

2023年度IRCプロジェクト活動報告書

1. プロジェクト申請所員名

- 1.申請所員： [品川大輔]

2. プロジェクト名称

- 2-1.プロジェクト名称（和）： [AA研『スワヒリ語辞典』のデジタル資源化]
- 2-2.プロジェクト名称（英）： [Building an online digital resource based on the Swahili dictionary published from ILCAA']

3. プロジェクト代表者（・副代表者）名・所属

- 3-1.プロジェクト代表者： [品川大輔]
- 3-2.プロジェクト代表者所属： [AA研]

4. プロジェクトの参加者名・所属

- 4-1.プロジェクト参加者1： [阿部優子]
- 4-2.プロジェクト参加者1所属： [蘭州大学]
- 4-1.プロジェクト参加者1： [加藤幹治]
- 4-2.プロジェクト参加者1所属： [データサイエンス共同利用基盤施設 人文学オープンデータ共同利用センター]

5. 今年度のプロジェクト活動報告

- 本プロジェクトは、1990年から1997年にかけてAA研から出版された守野・中島（編）『スワヒリ語辞典』を電子化したデータをもとに検索可能なデータベースを構築し、オンラインで公開することを目的としている。紙媒体で出版された同辞典の電子化作業自体は、二国間交流事業「バントゥ諸語の過去と現在：ミクロ類型論，歴史比較言語学，辞書学の統合による新展開」(2021-2022: JPJSBP12021230) によってすでに完了している。本プロジェクトでは、それを元にしたテキストデータの構造化とウェブサイト構築を行った。前者については、データの属性情報がタグ付けされているテキストデータを元に、オンラインデータベースとしてより利便性の高い表示を可能にすることを念頭においたデータベースの構築を行った。これは本来、上述の二国間プロジェクトにも関わった若手研究者に謝金業務として依頼する予定のものであったが、事情により、ウェブサイト構築を依頼した業者（君島結氏）にまとめて作業を依頼することになった。また、データの構造化（シンプルな.txtファイルから.jsonファイルへの変換）については、参加研究者の加藤幹治の技術支援によって行われた。一方、上述の事情によってテキストデータのノイズ除去などの細かな作業が十分には達成されていないという課題、また本プロジェクトとは別にオリジナル版の増補改訂作業が進んでおり、その情報を本データベースに統合するという計画もある。これらについては2024年度に本プロジェクトのphase-2として遂行する予定である。

6. プロジェクトの成果となるオンラインリソースの名称・公開URL

- 6-1.成果物1名称（日本語）： [守野・中島（編）『スワヒリ語辞典』オンライン・データベース]

- 6-2.成果物1名称（英語）： [Online lexical database of A dictionary of the Swahili language compiled by Mzee Morino T. and Nassor Nakazima H.]
- 6-3.成果物1名称（その他あれば）： []
- 6-4.成果物1公開URL： [<https://ilcaa-swahili.aa-ken.jp>]