

A Rule-Based Annotation for Ibibio Children's Play Songs

Aniefon Akpan¹, Moses Ekpenyong², Eno-abasi Urua², Margaret Okon³

¹National Institute of Nigerian Languages, Aba,

²University of Uyo, Uyo ³University of Calabar, Calabar

DOI: <https://doi.org/10.60787/jolan.vol6no1.463>

Abstract

We present a rule-based audio annotator developed for the efficient annotation and tagging of Ibibio words. The Ibibio audio annotator is a word tagger that addresses the challenge of time-consumption in annotating audio/video files that can benefit researchers that are faced with the task of annotating large audio/video files. In a preparatory step, the collected data from randomly selected nine Local Government Areas in Akwa Ibom State were transcribed using orthographic transcription. The transcribed Ibibio text was translated into language of larger communication and a bilingual mini Ibibio-English created. An interface was developed using Python software and could be modified to fit related Ibibio languages. The transcribed texts and bilingual mini dictionary were used as input files in the interface created with Python script to tag the file. The procedure of tagging was informed on series of decision-making processes like homonyms, morpheme and part of speech distinction. Eudico Language Annotator (ELAN) software sample textgrid file was manually modified to reflect the correct time start and time end of the selected video file. The audio annotator drastically reduced the time spent in tagging the large Ibibio corpus file.

Keywords: *Translation, Ibibio, Language, Tagging, Annotation*

1. Introduction

Language documentation concerns itself with keeping world's records of linguistic activities (Woodbury, 2003; Himmelmann, 1998) and these linguistic activities cut across all aspects of human interactions within a given language community. Within these linguistics activities exist some language aspects that are threatened and require documentation for posterity (Fort & Chelliah, 2026; Moeller, S. (2026; Seifart et al, 2018). Some of these linguistic activities require the use of video recorders to capture the intrinsic features of those language activities.

However, documentation of a language is insufficient as there is need for annotation of such language data in order to create a comprehensible text file for the

language community, linguists and researchers. Annotation, which is the addition of interpretive linguistic information to the data (Leech, 2005), adds linguistic values to the recorded data for appropriate understanding by the language users and further research into the thematic areas of such documentation.

One of the primary functions of language annotation is making language materials accessible to various users; from the language custodians to the archivists. However, annotation of language data consumes time (Schultze-Berndt, 2006), depending on some factors such as the hours of recording, levels of tiers and analysis (Shimoji, 2026).. In pre-annotation stage, the type of data and analysis intended for the research determines the recording equipment and annotation software. Although there are a handful of annotation software for language data, this research focuses on the Eudico Language Annotator (ELAN) software because ELAN has the pre-requisite and compatibility criteria required for the research.

The challenge of annotating hours of recorded data within a stipulated time frame is the primary concern of this study. This research uses rule-based machine annotation framework in reducing the time spent in the annotation of Ibibio data on a three-tier analysis in ELAN.

2. Review of Related Literature

The close interface developed between databases and computer processing has produced many corpus processing techniques (Dash, 2005). These developed techniques over the years have quickened language data processing in linguistics and one of the widely known linguistic data processing technique is annotation. Annotation is a term that is used to cover all types of annotation (including transcription) that relates to recording of communicative events (Bird and Liberman, 2001)

The interface between linguistics and computer studies has quickened language data processing. The most widely known type of linguistics data processing technique is annotation. Bird and Liberman (2001) assert that annotation is used to cover all types of information (including transcription) that can be related to the recording of a communicative event for which no recording exists. Leech (2005) proposes that the art of annotation helps in manual examination and automatic analysis of language corpus.

Schultze-Berndt (2006:215) identifies three main levels of linguistics annotation. By “main level”, it implies the importance of the three levels of annotation to be included in any language annotation to enhance understanding; the three levels

of annotations are transcription, translation and grammatical annotation. Transcription is the standardized representation of human sounds in a language. Ochs (1979:44) observes that transcription in annotation is influenced and filtered by the annotator's decision, often times by his theoretical goals. Hence, the type of transcription used in an analysis is dependent on the goals of the research. However, transcription should be represented as faithful as possible, (Schultze-Berndt, 2006; Liang, & Levow, 2025).

Three types of transcription are widely identified in annotation literature; orthographic, phonetic, and phonemic. Orthographic transcription enhances accessibility of the documentation by members of the speech community (Schultze-Berndt, 2006:220). This type of transcription is used if there is an already established orthography of the documented language. Phonemic Transcription which shares similar features with orthographic transcription, concerns itself with the sound that brings about a difference in meaning. Although it is not necessary to include the phonetic transcription when a phonological analysis has already been done (Schultze-Berndt, 2006:224).

Translation has to do with interpreting the original data of the communicative event into a target language, possibly language of wider communication. Schultze-Berndt (2006) identifies two types of translation which are literal and free translation. Literal Translation is concern with translating the word or sentence into the target language exactly as it is represented in the source language. Free Translation translates freely to fit the whole meaning of a phrase or sentence.

Grammatical Annotation is a type of annotation that provides information on the grammatical aspect of the language. Interlinear glossing, grammatical tagging at the word and morpheme level, and grammatical notes are all part of grammatical annotation. Interlinear Glossing is a (phonemic or orthographic) representation combined with morpheme by morpheme glosses. It involves the addition of orthography and morpheme tiers. The interlinear glossing will be applied in this study during Ibibio children's play songs/games annotation.

An illustration of interlinear glossing is provided below (Ibibio example).

\orth	etək okon atigha ukod ke itiad
\mo	etək okon a-tigha ukod ke itiad
\it	small PN 1SG – kick leg on stone
\ft	Little Okon kicks his leg on a stone.

The first tier makes use of the orthography symbols, the second tier highlights the morphemes in the Ibibio language, the third tier provides definition for the morphemes, and the last tier provides the free translation.

3. The Ibibio Language

The Ibibio language belongs to the Ibibiod group of languages (Essien, 2008; Urua, 2007) under the Lower Cross languages of the New Benue Congo language family (Greenberg, 1963). The Ibibio language is used as a lingua franca in Akwa Ibom State, and also taught as a subject in primary, and secondary schools in the state. There are in-depth scholarly works in the Phonology and Grammar of the language, with recent focus on documenting different aspects of the language.

Ibibio children's play songs are songs recited during a game session. The songs add values to the game and carries the traditional norms of the Ibibio people. Many songs sang during the Ibibio children's play cut across various aspects of life such as marriage, child bearing, education and so on. Some of the data collected in the field indicate instances of code mixing Ibibio and English in the play songs. However, this study focuses primarily on songs performed using the Ibibio language only.

4. Methodology

The Ibibio children's play songs were recorded in nine randomly selected primary schools in nine Local Government Areas in Akwa Ibom State. The documentation took place in a school environment during physical education session or long break with primary four, five and six pupils as participants in the study. The pupils of these classes were selected for the study because they have a fair knowledge of the Ibibio language and could also understand the purpose of the documentation. The Sony Handy video camera HDR-CX 360 was used in recording the Ibibio children's play songs. The recorder captured the bodily movement of the participants and audio of the play songs.

5. Input and Output Processing

The corpus was systematically processed to produce the expected output. The corpus referred to here is the textual corpus. This process is essential to the automation process of annotation. A framework developed for the rule-based annotation process is provided below.

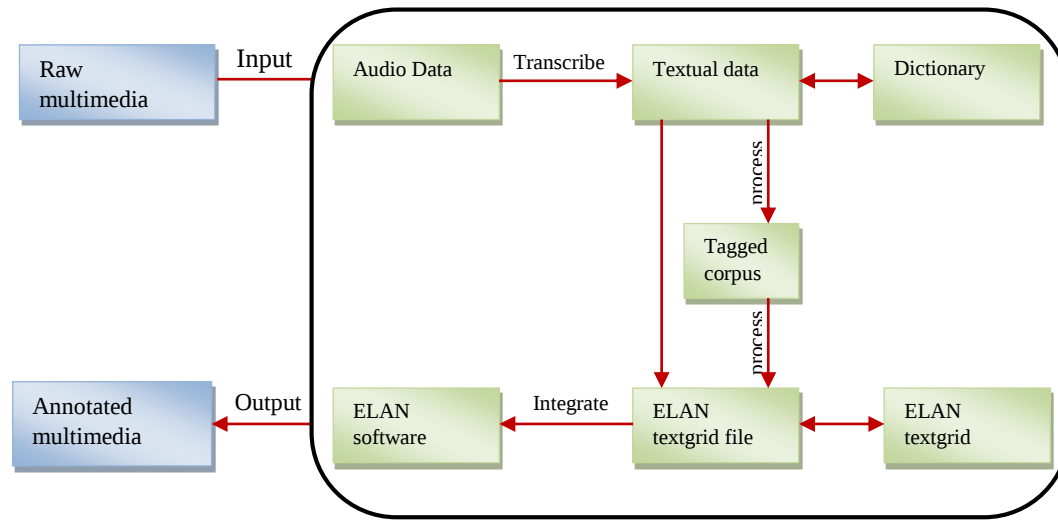


Figure 1. Framework for Rule-based Translation of Ibibio children's play songs

5.1. Text Grid File Preparation

Text grid file is a file that is required in ELAN software environment. Three (3) levels are required; the video stream, audio wave and the text. The text grid is displayed directly below the audio wave. The text grid deals with textual representations in different tiers (level). The processes involved in the preparation of the text grid file are discussed in the following Section.

5.1.1. Supervised Annotation

A wordlist was prepared from the song lyrics; prior to the supervised addition of the lexical category to the words in the wordlist to create an improvised dictionary. Morpheme boundaries were set in words that had more than one morpheme in the wordlist. The morpheme analysis is necessary at this level because it will be required during the interlinear translation tier for interlinear glossing. This was achieved using Microsoft Excel software. Three columns were created in Excel; word column, PoS column, and gloss column. Microsoft Excel was only used for creating a structured mini dictionary that was used in tagging the sentence in ELAN. A few constraints informed the decision-making process of tagging the data as discussed below;

Constraint 1

Homonyms – words with same pronunciation but different definition, constituted a major challenge to the development of the annotator. In addressing this challenge, numbers were assigned to each homonymic word in the improvised dictionary in addition to an IF clause introduced to the Python script to analyse, compare and extract the correct tag set for the homonym. The dictionary was primarily for the purpose of tagging the words in the recorded audio. It was about 250 entries.

Constraint 2

This involved the identification of words and bound morphemes. This process necessitates the appropriate translation. The following algorithm informed the decision of tagging free and bound morpheme.

```
if word is a free or bound morpheme:
    search part-of-speech/morpheme column
    locate the appropriate tag-set
else:
word is a free morpheme
```

5.1.2. Corpus file

This is a file that contains the lyrics of the Ibibio children's play songs/games. The corpus was transcribed from the recorded Ibibio children's play songs/games. This file was prepared using Microsoft Word 2007. This coding enables the insertion of special symbols that are not available on the computer keyboard imprinted characters.

6. Corpus Analysis

The Ibibio children's play songs/games were analysed using software tools selected to fit the framework of the analysis. Corpus analysis is necessary for proper interpretation of the corpus. The specific purpose of the software tools employed in this work is discussed in the next session.

6.1. Choice of Software

Five primary software tools were utilised in analysing the Ibibio children's play songs/games obtained through fieldwork. They include Microsoft Word and Excel software, Avidemux, Python and, ELAN. Avidemux is a video editing software that also has provision for extracting audio signal. In this work, the Avidemux 2.5 software was used to extract the audio wav needed for analysis in ELAN. To extract the wave file from the audio of the video file, the video file was imported to the Avidemux workspace by opening the file through the file 'drop-down' option on the menu bar. The audio 'drop-down' option by the left side of the workspace was changed to PCM (pulse-code modulation). PCM is a standard form of digital audio in computers. The audio drop-down on the menu bar was clicked and the fourth option 'save' was selected. The wave name of each of the song file in addition to '.wav' was keyed in before hitting the save button.

Python is an open source programming software tool that is employed in processing natural languages. Python 2.7 has provision for analysing utterances in wave form. This programming language was employed in the rule-based machine translation phase of the annotation process.

In addressing the multimedia data, the ELAN (Eudico Language Annotator) software was employed. ELAN is used in creating video and audio annotation of two or more people. For the documentation of Ibibio children's play songs, ELAN 4.6.0 was used for the annotation phase. ELAN is a professional annotation tool that has provision for time aligned annotation. In addition, the textual content of annotations in ELAN is stored using Unicode and the transcription stored in an XML (Extended Markup Language) format.

6.2. Word-Wave Alignment

To study the alignment format applied by ELAN, the sound file was played repeatedly in ELAN. Each word in the text grid was manually aligned to its specific audio wave stamp. The sample file was then used as a template for aligning other files. Where songs repeated more than once, the first occurrence of the song is aligned.

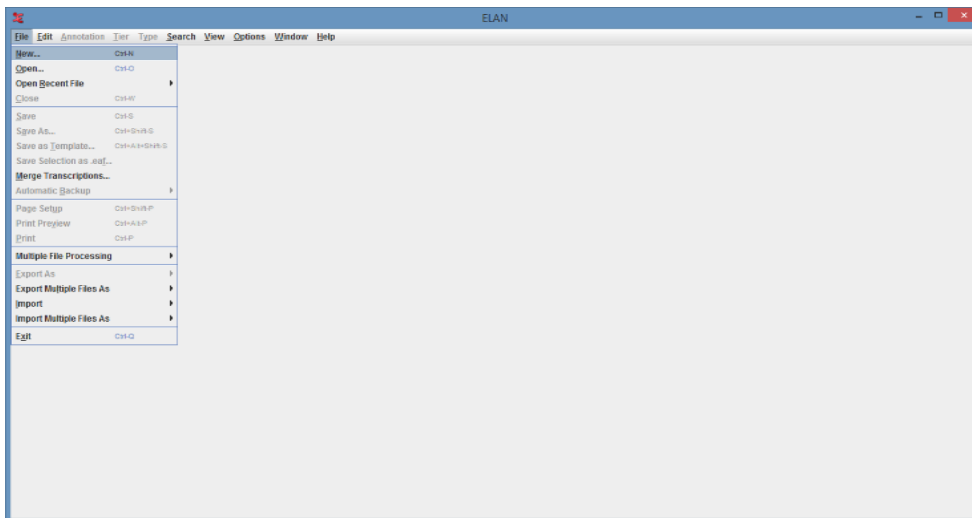
6.3. Integration of the generated NLP output into ELAN

An ELAN template was manually modified for the integration of the generated NLP output. The ELAN text grid file has three important requirements regarding audio/video recordings. These include;

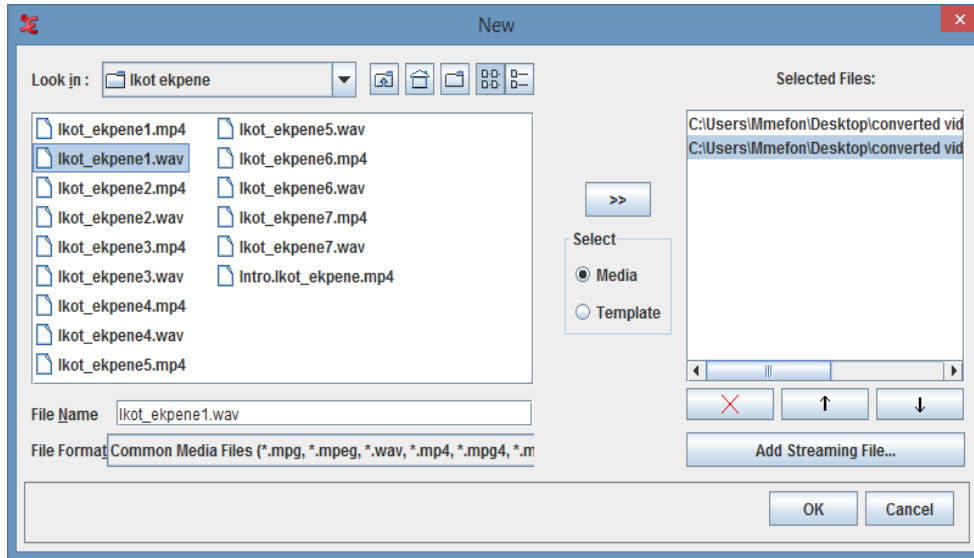
- i. Metadata of the recordings that indicate the encoding type, date, time stamps, media description and Uniform Resource Locator (URL) – a path that specifies the file’s location.
- ii. Catalogue of time alignments that indicates the various time boundaries in milliseconds (ms). The information assists in alignment of words with audio frames.
- iii. The Extensible Markup Language (XML) provides a structure for integrating the various tiers which confine them within the defined text boundaries to prevent text extending from the preceding and next tiers.

The following steps describe the creation of NLP output template into ELAN

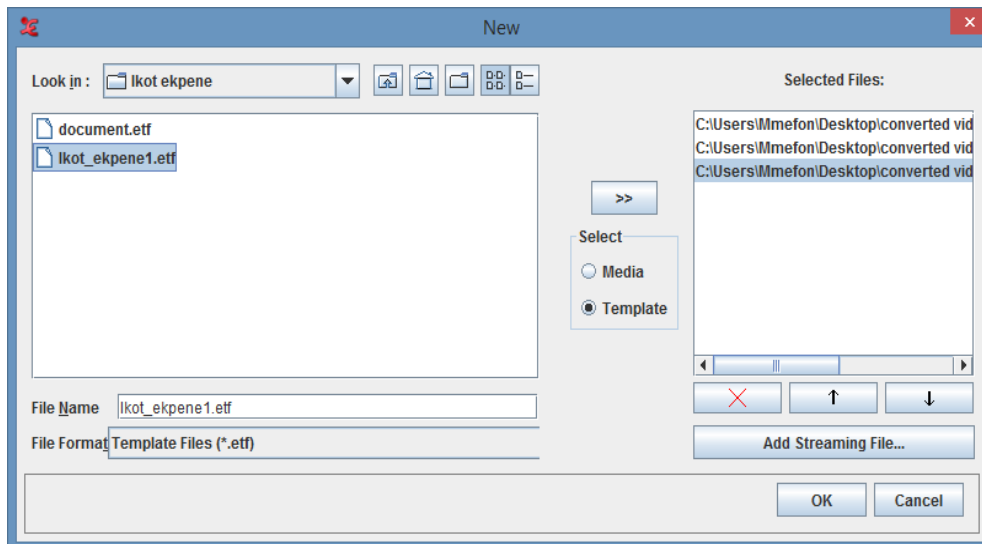
- i. ELAN software loaded to display ELAN’s workspace.



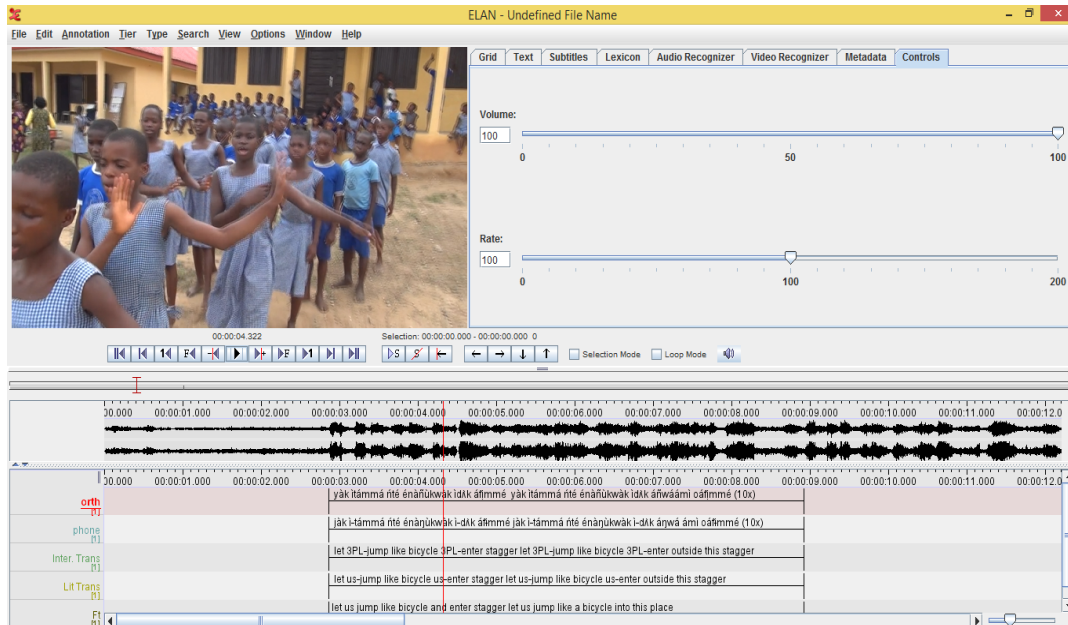
- ii. A new file was created by selecting the “New” option from the file dropdown menu.



- iii. The preceding action displayed a dialogue box with two options (media and template) for file creation. The media option was selected and it displayed the media files in the media folder selected and the respective media files (audio and video) were selected and loaded.



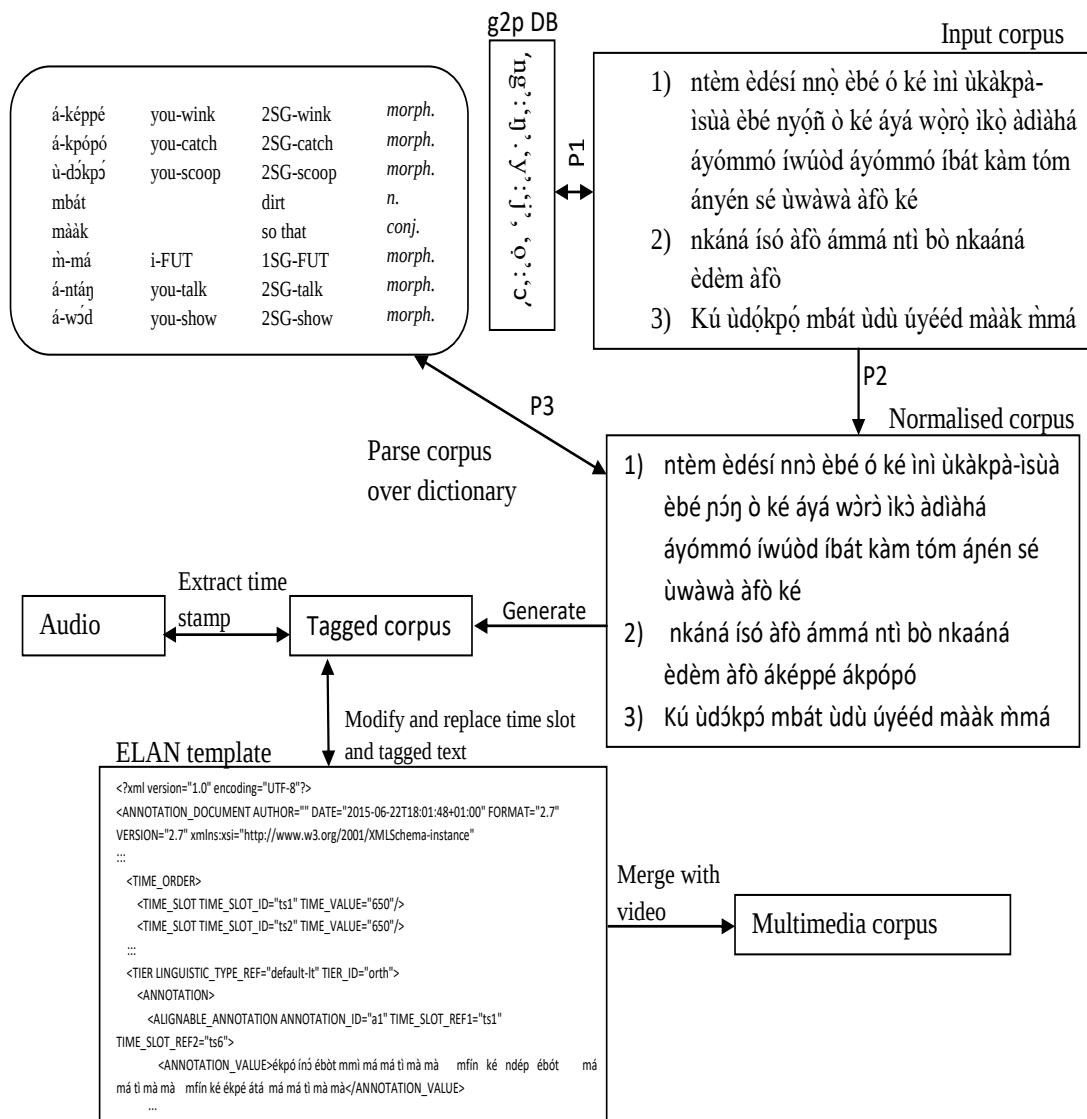
- iv. The template files were loaded by selecting the template radio button to reveal the template files.



- v. The ELAN workspace was loaded to reveal the video, audio, and template annotated file.

6.4. System Algorithm Flow

The system algorithm flow involved in the rule-based annotation of the Ibibio corpus with inputs and output including the integration process is shown in Figure 2.



7. Conclusion

To build an un-biased corpus that is comprehensive, a fieldwork covering the three senatorial districts of Akwa Ibom State was embarked upon. This process enabled the accumulation and documentation of the Ibibio children play songs – a very important heritage that requires urgent documentation to preserve its existence. The research

provides steps towards rule-based translation of the Ibibio language using Python programming language to automate the process. The rule-based translation adopted in the annotation of audio files in this study significantly reduced the time that would have been used in tagging the audio files. The study has provided an open-source remedy to rule-based annotation of data in ELAN using the Python programming language. The study can be applied to other Ibibio and tonal languages.

References

- Bird, S. and Liberman, M. (2001). A Formal Framework for Linguistic Annotation. *Speech Communication* 33:23-60.
- Dash, N. (2005). *Corpus Linguistics and Language Technology: with Reference to Indian Languages*. New Delhi: Mittal, p. 35.
- Essien, O. (2008). *A Grammar of the Ibibio Language*. Ibadan: University Press, p. 171.
- Fort, A., & Chelliah, S. L. (2026). *Customizing ASR for language documentation and resource prioritization*. In Proceedings of the Second Workshop on Customizable NLP: Progress and Challenges in Customizing NLP for a Domain, Application, Group, or Individual (CustomNLP4U) (pp. 149–159). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2026.customnlp4u-1.13>. <https://aclanthology.org/2026.customnlp4u-1.13.pdf> (aclanthology.org)
- Greenberg, J. (1963). *The Languages of Africa*. The Hague: Mouton and Co, p.171.
- Himmelmann, N. (1998). Documentary and Descriptive. *Linguistics* 36: 161-196.
- Himmelmann, N. (2006). Data and Language Documentation. In: Gippert, J., Himmelmann, N. P. and Mosel, U. (Eds.), *Essentials of Language Documentation*. Berlin: Mouton de Gruyter, pp. 1 - 30.
- Leech, G. (2005). Adding Linguistic Annotation In: Wynne, M. (Eds.), *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxford, pp. 17-29.

- Moeller, S. (2026). Computational methods for language documentation and revitalization. *Annual Review of Linguistics*, 12.
<https://doi.org/10.1146/annurev-linguistics-031522-101514> ([Annual Reviews](#))
- Ochs, E. (1979). Transcription as Theory. In: Ochs, E and Schieffelin, B. (Eds.). *Developmental Pragmatics*. New York: NY Academic Press, pp. 43-72
- Schultze-Berndt, E. (2006). Data and Language Documentation. In: Gippert, J., Himmelmann, N., and Mosel, U. (Eds.). *Essentials of Language Documentation*. Berlin: Morton de Gruyter, pp. 213-251.
- Seifart, F., Evans, N., Hammarström, H., & Levinson, S. C. (2018). Language documentation twenty-five years on. *Language*, 94(4), e324–e345.
- Shimoji, M. (2026). A POS tier is the key to automated annotation for low-resource language documentation: Neural interlinear glossing for Irabu, a Southern Ryukyuan language. *arXiv*. <https://arxiv.org/abs/2607.13372> ([arXiv](#))
- Urua, E. (2007). *Ibibio Phonetics and Phonology*. Portharcourt: Emhai Press, p. 218.
- Woodbury, A. (2003). Defining Documentary Linguistics. In: Austin (Ed.), *Language Documentation and Description*, pp. 35-51.

Aniefon Akpan¹, Moses Ekpenyong², Eno-abasi Urua², Margaret Okon³

¹National Institute of Nigerian Languages, Aba,

²University of Uyo, Uyo

³University of Calabar, Calabar