and signed) documented by governments, public institutions and academic communities, of which around 7,000 are still in use. However, most current NLP research focuses on 20 of the world's 7,000 languages (Magueresse et al., 2020). Most of the world's languages are therefore LRLs.
Over the past decade of efforts to create language resources for under-served languages, several terms have emerged to describe these languages, including 'low density', 'less commonly taught', 'under-resourced' and 'under-resourced' (Cieri et al., 2016). (Magueresse et al., 2020) defined Low-Resource Languages (LRLs) as languages that are *less studied, resource-scarce, underrepresented in digital formats, and less commonly taught*. In this paper, the term 'LRLs' refer to languages that exhibit one or more of these characteristics, with a particular emphasis on data scarcity. We focus on African LRLs.

African low-resource languages LRLs have unique characteristics and challenges that impact their representation in technology and natural language processing (NLP). Here is an overview of the main characteristics. Limited digital resources and data scarcity are major challenges for the development of NLP models for African languages (Thangaraj et al., 2024). The digital presence of many African languages is largely limited to informal sources, such as social media, which complicates data collection and processing. As a result, there is a critical lack of the large annotated datasets required for effective model training. This lack affects a variety of key resources, including digital text corpora, speech transcription datasets and labelled data tailored to specific NLP tasks, hampering the ability to develop robust linguistic technologies for these languages. In addition, African languages are often characterised by high linguistic complexity (Thangaraj et al., 2024). Many have agglutinative or highly inflectional morphology, where a single word can encode multiple layers of meaning through prefixes, suffixes or internal modifications. This morphological richness poses problems for NLP tasks such as tokenization and stemming, as standard techniques can struggle to break down these complex structures accurately. Many African languages also rely on tonal distinctions, i.e. variations in pitch that can completely change the meaning of a word. Accurately capturing pitch in written and spoken data is difficult, especially as pitch marks are often omitted from informal texts, leading to ambiguity and potential errors in training data. African languages are also often characterised by limited access to standardised writing systems, largely due to the predominance of oral traditions over written literacy. Many of these languages lack standardised orthographies and consistent conventions for spelling, punctuation and grammar (Thangaraj et al., 2024). This lack of widely accepted standards complicates data processing and poses consistency problems for NLP applications, as variations in written forms can lead to inconsistencies in model learning and evaluation. These languages also encompass a wide range of dialects, with significant regional variations in vocabulary, grammar and pronunciation (way). This linguistic diversity complicates the development of standardised NLP models that work reliably across all dialects, as models trained on one dialect do not necessarily generalise to others.

2.2 Conversational Agent

The concept of machines interacting with humans in a conversational way originated with the Turing test in 1950. The practical implementation of this concept began with early systems such as ELIZA (Weizenbaum, 1966) and PARRY (Colby, 1981), which relied on rule-based approaches; these systems used predefined rules and templates to process user input and generate responses. However, advances in artificial intelligence have since enabled the development of conversational AI, a specialised area of AI. Conversational AI is defined as "the study of techniques for developing software agents capable of engaging in natural conversational interactions with humans" (Khatri et al., 2018).

Conversational AI leads to AI-powered conversational agents (CAs), which are "software systems designed to mimic interactions with real people"

[4] https://unesdoc.unesco.org/ark:/48223/pf0000380132

through conversation in written and spoken natural language, as well as through gestures and other nonverbal expressions (Mariani et al., 2023). These systems are referred to by several terms based on their application and functionality, such as chatbots, smart bots, intelligent agents, conversational user interfaces, conversational AI systems, personal digital assistants, virtual personal assistants, or dialogue systems (Kusal et al., 2022). Conversational agents (CAs) are versatile tools employed across various domains to perform a wide range of valuable tasks. In the business sector, they are widely used for marketing, engaging customers through personalized interactions, and providing 24/7 customer support (Bavaresco et al., 2020). In healthcare, CAs function as personal health assistants, reminding patients to take medications, scheduling appointments, delivering medical information, and offering preliminary health (Laranjo et al., 2018). In the education sector, these agents serve as personal tutors, assisting students with homework, explaining complex concepts, offering study tips, and supporting language learning (Darvishi et al., 2024). Within the entertainment industry, CAs enhance user experiences by assisting players in digital games (Kusal et al., 2022).

2.3 Conversational Agents for Low-Resource Languages

While popular conversational agents such as Amazon Alexa, Apple Siri and Google Assistant primarily support high-resource languages (HLRs), they have begun to include a limited selection of LRLs. For example, Google Assistant[5] now supports Swahili, a language widely spoken in East Africa with over 16 million native speakers, as well as Hindi and Indonesian. Apple Siri[6] supports Malay and Thai, although functionality in these languages is more limited than in HLRs, often limiting users to simple voice commands. Siri also supports Hebrew and Arabic, languages that present unique challenges due to the distinct directionality of the script and complex phonetic structures. Amazon Alexa[7], although its range of low-resource languages is more limited, supports Hindi, improving accessibility for speakers in India. Although these platforms are making progress in terms of inclusion, support for LRLs remains limited. In particular, the availability of localized responses, the recognition of dialectal variations and the handling of complex linguistic features typical of African and other LRLs are often insufficient, resulting in less robust functionality than for HLRs.

African languages remain underrepresented in the field of conversational AI, although recent studies are increasingly exploring the feasibility of developing conversational agents for these languages. For example, (Awino et al., 2022) developed a Swahili conversational AI voicebot for customer support tasks, while (Adewumi et al., 2023) created a corpus to investigate cross-lingual transfer for dialogue generation in African LRLs. (Ogundepo et al., 2023) introduced a cross-lingual question-answering dataset with over 12,000 questions in 10 geographically diverse African languages. (Ogueji et al., 2021) examined the viability of Transformer-based multilingual language models, pretrained from scratch, for 11 African languages. Additionally, several community-driven initiatives and research groups, such as the Masakhane Research Foundation[8], KenCorpus[9], and Ghana NLP[10], are focused on building NLP models tailored for African languages.

3 Addressing Challenges in Low Resource Language Conversational Agents

Researchers have investigated various solutions for overcoming linguistic and resource limitations in developing conversational agents for low-resource languages (LRLs). This section presents some of these approaches.

3.1 Data Augmentation

The lack of data is a major obstacle to the development of effective conversational agents in LRLs. Data augmentation techniques come to the rescue by creating synthetic data or manipulating existing data to enrich the training dataset. This part presents some common techniques.

Back-translation is a technique that uses HRL to create synthetic data for the target LRL by translating and back-translating sentences. For example, (Adewumi et al., 2023) used this method to create a dialogue dataset for six African languages from MutiWOZ (Budzianowski et al., 2018), an English dialogue dataset. This technique allows additional training data to be created, capturing aspects

[5] https://assistant.google.com/

[6] https://www.apple.com/siri/

[7] https://www.amazon.com/b?node=21576558011&ref_=alxcom_lrnmore_btn_23

[8] https://www.masakhane.io/

[9] https://kencorpus.maseno.ac.ke/

[10] https://ghananlp.org/

of the original language structure, and is particularly useful when domain-specific data for LRLs is limited. However, back-translation can also introduce errors or biases from the HRL, hence the need for careful selection of the LRL to ensure structural compatibility and minimise potential distortions.

Synonym replacement (Kolomiyets et al., 2011) is a word substitution technique that describes the paraphrasing transformation of text instances by replacing certain words with synonyms to create variations in sentences.

Synthetic data generation is a technique used to create artificial data, such as text conversations or voice recordings, based on predefined rules or models. This approach creates new data from scratch by using generative models, such as GPT (Brown et al., 2020) or other transformer-based models, to produce text that simulates the characteristics and patterns of the target language.

Audio data augmentation techniques such as *noise injection*, *time stretching*, *pitch shifting*, and *reverberation* can be valuable methods for enhancing audio datasets (Wei et al., 2020).

It is essential to ensure that the source data used for augmentation is high quality and error-free. In addition, adapting data augmentation techniques to the specific domain of the conversational agent is essential to achieve optimal results.

3.2 Cross-Lingual Transfer Learning

Cross-linguistic transfer learning is an NLP approach in which knowledge from high-resource languages (such as English) is transferred to low-resource languages in order to improve the performance of models in those languages (Thangaraj et al., 2024). This technique is based on the assumption that languages, particularly those from the same language family, share certain underlying linguistic structures and semantic relationships. By transferring knowledge from a well-trained source language model, it may be possible to improve the learning process of the target language model, resulting in more accurate performance in LRL applications. Cross-lingual transfer capabilities are evaluated in different architectures, such as monolingual and multilingual.

Monolingual models are trained exclusively on a single language, allowing them to better capture linguistic details and nuances. By exploiting language-specific features and resources, these models can achieve higher accuracy in tasks such as translation, text generation, and classification (Thangaraj et al., 2024). (Gogoulou et al., 2022) investigates the feasibility of adapting existing monolingual models to the target language and examines their downstream performance compared to a model trained from scratch in that target language. Their results indicated that knowledge from the source language significantly enhanced the learning of both syntactic and semantic aspects in the target language. However, it can be difficult to find pre-trained models in HRLs for each corresponding LRL, as most pre-trained monolingual models are mainly trained in English, Mandarin, and so on.

Multilingual pre-trained language models such as mBERT (Pires et al., 2019) and XLM-R (Lample and Conneau, 2019), are trained on large datasets in multiple languages, enabling them to generalize and recognize patterns in different languages. By creating shared linguistic representations, these models facilitate the transfer of knowledge from HRLs to LRLs, thereby improving the performance of NLP tasks in low-resource contexts. However, these models are strongly influenced by the datasets on which they are trained. A biased training set that favours large corpora of specific languages may result in sub-optimal performance for under-represented languages (Thangaraj et al., 2024).

Case of African LRLs

African languages face a severe lack of training data and are often under-represented in multilingual datasets. Since the quality and quantity of multilingual data significantly influence the performance of cross-linguistic transfer learning models, the application of this method to African languages presents challenges. In addition, these languages often have complex grammatical structures and high linguistic diversity, further complicating the effectiveness of cross-linguistic transfer. However, recent research has begun to explore solutions for improving cross-linguistic transfer capabilities for African languages. (Ogueji et al., 2021) investigated the feasibility of pre-training multilingual language models exclusively on LRLs, without any transfer from HRLS. They presented AfriBERTa, a transformer-based multilingual language model trained on 11 African languages, which outperforms mBERT and XLM-R in tasks such as text classification and Named Entity Recognition (NER). This study paves the way for the development of multilingual models exclusively pre-trained on African languages.

3.3 Zero-shot and Few-Shot Learning

Zero-learning is a technique that allows a model trained in one language to perform tasks in another language without further fine-tuning (Pourpanah et al., 2023). This approach is both flexible, as it allows the model to perform tasks without task-specific training, and cost-effective, as it eliminates the need for additional training data. Few-shot learning is a technique in which the model requires only a small amount of data in the target language to achieve better performance (Pourpanah et al., 2023). This approach often outperforms zero-shot learning, as it allows to work with contextual data while requiring only a minimum amount of data.

These techniques are essential in cross-lingual transfer learning, enabling models to learn from only a few or even zero examples in the target language by leveraging knowledge from related languages. Multilingual models such as mBERT and XLM-R enable zero-shot or few-shot learning, often achieving strong performance in languages on which they have not been directly trained (Pires et al., 2019) (Lample and Conneau, 2019).

Zero-shot and few-shot learning are particularly advantageous in low-resource scenarios (Kuo and Chen, 2022), as they minimise the need for extensive target language data. These techniques hold promise for addressing data scarcity in the development of conversational agents for African LRLs. However, they come with limitations: zero-shot transfer may lack accuracy when handling language-specific expressions or or specialised and complex top. Furthermore, few-shot learning relies heavily on the quality of the example data provided; if these examples are suboptimal, the performance of the model may be compromised.

4 Methodology

4.1 Approach

In this work, we adopt a modular architecture 1 to design and build the conversational agent (CA), as it is particularly effective for task-based dialogue systems. This architecture decomposes the overall task into a series of sub-tasks, allowing each module to be trained independently, as suggested by (Razumovskaia et al., 2022). The system comprises three primary modules: (1) the Natural Language Understanding (NLU) module, which processes user input to accurately interpret intentions and extract relevant entities; (2) the Dialogue Management (DM) module, which determines the appropriate system actions based on the current state of the conversation; and (3) the Natural Language Generation (NLG) module, which generates contextually relevant responses based on user input. Since our system is voice-based, we also incorporate speech recognition and speech synthesis technologies to enable seamless spoken interactions.

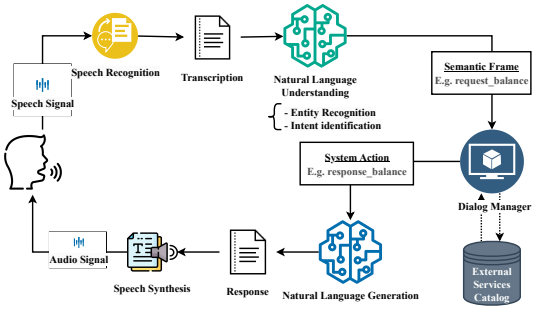

Figure 1: Conversation Agent - Modular Pipeline

Building on this modular foundation, we propose the development of a task-oriented AI conversational agent specifically designed for the commerce sector. The system aims to automate essential tasks such as sales management, stock tracking, and electronic transactions, offering a self-service interface that is particularly useful for underserved populations, including illiterate users. To enhance accessibility, we are considering a voice-based conversational agent that enables users to interact with the system through spoken Mooré. Given that Mooré is a low-resource language with limited digital resources, this approach is intended to bridge the gap in accessibility and usability, particularly in regions where literacy levels are low. The key challenges lie in both the scarcity of data and the linguistic complexity of Mooré. To address these challenges, we propose an innovative approach that combines data augmentation techniques with multilingual models and transfer learning. This strategy will help mitigate the lack of large-scale datasets for Mooré and improve the conversational agent's performance in understanding and generating responses. Currently, there is no publicly available Mooré dataset suitable for training a conversational AI system. As a result, we will undertake data collection efforts to build a comprehensive, domain-specific Mooré dataset. This will be complemented by the application of data augmentation techniques, such as synthetic data generation and language model fine-tuning, to further enhance the dataset's coverage and diversity. In addition, we will leverage

pre-trained multilingual models for various natural language processing (NLP) tasks, focusing on those trained on languages that are linguistically similar to Mooré. By fine-tuning these models, we aim to improve the conversational agent's ability to handle tasks such as intent recognition, slot filling, and dialogue management in the context of a low-resource language. This approach is expected to result in a robust, scalable AI-powered conversational agent that can be deployed in commercial settings to serve a wide range of users, including those with limited literacy skills, while overcoming the challenges posed by linguistic diversity and data scarcity.

4.2 Data Collection and Pre-Processing

Developing CAs for African LRLs like Mooré requires a robust and multifaceted data collection and prep-rocessing strategy to address the scarcity of linguistic resources. A foundational method involves leveraging publicly available text and speech corpora, such as transcripts of radio broadcasts, TV shows and interviews, folk tales, religious texts, and educational materials created in Mooré. These culturally rich sources provide a diverse linguistic base and can be digitized, segmented, and annotated for use in natural language processing (NLP) tasks. This effort can be significantly enhanced through partnerships with local organizations, including non-governmental organizations (NGOs), cultural institutions, and academic groups. Collaborations with these stakeholders offer access to linguistic and cultural expertise, facilitate targeted data collection in urban and rural areas, and ensure datasets are linguistically and culturally accurate through expert validation.

To further enrich the dataset, a community-driven approach via crowdsourcing initiatives is proposed. Engaging native speakers through digital platforms allows for the collection of conversational data and feedback. A mobile-friendly platform could be developed where users contribute voice recordings, translations, and annotations. This can encourage contributions from speakers of various dialects to ensure linguistic diversity. This approach not only expands the dataset but also empowers the community to actively participate in preserving and enhancing their language's representation in technology.

Collaboration with community-driven Projects such as Mashakane, is another critical element. Partnering with such projects enables the use of existing tools and frameworks tailored for low-resource language processing, expands datasets through collaborative community efforts, and provides pre-trained models that can serve as a starting point for developing Mooré conversational agents. These collaborations bring technical expertise and a network of contributors committed to the broader goal of advancing inclusivity in AI for African languages.

The collected data must undergo a rigorous preprocessing phase to ensure its quality and usability for training conversational agents. The first step involves tokenization and normalization of text data to address variations in spelling, grammar, and script usage, creating a standardized format for analysis. This step is crucial for languages like Mooré, which may exhibit significant orthographic and syntactic variations across different speakers and contexts. Additionally, dialect tagging will be employed, where annotators label data according to regional or dialectal differences. This nuanced approach ensures that the final models can capture and respond to the linguistic diversity inherent in Mooré. For audio data, noise reduction is a critical step. Speech recordings will be processed to remove background noise, interruptions, and other non-linguistic sounds that could interfere with the performance of speech recognition systems. This ensures clarity and accuracy in subsequent processing stages. To maintain the dataset's integrity, a robust quality assurance process will be implemented. Linguists and native speakers will validate the dataset to ensure that it is linguistically accurate, culturally relevant, and representative of the Mooré language. This validation step is essential for creating conversational agents that are not only effective but also respectful of the cultural and linguistic nuances of the target community.

4.3 Data Augmentation

We address the data scarcity challenges in developing a conversational agent for low-resource languages by building two distinct datasets, each tailored to specific Natural Language Processing (NLP) tasks essential for our system.

4.3.1 Speech Recognition & Synthesis Dataset

The first dataset (cf. Figure 2) comprises audio recordings paired with corresponding text transcriptions, facilitating the training of both speech recognition and speech synthesis models. Each audio file includes an alignment file that maps audio seg-

ments with their respective transcriptions, ensuring precise matching for effective training. The dataset will be used in the Speech Recognition and Speech Synthesis modules in the system.

To enhance this dataset despite limited data availability, we apply various audio data augmentation techniques, including:

- **Noise Injection**: Adding background noise to audio samples to simulate different environments.
- **Time Stretching**: Modifying the speed of audio without affecting pitch, allowing the model to handle variations in speaking rates.
- **Pitch Shifting**: Changing the pitch of audio samples to account for variations in speaker pitch.
- **Reverberation**: Adding echo effects to simulate different acoustic environments.

These augmentations aim to diversify the dataset, improving the robustness and generalizability of the Speech Recognition and Speech Synthesis models.

4.3.2 Textual Data for NLP Tasks

The second dataset (cf. Figure 3), focusing on textual data, is designed to support tasks like Natural Language Understanding (NLU) and Natural Language Generation (NLG), which are essential for modules such as Natural Language Understanding, Semantic Frame construction, System Action selection, and Natural Language Generation.

To expand this textual dataset and overcome data scarcity, we employ text-based data augmentation techniques, including:

- **Synonym Replacement**: Replacing words with their synonyms to create varied expressions while retaining the original meaning. This technique is particularly useful in Mooré, where NLP resources are scarce, making it an effective yet straightforward augmentation method.
- **Paraphrasing**: Rewriting sentences with alternative phrasings to increase linguistic diversity, providing additional training samples for robust language understanding and generation.

These techniques will enrich the dataset, enabling the Natural Language Understanding and Natural Language Generation modules to better identify user intent, recognize entities, and generate coherent responses in Mooré.

In summary, both datasets and their respective augmentation techniques are designed to address specific challenges in low-resource language processing, enhancing the performance of each module within the conversational agent system.

4.4 Natural Language Processing (NLP) Tasks

To develop a conversational agent (CA) using a modular architecture, several essential NLP tasks are distributed across specialized modules. Each module is designed to handle a specific aspect of language processing, enabling the CA to function effectively by training specialized NLP models independently for each task.

Among all modules, the the Natural Language Understanding (NLU) module is the most challenging. It is responsible for two main sub-tasks: intent classification and slot filling (Razumovskaia et al., 2022).

- **Intent Classification**: This task identifies the user's goal or intent in a conversation, enabling the CA to interpret the purpose behind user input. It can be approached as a classification problem, where each user input is categorized into a predefined intent class, or as a question-answering task to extract specific responses based on user queries.
- **Slot Filling**: This task involves extracting relevant entities or "slots" from user input, such as names, dates, or locations, which are necessary for generating accurate responses. Slot filling is commonly modeled as a span extraction task, where the model identifies and labels key pieces of information in the input text.

For both tasks, we employ cross-lingual transfer learning in zero-shot or few-shot learning settings. This approach leverages pre-trained multilingual models, which have proven effective in low-resource language (LRL) contexts. By transferring knowledge from high-resource languages to Mooré, our target language, we can bypass the scarcity of labeled data. The pre-trained models will be carefully selected based on the linguistic similarity be-

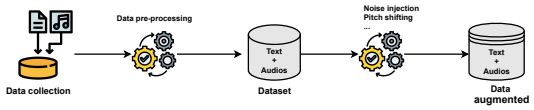

Figure 2: Aligned Text-Audio Data Augmentation

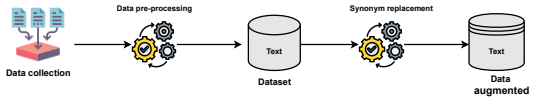

Figure 3: Text Data Augmentation

tween their source languages and Mooré, ensuring effective adaptation.

To inform this model selection, the initial phase of our work involves a linguistic similarity analysis between African languages — particularly Mooré — and various high-resource languages. By calculating similarities in structure, vocabulary, and grammar, this analysis will identify languages that share structural or lexical characteristics with Mooré. This step facilitates the adaptation of resources and methodologies from high-resource languages to low-resource African languages, improving model performance in the CA.

5 Conclusion

This position paper underscores the urgent need for conversational agents (CAs) tailored to low-resource languages (LRLs) in Africa to improve accessibility and security in digital tools for informal commerce. By introducing CAs in languages like Mooré, entrepreneurs could gain independence in managing financial transactions, reducing reliance on third parties and lowering fraud risks.

The paper proposes strategies for addressing challenges such as data scarcity and linguistic complexity, including cross-linguistic transfer learning and data augmentation tailored to low-resource settings. These ideas aim to bridge the digital divide, empowering African language speakers with greater access to technology and financial autonomy. Building on the ideas presented in this position paper, future work will focus on investigating the concrete implementation of these strategies, with an emphasis on data collection and model refinement for African languages to foster digital equity in Africa's informal economy.

6 Limitations and Future work

We acknowledge that our future implementation may present several limitations given the challenges and gaps in the current proposal. One significant limitation lies in data availability and quality as the success of data collection strategies depends on access to high-quality resources, community engagement, and effective partnerships. Furthermore, ensuring linguistic diversity and accuracy across dialects remains a persistent challenge that requires ongoing refinement.

The modular architecture proposed in this work, while flexible, introduces the potential for error propagation, as each module contributes its own error rate to the overall system. This cumulative effect can compromise the accuracy and performance of the conversational agent. Addressing this limitation requires robust monitoring and evaluation mechanisms at every stage of the pipeline.

Finally, this study is focused exclusively on the Mooré language and its application in informal commerce. While this targeted scope facilitates in-depth exploration, it limits the generalizability of the proposed strategies to other African languages. Future work should investigate the scalability and adaptability of these approaches to a broader range of African languages and diverse use cases, thereby enhancing their wider applicability and impact.

7 Ethical considerations

The proposed CA system entails ethical and societal considerations, including ensuring informed consent, data privacy, and fair compensation in data collection strategies such as crowdsourcing and partnerships. Additionally, the agent must prioritize security and reliability to maintain user trust when handling sensitive data like banking information. For instance, integrating voice recognition can enhance security by enabling users to protect their actions and ensure authorized access.

Acknowledgments

This work is supported by the Luxembourg Ministry of Foreign and European Affairs through their Digital4Development (D4D) portfolio under the project LuxWAyS (Luxembourg/West-Atrica Lab for Higher Education Capacity Building in Cyber-Security and Emerging Topics in ICT4Dev).

References

Challenges Dialects Present to Speech Recognition Systems in African Languages — waywithwords.net. https://waywithwords.net/resource/dialects-speech-recognition-systems/. [Accessed 28-10-2024].

Tosin Adewumi, Mofetoluwa Adeyemi, Aremu Anuoluwapo, Bukola Peters, Happy Buzaaba, Oyerinde Samuel, Amina Mardiyyah Rufai, Benjamin Ajibade, Tajudeen Gwadabe, Mory Moussou Koulibaly Traore, Tunde Oluwaseyi Ajayi, Shamsuddeen Muhammad, Ahmed Baruwa, Paul Owoicho, Tolulope Ogunremi, Phylis Ngigi, Orevaoghene Ahia, Ruqayya Nasir, Foteini Liwicki, and Marcus Liwicki. 2023. Afriwoz: Corpus for exploiting cross-lingual transfer for dialogue generation in low-resource, african languages. In *2023 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8.

Aubra Anthony, Nanjira Sambuli, and Lakshmee Sharma. 2024. Security and trust in africa's digital financial inclusion landscape.

Ebbie Awino et al. 2022. *Swahili Conversational Ai Voicebot for Customer Support*. Ph.D. thesis, University of Nairobi.

Rodrigo Bavaresco, Diórgenes Silveira, Eduardo Reis, Jorge Barbosa, Rodrigo Righi, Cristiano Costa, Rodolfo Antunes, Marcio Gomes, Clauter Gatti, Mariangela Vanzin, Saint Clair Junior, Elton Silva, and Carlos Moreira. 2020. Conversational agents in business: A systematic literature review and future research directions. *Computer Science Review*, 36:100239.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. *CoRR*, abs/2005.14165.

Paweł Budzianowski, Tsung-Hsien Wen, Bo-Hsiang Tseng, Iñigo Casanueva, Stefan Ultes, Osman Ramadan, and Milica Gašić. 2018. MultiWOZ - a large-scale multi-domain Wizard-of-Oz dataset for task-oriented dialogue modelling. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 5016–5026, Brussels, Belgium. Association for Computational Linguistics.

Christopher Cieri, Mike Maxwell, Stephanie Strassel, and Jennifer Tracey. 2016. Selection criteria for low resource language programs. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 4543–4549, Portorož, Slovenia. European Language Resources Association (ELRA).

Kenneth Mark Colby. 1981. Modeling a paranoid mind. *Behavioral and Brain Sciences*, 4(4):515–534.

Ali Darvishi, Hassan Khosravi, Shazia Sadiq, Dragan Gašević, and George Siemens. 2024. Impact of ai assistance on student agency. *Computers Education*, 210:104967.

Evangelia Gogoulou, Ariel Ekgren, Tim Isbister, and Magnus Sahlgren. 2022. Cross-lingual transfer of monolingual models. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 948–955, Marseille, France. European Language Resources Association.

Chandra Khatri, Anu Venkatesh, Behnam Hedayatnia, Raefer Gabriel, Ashwin Ram, and Rohit Prasad. 2018. Alexa prize—state of the art in conversational ai. *AI magazine*, 39(3):40–55.

Oleksandr Kolomiyets, Steven Bethard, and Marie-Francine Moens. 2011. Model-portability experiments for textual temporal analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 271–276, Portland, Oregon, USA. Association for Computational Linguistics.

Chia-Chih Kuo and Kuan-Yu Chen. 2022. Toward zero-shot and zero-resource multilingual question answering. *IEEE Access*, 10:99754–99761.

Sheetal Kusal, Shruti Patil, Jyoti Choudrie, Ketan Kotecha, Sashikala Mishra, and Ajith Abraham. 2022. Ai-based conversational agents: A scoping review from technologies to future directions. *IEEE Access*, 10:92337–92356.

Guillaume Lample and Alexis Conneau. 2019. Cross-lingual language model pretraining. *CoRR*, abs/1901.07291.

Liliana Laranjo, Adam G Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian, Blanca Gallego, Farah Magrabi, Annie YS Lau, et al. 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, 25(9):1248–1258.

Alexandre Magueresse, Vincent Carles, and Evan Heetderks. 2020. Low-resource languages: A review of past work and future challenges. *Preprint*, arXiv:2006.07264.

Marcello M. Mariani, Novin Hashemi, and Jochen Wirtz. 2023. Artificial intelligence empowered conversational agents: A systematic literature review and research agenda. *Journal of Business Research*, 161:113838.

Lina Martínez and John Rennie Short. 2022. The informal city: Exploring the variety of the street vending economy. *Sustainability*, 14(12).

Kelechi Ogueji, Yuxin Zhu, and Jimmy Lin. 2021. Small data? no problem! exploring the viability of pretrained multilingual language models for low-resourced languages. In *Proceedings of the 1st Workshop on Multilingual Representation Learning*, pages 116–126, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Odunayo Ogundepo, Tajuddeen R Gwadabe, Clara E Rivera, Jonathan H Clark, Sebastian Ruder, David Ifeoluwa Adelani, Bonaventure FP Dossou, Abdou Aziz Diop, Claytone Sikasote, Gilles Hacheme, et al. 2023. Afriqa: Cross-lingual open-retrieval question answering for african languages. *arXiv preprint arXiv:2305.06897*.

Ellis L.C. Osabutey and Terence Jackson. 2024. Mobile money and financial inclusion in africa: Emerging themes, challenges and policy implications. *Technological Forecasting and Social Change*, 202:123339.

Telmo Pires, Eva Schlinger, and Dan Garrette. 2019. How multilingual is multilingual bert? *CoRR*, abs/1906.01502.

Farhad Pourpanah, Moloud Abdar, Yuxuan Luo, Xinlei Zhou, Ran Wang, Chee Peng Lim, Xi-Zhao Wang, and Q. M. Jonathan Wu. 2023. A review of generalized zero-shot learning methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):4051–4070.

Evgeniia Razumovskaia, Goran Glavas, Olga Majewska, Edoardo M Ponti, Anna Korhonen, and Ivan Vulic. 2022. Crossing the conversational chasm: A primer on natural language processing for multilingual task-oriented dialogue systems. *Journal of Artificial Intelligence Research*, 74:1351–1402.

Harish Thangaraj, Ananya Chenat, Jaskaran Singh Walia, and Vukosi Marivate. 2024. Cross-lingual transfer of multilingual models on low resource african languages. *arXiv preprint arXiv:2409.10965*.

Shengyun Wei, Shun Zou, Feifan Liao, and weimin lang. 2020. A comparison on data augmentation methods based on deep learning for audio classification. *Journal of Physics: Conference Series*, 1453(1):012085.

Joseph Weizenbaum. 1966. Eliza—a computer program for the study of natural language communication between man and machine. *Commun. ACM*, 9(1):36–45.