

Cascades, Journal of the Department of French and International Studies

Cascades : Revue Internationale Du Département De Français Et D'études Internationales

ISSN (Print): 2992-2992; E-ISSN: 2992-3670

www.cascadesjournals.com; Email: cascadejournals@gmail.com

VOLUME 3; NO. 2; OCTOBER, 2025 ; PAGE 116-124



L'intelligence Artificielle au Service de la Revitalisation Linguistique : une étude de cas de la Langue Kalabari

GEORGEWILL DARLYN GOSPELBA

Department of French and International Studies

Ignatius Ajuru University of Education

Rumolumeni, Port Harcourt

Email: darlyngeorgewill@gmail.com

08062882887/08091896658

Abstract

In the face of the global threat of language extinction, Nigerian languages—particularly those of the Ijoid group such as Kalabari—are undergoing a rapid decline, driven by globalization, the dominance of official languages, and the scarcity of appropriate technological tools. This study examines the potential of Artificial Intelligence (AI) in the documentation, preservation, and revitalization of the Kalabari language. The adopted methodology relies on a conceptual analysis that integrates insights from computational linguistics and language revitalization studies, applied within the sociolinguistic and structural context of Kalabari. This interdisciplinary approach enables an exploration of the interactions between technological innovation and local linguistic realities, while identifying practical opportunities for deploying AI in support of linguistic diversity. The findings highlight three major domains where AI can effectively contribute to Kalabari language revitalization. First, the creation of digital resources—such as corpora, dictionaries, and oral archives—facilitates systematic documentation and standardization. Second, the development of adaptive learning tools powered by AI can enhance intergenerational language transmission. Third, integrating Kalabari into everyday digital communication through real-time translation and voice-assisted technologies promotes its active use in contemporary contexts. In conclusion, the study demonstrates that AI provides scalable and adaptable solutions for endangered language revitalization. However, the long-term success of such initiatives depends on the availability of robust linguistic data, the cultural and contextual relevance of technological systems, and the sustained engagement of Kalabari-speaking communities at every stage of implementation. Ultimately, AI emerges as a strategic instrument for safeguarding linguistic heritage—when guided by collaborative, inclusive, and culturally informed practices.

Résumé

Face à la menace mondiale d'extinction des langues, les langues nigérianes, en particulier celles du groupe ijoïde comme le kalabari, connaissent un déclin rapide, sous l'effet de la mondialisation, de la domination des langues officielles et de la rareté des outils technologiques appropriés. Cette étude examine le potentiel de l'intelligence artificielle (IA) dans la documentation, la préservation et la revitalisation de la langue kalabari. La méthodologie adoptée s'appuie sur une analyse conceptuelle intégrant les connaissances issues de la linguistique computationnelle et des études sur la revitalisation linguistique, appliquées au contexte sociolinguistique et structurel du kalabari. Cette approche interdisciplinaire permet d'explorer les interactions entre l'innovation technologique et les réalités linguistiques locales, tout en identifiant des opportunités concrètes de déploiement de l'IA au service de la diversité linguistique. Les résultats mettent en évidence trois domaines majeurs dans lesquels l'IA peut contribuer efficacement à la revitalisation de la langue kalabari. Premièrement, la création de ressources numériques, telles que les corpus, les dictionnaires et les archives orales, facilite la documentation et la normalisation systématiques. Deuxièmement, le développement d'outils d'apprentissage adaptatif alimentés par l'IA peut améliorer la transmission intergénérationnelle de la langue. Troisièmement, l'intégration du kalabari dans la

communication numérique quotidienne grâce à la traduction en temps réel et aux technologies d'assistance vocale favorise son utilisation active dans les contextes contemporains. En conclusion, l'étude démontre que l'IA offre des solutions évolutives et adaptables pour la revitalisation des langues menacées. Cependant, le succès à long terme de telles initiatives dépend de la disponibilité de données linguistiques fiables, de la pertinence culturelle et contextuelle des systèmes technologiques, et de l'engagement soutenu des communautés parlant le kalabari à chaque étape de la mise en œuvre. En fin de compte, l'IA apparaît comme un instrument stratégique pour la sauvegarde du patrimoine linguistique, lorsqu'elle est guidée par des pratiques collaboratives, inclusives et culturellement éclairées.

Mots-clés: intelligence artificielle, revitalisation linguistique, langue kalabari, ijoïde, préservation culturelle

Introduction

Selon les données de l'UNESCO (2023), près de 40 % des langues du monde sont aujourd'hui menacées d'extinction, et l'on estime qu'environ une langue disparaît toutes les deux semaines. Cette tendance alarmante révèle une crise mondiale du patrimoine linguistique et culturel, où chaque disparition efface un mode unique de pensée, d'expression et de transmission des savoirs.

Au Nigeria, qui compte plus de 500 langues, cette menace prend une dimension particulièrement préoccupante. Si les grandes langues comme le hausa, le yoruba et l'igbo bénéficient d'une forte vitalité démographique et sociopolitique, les langues minoritaires du Delta du Niger subissent une marginalisation croissante. Les effets combinés de l'urbanisation, de la mondialisation, de l'éducation monolingue en anglais et de l'absence de politiques linguistiques efficaces fragilisent fortement ces idiomes.

Parmi elles, la langue kalabari, appartenant au groupe ijoïde, illustre parfaitement cette fragilité. Jadis langue de communication intercommunautaire et de transmission culturelle, le kalabari connaît aujourd'hui une érosion marquée de la transmission intergénérationnelle. De plus en plus de jeunes locuteurs se tournent vers l'anglais ou le pidgin nigérian, reléguant leur langue maternelle à des usages rituels, symboliques ou occasionnels.

Malgré les progrès considérables de l'intelligence artificielle (IA) dans le domaine des technologies linguistiques, un vide académique majeur persiste : la plupart des recherches ignorent les langues africaines peu documentées, en particulier les langues tonales comme le kalabari. Ces langues présentent pourtant des structures phonologiques, morphologiques et syntaxiques complexes qui requièrent des approches technologiques adaptées.

Les technologies d'IA — telles que les systèmes de reconnaissance automatique de la parole (ASR), la traduction neuronale automatique (NMT) et les plateformes d'apprentissage basées sur l'IA — offrent un potentiel considérable pour la documentation, l'enseignement et la revitalisation du kalabari. Toutefois, pour être véritablement efficaces, ces solutions doivent être contextualisées et sensibles à la complexité linguistique et culturelle de la langue.

Cet article propose une analyse du rôle que peut jouer l'intelligence artificielle dans la revitalisation linguistique de la langue kalabari. Après avoir exposé le contexte du déclin linguistique au Nigeria et dans le Delta du Niger, il présentera le cadre conceptuel et méthodologique fondé sur la linguistique computationnelle et la revitalisation linguistique. Ensuite, il examinera trois axes d'application de l'IA — la création de ressources numériques, les outils d'apprentissage adaptatifs et la traduction en temps réel — avant de conclure sur les conditions nécessaires à une mise en œuvre efficace, notamment la disponibilité des données, la pertinence culturelle et l'implication communautaire.

Contexte linguistique :

Le kalabari est une langue nigériane appartenant à la famille des langues ijoid, un sous-groupe de la grande famille niger-congo. Elle est principalement parlée dans l'État de Rivers, au sud du Nigeria, notamment dans les zones côtières du Delta du Niger, incluant les communautés de Buguma, Abonnema, Bakana et Tombia.

Malgré son importance culturelle et historique, le kalabari est aujourd'hui classé parmi les langues menacées selon les critères de l'UNESCO. La langue subit une forte pression sociolinguistique du pidgin nigérian, de l'anglais (langue officielle du pays), et des autres langues majoritaires de la région. Ce contact intensif provoque un déclin de la transmission intergénérationnelle, notamment chez les jeunes, qui privilégient souvent l'anglais ou le pidgin pour des raisons d'éducation, d'emploi et de prestige social.

Classification linguistique et caractéristiques

Famille linguistique : Niger-Congo

Sous-famille : Ijoid

Groupe : Ijo (branche occidentale)

Langue : Kalabari (ou Awome, selon certaines variantes dialectales)

Le kalabari est une langue tonale, ce qui signifie que le ton (hauteur de la voix) joue un rôle essentiel dans la distinction du sens des mots. Elle présente généralement deux à trois tons principaux : haut, bas et moyen. Sur le plan typologique, la langue suit l'ordre SVO (Sujet–Verbe–Objet), typique de nombreuses langues nigérianes modernes.

Contexte régional et sociolinguistique

Dans son contexte géographique, le kalabari est entouré de langues voisines telles que l'Okrika, le Nembe, et l'Ikwerre, avec lesquelles il entretient des interactions linguistiques et culturelles constantes. Cette proximité a favorisé le multilinguisme, mais aussi des phénomènes de mélange linguistique et de bilinguisme déséquilibré au détriment du kalabari.

Les efforts de revitalisation linguistique restent encore limités : la langue n'est pas largement enseignée dans les écoles, et peu de ressources technologiques (corpus numériques, dictionnaires modernes, outils d'apprentissage) existent pour la soutenir. Cette situation justifie l'urgence de programmes de documentation, de normalisation et de numérisation du kalabari, notamment à travers l'usage des technologies de l'intelligence artificielle et des plateformes éducatives numériques.

Statut Sociolinguistique du Kalabari

Selon les données de l'Ethnologue (2024), le kalabari compte environ 250 000 à 300 000 locuteurs natifs, principalement concentrés dans l'État de Rivers, au sud du Nigeria. Bien que ce chiffre puisse sembler élevé, la vitalité linguistique est en déclin rapide, car une grande partie des jeunes générations ne maîtrise plus la langue de manière active.

Domaines d'usage

Le kalabari demeure principalement utilisé dans les contextes familiaux et communautaires, notamment :

À la maison, pour la communication entre parents âgés et enfants ;

Dans la culture traditionnelle, à travers les chants rituels, les proverbes, les contes et les cérémonies religieuses ;

Dans les médias locaux, où quelques émissions radiophoniques communautaires conservent un usage partiel de la langue, souvent en alternance avec le pidgin nigérian ou l'anglais.

Cependant, la langue reste quasi absente des domaines formels tels que l'éducation, l'administration, la presse nationale et le numérique. Cette marginalisation contribue à sa perte de prestige et à son recul progressif dans la vie quotidienne.

Facteurs du déclin linguistique

Plusieurs facteurs expliquent la régression du kalabari :

1. Le déplacement linguistique urbain : Dans les zones urbaines comme Port Harcourt, les familles kalabari adoptent majoritairement le pidgin ou l'anglais, perçus comme symboles de modernité et d'opportunité économique.

2. Le manque de ressources éducatives : Absence de manuels scolaires, de dictionnaires actualisés et d'enseignants formés à la didactique du kalabari.

3. La faible politique de valorisation linguistique : Les programmes gouvernementaux favorisent les grandes langues nationales (yoruba, igbo, haoussa), laissant les langues minoritaires comme le kalabari sans soutien institutionnel.

4. La rupture intergénérationnelle : Les jeunes générations comprennent souvent la langue mais ne la parlent plus couramment, limitant sa transmission naturelle.

Défis technologiques et linguistiques

Le kalabari présente des défis particuliers pour les technologies linguistiques modernes

Système tonal complexe : La langue repose sur deux à trois tons distinctifs, ce qui rend difficile la reconnaissance automatique de la parole par les systèmes d'IA standard.

Phonèmes spécifiques : Le kalabari possède des sons rares comme /gb/ et /kp/ (consonnes labio-vélaires), souvent absents des jeux de données phonétiques internationaux.

Absence de grands corpus numériques : Il n'existe presque pas de bases de données textuelles, vocales ou orthographiques suffisantes pour l'entraînement des modèles d'intelligence artificielle.

Variation dialectale interne : Les différences entre les sous-dialectes (par exemple, Abonnema et Buguma) compliquent la standardisation orthographique et la création d'outils automatiques.

Enjeux et nécessité d'outils d'IA personnalisés

Ces obstacles montrent que le kalabari ne peut être intégré efficacement aux plateformes numériques et éducatives qu'à travers des outils d'intelligence artificielle spécifiquement adaptés à sa structure linguistique.

Un modèle linguistique général (comme ceux développés pour l'anglais ou le français) ne peut ni reconnaître correctement les tons, ni interpréter les phonèmes labio-vélaires, ni gérer l'absence de corpus formel.

La revitalisation du kalabari dépend donc d'un effort ciblé : documenter, numériser et entraîner des modèles d'IA locaux capables de respecter ses spécificités linguistiques et culturelles.

Cadre théorique : technologies de l'intelligence artificielle pour la réalisation linguistique

Les technologies de l'intelligence artificielle (IA) jouent aujourd'hui un rôle central dans le développement, la documentation et la revitalisation des langues. Leur application dans le domaine de la linguistique – souvent appelée Traitement Automatique du Langage (TAL) – vise à permettre aux machines de comprendre, analyser et produire des langues naturelles. Ces technologies comprennent notamment la traduction automatique, la reconnaissance vocale, la synthèse de la parole, la modélisation sémantique et les systèmes d'apprentissage linguistique.

Dans le cadre des langues africaines, et plus particulièrement du kalabari, ces outils offrent un potentiel considérable pour compenser l'absence de ressources linguistiques formelles, à condition qu'ils soient adaptés à la complexité phonologique et tonale de ces langues.

Les langues peu dotées et le TAL

Le concept de langues peu dotées (ou low-resource languages) désigne les langues qui disposent de peu ou pas de ressources numériques, telles que des corpus textuels, bases vocales, grammaires numériques ou modèles d'apprentissage supervisé.

Selon Joshi et al. (2020), plus de 85 % des langues du monde sont considérées comme low-resource pour les systèmes de TAL. Dans ce contexte, la majorité des langues africaines – dont le kalabari – ne disposent pas encore des outils nécessaires pour être intégrées dans les plateformes linguistiques modernes.

Les approches de TAL pour les langues peu dotées cherchent à :

Collecter et annoter des données linguistiques locales, développer des modèles d'apprentissage multilingues capables de transférer les connaissances d'une langue riche vers une langue pauvre, et intégrer la reconnaissance automatique de la parole dans des programmes de documentation linguistique.

La reconnaissance vocale automatique et la documentation orale

La reconnaissance vocale automatique (RVA) est une branche essentielle de l'IA linguistique. Elle permet à un système d'identifier et de transcrire la parole humaine en texte. Pour les langues orales menacées comme le kalabari, la RVA offre une opportunité unique de documenter et préserver le patrimoine oral – chants, proverbes, récits, discours – avant leur disparition.

Des projets internationaux comme le Mozilla Common Voice Project illustrent bien cette dynamique. Ce programme open source invite les locuteurs du monde entier à enregistrer leur voix dans leur langue maternelle, afin de constituer une base de données vocale librement accessible pour la recherche et la création d'outils de reconnaissance vocale.

Certaines langues africaines telles que le luganda, le swahili ou le wolof y sont déjà représentées, ce qui démontre la faisabilité d'une participation communautaire pour la création de corpus vocaux.

L'intégration du kalabari dans un projet similaire permettrait de :

sauvegarder ses expressions orales et formes tonales uniques,

créer une base de données linguistique ouverte,

et alimenter des systèmes d'apprentissage automatique capables de reconnaître la structure phonologique particulière de la langue.

Vers une intelligence artificielle adaptée aux langues autochtones

Les avancées récentes en apprentissage profond (deep learning) et en modèles multilingues (comme Whisper d'OpenAI, Facebook XLS-R, ou Google Translate) démontrent la possibilité d'entraîner des systèmes à partir de données limitées, en exploitant des transferts interlinguistiques.

Toutefois, pour les langues comme le kalabari, il est indispensable de concevoir des approches contextualisées et communautaires, qui associent les linguistes locaux, les parlants natifs et les développeurs d'IA.

Ces collaborations représentent une étape décisive vers une réalisation linguistique durable, où la technologie ne se substitue pas à la culture, mais devient un outil de préservation et de transmission du patrimoine linguistique.

Traduction automatique neuronale et limites pour la langue kalabari

La traduction automatique neuronale (TAN) — ou Neural Machine Translation (NMT) — constitue aujourd'hui l'un des progrès les plus remarquables dans le domaine du traitement automatique du langage (TAL). Basée sur des réseaux de neurones profonds, elle permet à une machine d'apprendre les correspondances entre deux langues à partir d'un grand volume de données bilingues (corpus parallèles). Ces modèles apprennent les structures grammaticales, les relations sémantiques et les règles syntaxiques sans intervention humaine explicite.

Des systèmes comme Google Translate, DeepL ou Meta NLLB (No Language Left Behind) reposent sur ce principe et couvrent déjà des centaines de langues. Cependant, ces modèles dépendent fortement de la quantité et de la qualité des données disponibles pour chaque langue.

Limites pour le Kalabari

Pour le kalabari, la traduction automatique neuronale rencontre plusieurs obstacles majeurs :

1. Absence de corpus parallèles : il n'existe pratiquement aucun texte bilingue kalabari-anglais ou kalabari-français disponible en ligne.
2. Complexité tonale et phonologique : les tons, essentiels au sens lexical, ne sont pas représentés dans la majorité des systèmes de transcription actuels, rendant difficile la correspondance entre texte et signification.
3. Orthographe non standardisée : la coexistence de variantes dialectales et d'écritures non uniformes empêche la constitution de données homogènes pour l'apprentissage automatique.
4. Faible présence numérique : contrairement aux langues dominantes, le kalabari reste quasi absent des plateformes éducatives, des médias numériques et des réseaux sociaux, limitant la collecte automatique de données textuelles.

Ainsi, bien que la TAN soit performante pour les langues dotées de vastes corpus (comme l'anglais ou le français), elle demeure inefficace pour le kalabari sans une étape préalable de documentation, normalisation orthographique et numérisation.

Outils pédagogiques assistés par l'intelligence artificielle

L'essor de l'IA a également favorisé la création d'outils d'apprentissage linguistique intelligents, capables d'adapter le contenu au profil et au rythme de l'apprenant. Ces technologies peuvent jouer un rôle stratégique dans la revitalisation du kalabari à travers l'éducation linguistique.

1. Chatbots conversationnels: Les chatbots linguistiques utilisent des modèles de compréhension du langage pour simuler une conversation avec l'apprenant.

Dans le cas du kalabari, un chatbot pourrait proposer des dialogues quotidiens contextualisés (salutations, proverbes, expressions idiomatiques) ;

corriger automatiquement les erreurs grammaticales et lexicales, et fournir une immersion linguistique virtuelle là où les locuteurs natifs sont rares.

Des prototypes similaires existent déjà pour des langues africaines, comme Yorùbá VoiceBot ou Kiswahili Tutor, démontrant la faisabilité d'un modèle comparable pour le kalabari.

2. Jeux linguistiques interactifs (apprentissage ludique)

Les plateformes gamifiées, inspirées d'applications comme Duolingo ou Memrise, rendent l'apprentissage plus engageant grâce à des quizz, des points de progression et des récompenses virtuelles.

Un module kalabari pourrait introduire des mini-jeux basés sur :

la prononciation tonale correcte,

la construction de phrases simples,

la reconnaissance auditive de mots-clés.

Ces jeux encourageraient les jeunes locuteurs à redécouvrir leur langue dans un cadre numérique attractif.

3. Outils de prononciation et de correction vocale

Les outils d'analyse vocale fondés sur la reconnaissance automatique de la parole permettent de comparer la prononciation d'un apprenant à celle d'un locuteur natif.

Pour le kalabari, un tel système aiderait à :

distinguer les variations tonales,

corriger les consonnes labio-vélaires spécifiques (/gb/, /kp/),

et renforcer la maîtrise de la prosodie orale.

Combinés à des bases de données vocales (comme celles issues du Mozilla Common Voice Project), ces outils constitueraient un levier essentiel pour enseigner et revitaliser la langue à grande échelle.

Vers une pédagogie numérique du kalabari

L'intégration de ces outils d'IA — traduction neuronale, chatbots, plateformes ludiques et systèmes de prononciation — ouvre la voie à une pédagogie numérique adaptée aux réalités linguistiques du kalabari.

Toutefois, leur efficacité dépendra d'une collaboration interdisciplinaire entre communautés locales, chercheurs linguistes, ingénieurs en IA et institutions éducatives, afin de garantir la fidélité linguistique et la valeur culturelle des contenus générés.

Application des technologies de l'IA au cas de la langue Kalabari

1. Création d'une base de ressources numériques

La première étape essentielle pour la revitalisation de la langue Kalabari consiste à créer une base de ressources numériques complète. Grâce à la reconnaissance vocale automatique (ASR – Automatic Speech Recognition), il est possible de constituer un corpus oral kalabari à partir des récits, chants et proverbes des anciens. Ces enregistrements, soigneusement transcrits et annotés, formeront la base de données nécessaire pour entraîner des modèles linguistiques spécifiques à la langue.

Par ailleurs, un dictionnaire numérique intelligent pourrait être développé à partir de cette base. À l'aide d'outils d'analyse assistée par l'IA, les mots seraient automatiquement classés selon leur sens, tonalité, usage contextuel et équivalents en anglais. Ce dictionnaire servirait à la fois aux chercheurs, aux traducteurs et aux apprenants.

2. Développement d'outils d'apprentissage interactifs

L'intelligence artificielle offre des opportunités uniques pour l'enseignement du kalabari. On peut envisager la conception d'une application d'apprentissage du kalabari intégrant un retour tonal en temps réel : les apprenants prononcent un mot et l'IA évalue la justesse des tons, un aspect fondamental dans cette langue.

De plus, les exercices linguistiques pourraient s'appuyer sur des contes et fables traditionnels kalabari, favorisant ainsi l'apprentissage contextuel tout en valorisant la culture locale. L'application inclurait également des mini-jeux, des quiz de vocabulaire et un assistant conversationnel virtuel simulant des dialogues de la vie quotidienne.

3. Usage quotidien et communication

Les technologies d'IA peuvent aussi renforcer l'usage quotidien et communicatif du kalabari. Un système de traduction automatique en temps réel Kalabari–Anglais pourrait faciliter la communication entre locuteurs natifs et non-natifs, notamment dans les milieux scolaires, administratifs et touristiques.

De même, la génération automatique de sous-titres en kalabari dans les vidéos permettrait aux membres de la diaspora kalabari de rester connectés à leur langue maternelle. Ces outils encourageraient une pratique vivante et moderne du kalabari dans la vie quotidienne, contribuant ainsi à sa préservation et à sa transmission intergénérationnelle.

Défis, limites et considérations

1. Problèmes de données

L'un des principaux défis dans l'application de l'IA à la langue kalabari réside dans la rareté des ressources numériques. Ce phénomène, souvent appelé le « défi du démarrage à froid » (cold start challenge), se manifeste par le manque de corpus oraux, de textes écrits et de données annotées nécessaires à l'entraînement des modèles d'intelligence artificielle. Sans un volume suffisant de données représentatives, les systèmes d'IA risquent de produire des résultats incomplets ou imprécis, limitant ainsi leur efficacité dans la documentation et l'apprentissage du kalabari.

2. Biais et représentations

Un autre enjeu majeur concerne les biais linguistiques et dialectaux. La couverture inégale des variétés dialectales du kalabari peut entraîner une surreprésentation de certaines formes linguistiques au détriment d'autres, faussant ainsi les analyses et les outils développés. Cette asymétrie menace la richesse interne de la langue et peut contribuer à l'exclusion de certaines communautés locales dont les parlers diffèrent du kalabari standardisé. Il est donc essentiel d'assurer une représentation équilibrée des variantes régionales dans les bases de données et les applications d'IA.

3. Collaboration communautaire

Enfin, la participation active des communautés kalabari est une condition incontournable du succès des projets de revitalisation linguistique. L'approche doit être participative et inclusive, suivant le principe éthique : « Rien sur nous sans nous ». Les locuteurs natifs, enseignants, chercheurs et leaders culturels doivent être impliqués à chaque étape — de la collecte des données à la conception des outils numériques. Cette collaboration garantit la légitimité culturelle des technologies développées et renforce leur acceptabilité au sein de la communauté.

Conclusion et perspectives futures

Synthèse générale

L'intelligence artificielle (IA) représente aujourd'hui un outil stratégique pour la préservation, l'enseignement et la communication des langues en danger telles que le kalabari. En facilitant la documentation numérique à travers la reconnaissance vocale et la transcription automatique, l'IA permet de conserver la richesse orale du patrimoine linguistique. Dans le domaine éducatif, elle favorise la création d'applications interactives d'apprentissage, adaptées aux besoins des jeunes générations. Enfin, sur le plan communicatif, les systèmes de traduction et d'assistance vocale offrent de nouvelles passerelles entre le kalabari et d'autres langues, soutenant ainsi son usage quotidien et sa visibilité à l'échelle mondiale.

Feuille de route en trois phases

Pour concrétiser ce potentiel, une feuille de route communautaire en trois phases est proposée :

Phase 1 : Collecte de données dirigée par la communauté

Organisation de campagnes locales d'enregistrement des récits oraux, chants traditionnels et conversations spontanées avec la participation active des anciens, enseignants et jeunes locuteurs. Ces données constitueront la base d'un corpus linguistique authentique et diversifié.

Phase 2 : Développement de modèles de base ASR et TTS

Conception de prototypes de reconnaissance vocale automatique (ASR) et de synthèse vocale (TTS) adaptés aux particularités tonales du kalabari. Ces outils permettront la transcription et la génération audio en langue locale, facilitant la création d'autres applications éducatives.

Phase 3 : Application d'apprentissage interactive

Déploiement d'une application prototype d'apprentissage du kalabari, intégrant des exercices de prononciation, de compréhension orale et de traduction instantanée. Cette application servirait à la fois d'outil éducatif et de plateforme de valorisation culturelle, renforçant la transmission intergénérationnelle.

Vision d'avenir

En perspective, la réussite de cette initiative dépendra d'une synergie durable entre chercheurs, développeurs, institutions éducatives et communautés kalabari. En combinant innovation technologique et engagement culturel, l'IA peut devenir un véritable levier de renaissance linguistique, assurant que le kalabari continue de vivre, d'évoluer et de s'exprimer dans le monde numérique de demain.

Conclusion

L'étude met en évidence le rôle déterminant que l'intelligence artificielle (IA) peut jouer dans la préservation, la revitalisation et la valorisation de la langue Kalabari, aujourd'hui menacée de disparition. En exploitant des technologies telles que la reconnaissance vocale automatique (ASR), la synthèse vocale (TTS) et la traduction neuronale, il devient possible de documenter le patrimoine oral, de faciliter l'apprentissage intergénérationnel et d'intégrer la langue dans la communication quotidienne.

Toutefois, le succès d'une telle entreprise dépend de plusieurs facteurs clés : la disponibilité des données linguistiques, la représentation équitable des variétés dialectales, ainsi que la participation active des communautés kalabari à chaque étape du processus. Une approche purement technologique, déconnectée des réalités culturelles locales, risquerait de reproduire les inégalités existantes dans le domaine des langues à faibles ressources.

Ainsi, l'IA ne doit pas être perçue comme une fin en soi, mais comme un outil collaboratif au service de la diversité linguistique. En combinant innovation scientifique, implication communautaire et sensibilité culturelle, la revitalisation du kalabari peut devenir un modèle de développement durable pour d'autres langues africaines en danger. L'avenir du kalabari dépendra donc non seulement des avancées technologiques, mais surtout de la volonté collective de transmettre et de faire vivre la langue dans le monde numérique de demain.

Références

- Adegbija, E. (2004). Language policy and planning in Nigeria. *Current Issues in Language Planning*, 5(3), 181–246. <https://doi.org/10.1080/14664200408668258>
- Ajao, T. S., & Abidoye, J. A. (2022). Artificial intelligence and the future of African languages: Potentials and challenges. *African Journal of Language and Technology*, 8(2), 45–63.
- Bamgbose, A. (2011). African languages today: The challenge of and prospects for empowerment under globalization. In *Selected proceedings of the 40th Annual Conference on African Linguistics* (pp. 1–14). Cascadilla Proceedings Project.
- Bird, S. (2020). Decolonising speech and language technology. In *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 3504–3519). Association for Computational Linguistics.

- Blasi, D. E., Anastasopoulos, A., & Neubig, G. (2022). Systematic inequalities in language technology performance across the world's languages. *Frontiers in Artificial Intelligence*, 5, 62–79. <https://doi.org/10.3389/frai.2022.841532>
- Chelliah, S. L., & de Reuse, W. J. (2011). *Handbook of descriptive linguistic fieldwork*. Springer.
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2023). *Ethnologue: Languages of the world* (26th ed.). SIL International.
- Ezeani, I., He, Y., & Xu, Y. (2020). IgboNLP: Building resources for Igbo language technology. In *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 4235–4244). European Language Resources Association.
- Good, J. (2021). The challenges of African language documentation in the digital age. *Language Documentation & Conservation*, 15, 445–470.
- Hale, K. L., Krauss, M., Watahomigie, L. J., Yamamoto, A. Y., Craig, C., Jeanne, L., & England, N. (1992). Endangered languages. *Language*, 68(1), 1–42. <https://doi.org/10.2307/416368>
- Hovy, E., & Spruit, S. L. (2016). The social impact of natural language processing. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (pp. 591–598).
- Mozilla Foundation. (2021). *Common Voice Project: Democratizing speech technology*. Mozilla. <https://commonvoice.mozilla.org>
- Ndimele, O. (2019). *An overview of the Ijoid languages of the Niger Delta*. Port Harcourt: University of Port Harcourt Press.
- Nwankwo, S. N. (2021). Digital revitalization of Nigerian minority languages: Strategies and sustainability. *Journal of African Linguistics and Culture*, 4(1), 67–88.
- Ofulue, C. I. (2020). Technology and African language preservation in the 21st century: A Nigerian perspective. *Nigerian Journal of Linguistics*, 37(1), 29–52.
- Onyima, B. N. (2023). Community participation in indigenous language preservation: Case study of the Kalabari people. *Journal of Sociolinguistic Studies*, 11(2), 122–140.
- Simons, G. F., & Lewis, M. P. (2013). The world's languages in crisis: A 20-year update. In Austin, P. K. & Sallabank, J. (Eds.), *The Cambridge handbook of endangered languages* (pp. 13–34). Cambridge University Press. <https://doi.org/10.1017/CBO9780511975981>
- UNESCO. (2022). *World Atlas of Languages: Language vitality and endangerment framework*. Paris: UNESCO Publishing. <https://www.unesco.org/en/world-atlas-languages>
- Wagner, C., Behnke, F., & Duong, T. (2022). Language technology for low-resource African languages: A roadmap for future research. *Journal of Language Technology and Computational Linguistics*, 37(2), 55–78.
- Wardhaugh, R., & Fuller, J. M. (2021). *An introduction to sociolinguistics* (8th ed.). Wiley-Blackwell.
- Yimam, S. M., Alemayehu, H. M., & De Pauw, G. (2020). Low-resource speech recognition for African languages. In *Proceedings of the 1st Workshop on NLP for African Languages* (pp. 21–29). Association for Computational Linguistics.