Malaysian Journal of Computing, 6 (1): 727-744, 2021

eISSN: 2600-8238 online

A HINDU-ARABIC TO *HAUSA* NUMBER TRANSCRIPTION SYSTEM

Safiriyu Ijiyemi Eludiora[1] and Muhammad Auwal Abubakar[2*]

[1]*Department of Computer Science and Engineering, Obafemi Awolowo University, Nigeria*
[2]*Department of Computer Engineering Technology, Federal Polytechnic Ede, Nigeria*
[1]eludioraomolola@gmail.com, [2]*mauwal07@gmail.com

ABSTRACT

The invention of the numeration system is regarded as one of man's great accomplishments. It greatly helps man express his communication needs and serves as an important tool in language pedagogy, historical linguistics, comparative study of African languages, and computational linguistics. However, the numeral system is reported to be an endangered area being identified in the use and study of language, and in no distant time, the traditional number system of the African indigenous counting systems may lose its contact with the new generation. This paper presents a Hindu-Arabic to Hausa number transcription system. Secondary data used was sourced from literature. Context-Free Grammar (CFG) and Unified Modelling Language (UML) was used to design the system. The system designed was implemented using the Python programming language. Mean Opinion Score (MOS) evaluation approach was used to evaluate the implemented system. The result of the evaluation on Numbers with Single Representations (NSR), and Numbers with Multiple Representations (NMR) is based on three (3) metrics: syllable accuracy, orthography accuracy and syntax accuracy. The experimental respondents', system developed, and human expert average scores on NSR were respectively 0%, 100% and 100% for syllable accuracy, 40.1%, 100% and 100% for orthography accuracy, and 62.8%, 100% and 100% for syntax accuracy. Similarly, the experimental respondents', system developed, and human expert average scores on NMR were respectively 0%, 100% and 100% for syllable accuracy, 21.4%, 100% and 100% for orthography accuracy, and 31.7%, 100% and 100% for syntax accuracy. The system is capable of transcribing cardinal numbers, 1 to 1-billion, and the expert response confirmed its accuracy. The study concluded that among others, the developed system is of great importance in the teaching and learning of the traditional Hausa counting system. Future work on contextual Hausa numeral system analysis is recommended.

Keywords: *Hausa Number, Hindu-Arabic, Transcription System*

Received for review: 20-12-2020; Accepted: 09-02-2021; Published: 02-03-2021

1. Introduction

A number is a word or symbol representing an amount or quantity (Longman, 2009). A numeral system consists of a set of symbols used for recording numbers and number names (Hess, 1970). Today, numbers are seen as tools that are absolutely necessary for our day-to-day activities (Agbeyangi *et al.*, 2016), as they are used in trade, cosmology, mathematics, divination, music and medicine (Akinade & Odejobi, 2014). The invention of a numeration system is regarded as one of the great accomplishments of man, and the Hindu-Arabic, simply Arabic (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) is one of the most widely used numerals (Hess, 1970). Numeral system greatly assists man in expressing his communication needs (Agbeyangi *et al.*, 2016; Babarinde, 2013), and also serve as an important tool in language pedagogy, historical linguistics, comparative study of African languages and computational linguistics (Babarinde, 2013). However, according to Babarinde (2013), numeral system is an endangered area being identified in language use and study. Affirming Babarinde (2013) statement, Mbah *et al.* (2014) reported that in no distant time, the traditional number system of the African indigenous counting system might lose its contact with the new generation, due to the fast influence of the western decimal numeral system. It is now common that many Africans cannot count from 1to 100 (not even incorrectly, let alone thousands or millions) in their native languages, which is very unfortunate and disgraceful. Furthermore, report of Akinade & Odejobi (2014) also noted number conversion into textual form as an important task in Text-to-Speech (TTS), Machine Translation (MT), and related applications but understudied by researchers. The aforementioned issues necessitated an urgent response to rescue endangered *Hausa* numerals specifically its traditional cardinal numbers from getting into extinction.

Hausa is a tone language with three distinctive tones; High (H), Low (L), and Falling tone (F) and vowel length (single and double), and a major indigenous language in Nigeria (Adelani *et al.*, 2020; Newman & Newman, 2001). The language is the most important lingua franca (indigenous) in West and Central Africa, and spoken by (40 – 50) million people as a first or second language (Wolff, 2013). Nigeria (Northern) and Niger (South-Eastern) have the largest population of the *Hausa* people (Bashir *et* al., 2017; Inuwa-Dutse, 2021; Shuaibu, 2014; Waziri *et al*., 2010). Popular World radio stations like Radio France (RF) Internationale, British Broadcasting Corporation (BBC), Voice of America (VOA), and Deutsche Welle (DW) also broadcast in *Hausa* language (Eludiora *et al*., 2016).

The *H*ausas have two numeral systems (Amfani, 2016): traditional and the present day. The *Hausa* traditional numeral system is influenced by both English (Amfani, 2016) and Arabic language (Amfani, 2016; Zaslavsky, 1973), due to their contacts with these languages. Such that *Hausa* borrowed several words from the two languages and hausanized them. E.g., 'Million' is *mìlíyàn* and *thalaathuun* (thirty in Arabic) is *tàlàatín*. Table 1 shows some cardinal Arabic numerals with their *Hausa* numbers equivalent.

The rest of this paper is organised as follows: Section 2 presents works related to the study, and Section 3 describes the methodology for the study. While the results of the study are presented and discussed in Section 4. Lastly, Section 5 concludes the study.

2. Literature Review

Akinade & Odejobi (2014) proposed a computational model of *Yorùbá* numerals in a number-to-text conversion system. The system is capable of converting cardinal numbers to their equivalent standard *Yorùbá* number names. Context-Free Grammar (CFG) and Unified Modelling Language (UML) were used in designing the system. The system was implemented using Python programming language, and evaluated using mean opinion score (MOS). A recall of 100% with respect to the collected corpus was produced by the system. The system is limited to *Yorùbá* language. The proposed work adopted some methods used in this work.

Ninan *et al.* (2017) proposed a computational analysis of *Igbo* numerals in a number-to-text conversion system. The system is capable of converting cardinal numbers (from 1-1000) to their equivalent standard *Igbo* number names. A Context-Free Grammar (CFG), sequence and activity diagrams were used in designing the system. System implementation was done using python programming language. The system evaluation was done by the administration of questionnaires to some chosen speakers of *Igbo* language (native) and experts to give a preferred representation of some selected numbers. The system result showed 100% accuracy for some randomly selected numbers. The authors further stated that the system would be useful in learning and teaching of *Igbo* language and thus fitted to other languages. This system is limited to *Igbo* language only (with scope 1 - 1000). The proposed work adopted some of the methods used in this work.

Eludiora & Oluwa (2017) proposed a *Yorùbá* arithmetic multimedia learning system. The system is capable of teaching speakers of *Yorùbá* language and other learners how to count, perform basic calculations (addition, subtraction, multiplication and division) on integer numbers (ranging from 0 - 1000) and also pronounce them in *Yorùbá* language. The system design was done using UML, and implemented using Python programming language. System evaluation was done by administrating questionnaires to potential users to determine how the system is perceived. The result showed that 90% of the users could not evaluate the addition expressions, subtraction expressions (55%), multiplication expressions (65%) and lastly division expressions (90%). The work is limited to *Yorùbá* language. The proposed work adopted some of the methods used in work.

Agbeyangi *et al.* (2016) developed a Web-Based *Yorùbá* numeral translation system. The system is an online application (web) capable of translating cardinal numbers both text and figure to standard *Yorùbá* equivalent form. Unified Modelling Language (UML) and Automata theory were used to design the system. System implementation was done using Google Web App Engine with python support. Mean Opinion Score (MOS) method was used in the system evaluation, where the authors considered each number (higher numerals included) to be either correctly written or not as criteria. The evaluation result gives 100% recall on corpus considered. Though the system is limited to *Yorùbá* language, the proposed work adopted some methods used in work.

Table 1. Cardinal Arabic numerals with their *Hausa* numbers equivalent

Hindu-Arabic	Traditional (Hausa)	Present (Hausa)
1.	Ɗáyá	Ɗáyá
2.	Bíyú	Bíyú
3.	Úkù	Úkù
4.	Húɗű	Húɗű
5.	Bìyár	Bìyár
6.	Shídà	Shídà
7.	Bákwài	Bákwài
8.	Tákwàs	Tákwàs
9.	Tárà	Tárà
10.	Goómà	Goómà
20.	Goómà bíyú	Ashìrín
25.	Goómà bíyú dà bìyár	Ashirin dà bìyár
50.	Goómà bìyár	Hàmsín
100.	Ɗàrií	Ɗàrií

Okafor *et al*. (2014) proposed an Indigenous language translator application for major ethnic groups in Nigeria (NigLT Ver 1.0). The application converts English words to *Hausa*, *Igbo* and *Yorùb*á, and vice versa. UML was used to design the software. System implementation was done using the JAVA programming language. The application does not support numerals. However, the proposed work gained some insight into the method used in work.

Eludiora & Nwokocha (2017) worked on the development of an interactive *Hausa* language learning system. The system can be able to teach *Hausa* language basics (alphabet and some numerals) to beginners, foreigners and others having an interest in the language. Unified Modeling Language (UML) was used in the system design. System implementation was done using Python programming language, and SQLite. The numerals are limited to 1-100. The authors also stated that the likes of this study would assist in preserving and preventing a language from the possibility of being into extinction. The proposed work gained some insight into *Hausa* numeral and also adopted some methods.

Newman (2007) proposed a *Hausa*-English dictionary. The author primarily presented the work in the standard *Hausa* orthography, and words definition are majorly in American English. *Hausa* long vowels and tones are being indicated. Long vowel by a macron, high tone unmarked, low tone by a grave accent, and falling tone by a circumflex accent. The definition of words is generally given in the order of most commonly used to less. The work is relevant to the proposed work in the data aspect.

Amfani (2016) proposed an outline of the *Hausa* numeral system. The works discuss the basic numbers and the formation of the derived ones and the influence of both English and

Arabic languages on the *Hausa* numeral system. The study grouped *Hausa* numeral into traditional and present-day. *Hausa* long vowels and tones are being indicated. Long vowel by double letters of the vowel, high tone by acute accent, low tone by a grave accent, and falling tone by double letters of a vowel in a long vowel, with the first having a high tone, and low tone for the other. This work serves as the major source of data for the proposed study.

Eludiora & Odetjobi (2016) developed an English to *Yorùb*á Translator. The system translates simple sentences (Subject-Verb-Object (SVO)). Phrase structure grammar, rewrite rules, and Unified Modelling Language (UML) was used in designing the system. The authors used three criteria in the evaluation process: sentence syntax correctness, word orthography, and word syllable accuracy. In syntax correctness, words positions are considered, and a properly positioned word (in a sentence) has a good score. Syntax correctness example is seen in the wrongly written sentence, *Musa ilé wọ,* as opposed to the correct one (*Musa wọ ilé*). Thus, in the former sentence, only the word, *Musa* will have a good score, as it is rightly positioned. In word orthography, the spellings of each word within a sentence are considered. Word syllable accuracy is about correctly tone-marking and under-dotting a word in a sentence. The system result shows 66.7%, 82.3% and 100% average accuracies respectively for the total experimental subject respondents (ESRs), machine translator and human expert for word syllable, word orthography, and sentence syntax. The work is limited to *Yorùb*á language. Some of the methods and evaluation process gave insight into the proposed work.

The reviewed research includes numerals transcription systems, language translators, multimedia/interactive learning systems, numeral system description, and a dictionary. The numerals transcription systems are based on *Yoruba,* and *Igbo* languages. One of the language translators is based on *Yoruba* language and the other one on the three indigenous languages (*Hausa, Yoruba, and Igbo*) in Nigeria. The multimedia/interactive learning systems are based on *Yoruba* and *Hausa* languages. While the numeral system description is based on *Hausa* present-day and traditional numeral system, the dictionary is a *Hausa*-English dictionary.

3. Methodology

3.1 Data Collection

Data used in this work were collected from Amfani (2016), Bunza (2018) and Newman (2007). However, Amfani (2016) presentation is adopted due to its simplicity and usage in a formal medium of communication. It is analysed and explained in Table 2.

3.2 System Design

The system design was done using Context-free grammar (CFG) to capture the structure of the *Hausa* numerals and Unified Modelling Language (UML) for the software design. The system flow chart is, as shown in Figure 1. The first stage is where a user inputs the Arabic number (e.g., 5, 10, etc.) to be converted into its *Hausa* textual equivalent. Thereafter, the system checks to see if the number is a valid one or not. For a valid number, the system proceeds to check if it is basic or derived. Otherwise, it goes back to the initial stage and continues to wait for a valid number. If found to be a basic number, the system converts it directly to its *Hausa* textual equivalent with the help of a look-up table. Otherwise, the system first decomposes the number and convert all the lexical items (which includes the arithmetic operators involved) to their *Hausa* equivalents using the rules in Table 2. Finally, the converted lexical items are combined still using the rules in Table 2 to have the *Hausa* textual equivalent of the Arabic number. The system transcribes Hindu-Arabic cardinal numbers from 1 to 1-billion to their *Hausa* equivalents.

Table 2. *Hausa* Numeral Derivation Rule/Technique

Numbers	Rule/Technique
1-10	These are basic numbers that cannot be broken or decomposed further. They are: *ɗáyá* (1), *bíyú* (2), *úkù* (3), *húɗú* (4), *bìyár* (5), *shídà* (6), *bákwài* (7), *tákwàs* (8), *tárà* (9), *goómà* (10)
11 – 19	Derived by the addition of 1 - 9 to 10. The *Hausa* word *sháà* is used for the addition operator. Examples: 11 = 10 + 1 = *goómà sháà ɗáyá*, 19 = 10 + 9 = *goómà sháà tárà*
Decades (20 -90)	Derived by multiplying 10 by 2 - 9, and multiplication is implied. Examples: 20 = 10 x 2 = *goómà bíyú*, 50 = 10 x 5 = *goómà bìyár*.
21 - 99 (Decades excluded)	Derived by the addition of 1 - 9 to Decades (20 - 90). The *Hausa* word "*dà*" is used for the addition operator as opposed to "*sháà*" used for 11 - 19 above. Examples: 21 = (10 x 2) + 1 = *goómà bíyú dà ɗáyá*, 99 = (10 x 9) + 9 = *goómà tárà dà tárà*
Between 10 - 100 (ending with 8 or 9). That is, 18, 19 … 28, 29… 98, 99.	Derived by subtracting "1" or "2" from their next "tenths"). The *Hausa* word *báà* is for the subtraction operator. Examples: 18 = 20 – 2 = goómà bíyú bàa bíyú, 59 = 60 – 1 = goómà shídà bàa ɗáyá.
100 – 900	Derived as multiple of 100 and employed the used of the *Hausa* word "*dàrií*" for hundred. Multiplication is implied. Examples: 100 = 100 x 1 = *dàrií ɗáyá*, 220 = 200 + 20 = *dàrií bíyú dà goómà bíyu*.
1000 and above	Derived as multiple of 1000 and employed the used of the *Hausa* word "*dúbuú*" for 1000. The *Hausa* word *sàu* is for the multiplication operator. Examples: 1000 = 1000 x 1 = *dúbuú ɗáyá*. 512,000 = (1000 x 500) + 12 = *dúbuú ɗàrií bìyár dà goómà sháà bíyú*. 1000,000 = 1000 x 1000 = *dúbuú sàu dúbuú*. *Zámbàr* is another *Hausa* word for 1000 but only used for one million upward. Example: 1,000,000 = 1000 x 1000 = *zámbàr dúbuú* **Note:** The multiplication operator between *zámbàr* and *dúbuú* is always implied.

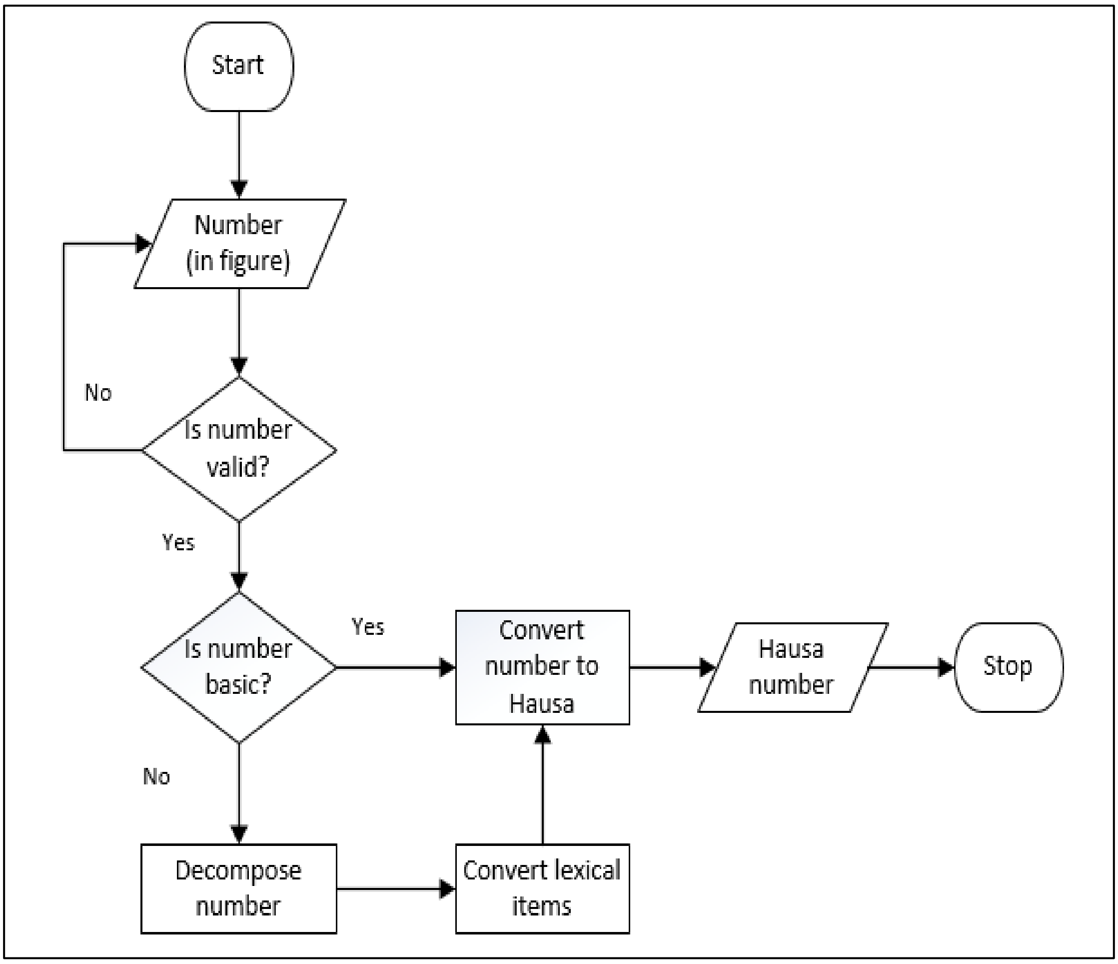


Figure 1. System flow chart.

3.2.1 Context-Free Grammar (CFG) for *Hausa* numeral system

The design of the CFG for *Hausa* numerals starts by identifying the set of the terminal:

(a) The set of lexemes:

i) DIGIT = {ɗáyá (1), bíyú (2), úkù (3), húɗú (4), bìyár (5), shídà (6), bákwài (7), tákwàs (8), tárà (9)}

ii) DIGIT1000 ={zámbàr (D1000)}

iii) The set of multiplicative bases i.e., M = {*goómà* (10), *ɗàrií* (100), *dúbuú* (1000)}

(b) The sets of lexical affixes indicating arithmetic operations in *Hausa* numerals. These sets are two:

i) The addition operator (+) used between tens and unit between the numbers 10 and 20. The *Hausa* word *sháà* is used as discussed in Table 2. A typical example is seen in *goómà sháà úkù* (13). Hence, V = {*sháà* (+)}.

ii) A set of operators used between thousand and hundred, hundred and tens, and tens and unit. This includes:

- Addition operator (+), represented by the *Hausa* word *dà*, used between thousand and hundred, hundred and tens, and tens and unit (other than the one mentioned in b (i) above). A typical example is seen in *dúbuú ɗáyá dà ɗàrií bìyár dà goómà bíyú dà úkù* (1,523)

- Subtraction operator (-), represented by the *Hausa* word *báà*, used between tens and unit, and only applicable to over-counting in *Hausa*. A typical example is seen in *goómà úkù báà* bíyú (28)
- Multiplication operator (x), represented by the *Hausa* word *sàu* and only used in representing millions and above. That is, it only appears between two or more *dúbuú*. A typical example is seen in *dúbuú sàu dúbuú úkù* (3,000,000).

Hence, VV = {*dà* (+), *báà* (-), *sàu* (x)}.

Thus, all elements in: DIGIT, DIGIT1000, M, V, and VV, formed the set of terminal symbols, T.

Hence, the *Hausa* numeral system production rules are as shown below:

S → NUM (1)

NUM → NP | NUM SN (2)

SN → VV NP (3)

NP → DIGIT | DIGIT1000 | MP | NP VP (4)

MP → M | MP NP | NP MP (5)

VP → V DIGIT (6)

Where NP = Noun phrase, VP = Verb phrase, SN = Sub number, NUM = Number form.

- Rule (1) means that the start symbol S, is a number denoted by NUM as shown in Figure 2.
- Rule (2) is interpreted as a number (NUM) could be formed by combining an existing number (NUM) with a phrase (SN) or from just a phrase (NP), as shown in Figure 2.
- Rule (3) is interpreted as the formation of the phrase (SN) from the combination of the verb (VV) and another phrase (NP), as shown in Figure 2.
- Rule (4) means that a phrase (NP) could be a DIGIT, DIGIT1000, phrase formed by multiplication (MP), or a combination of noun phrase (NP) and verb phrase (VP), as shown in Figure 3.
- Rule (5) indicates that a phrase formed by multiplication (MP) could be a base (multiplicative), recursive multiplication of MP and a noun phrase (NP) or vice versa, as shown in Figure 3.
- Rule (6) indicated that a verb phrase (VP) could be formed by combining a verb (V) and a DIGIT, as shown in Figure 4.

The proposed *Hausa* numeral grammar was tested using the python library, Natural Language Toolkit (NLTK). *Hausa* numerals input strings were parsed using a chart parser to see if they are in conformity with the designed grammar. Figure 2 to Figure 7 are some of the parse trees that resulted from the verification.

The rules below regulate well-formed *Hausa* (numeral) structures:

i. Anytime a phrase (MP) is formed by multiplying two numbers, the multiplier (NP) must be less than the multiplicand (MP). This is illustrated in Figure 5.
ii. Anytime the rule SN → VV NP is used, and VV has *sàu* value, then NP can only be a multiple of 1000 (*dúbuú*). It is as illustrated in Figure 6.
iii. Anytime the rule MP → NP MP is used, and NP is D1000 (*zámbàr*), then MP must be a multiple of *dúbuú* (1000) and that multiplication implied as depicted in Figure 7.

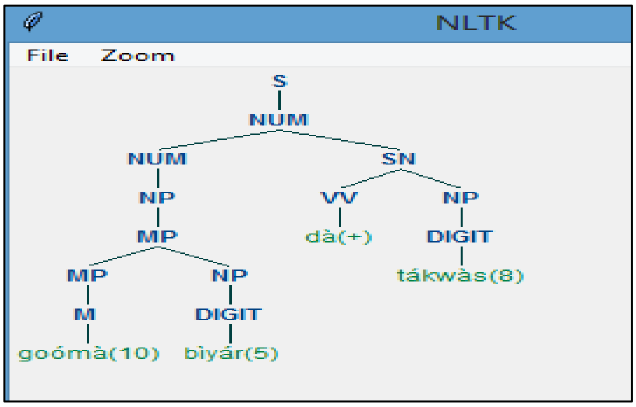


Figure 2. Parse tree for 58

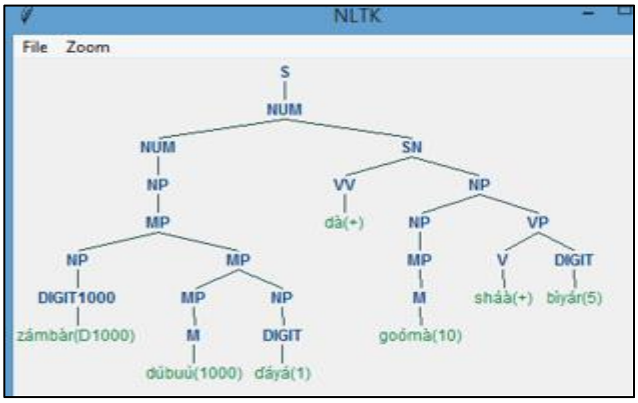


Figure 3. Parse tree for 1,000,015

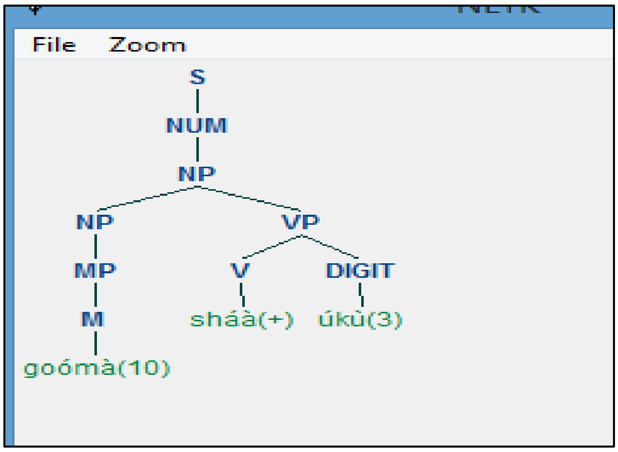


Figure 4. Parse tree for 13

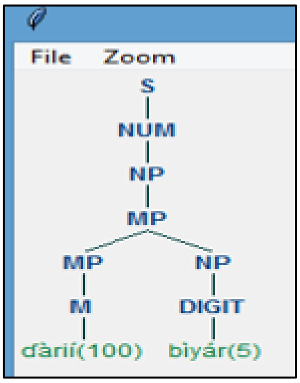


Figure 5. Parse tree for 500

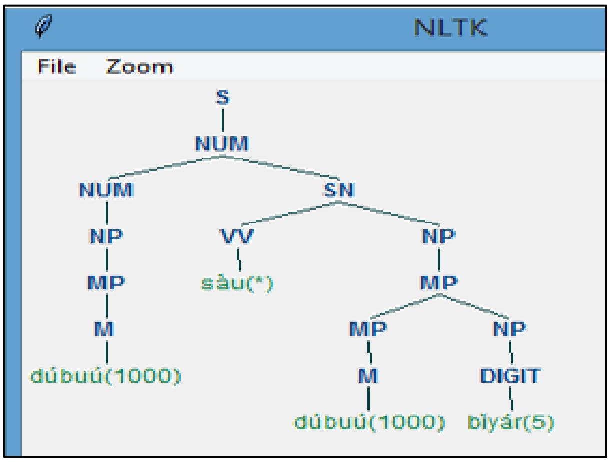


Figure 6. Parse tree for 5,000,000 using *dúbuú*

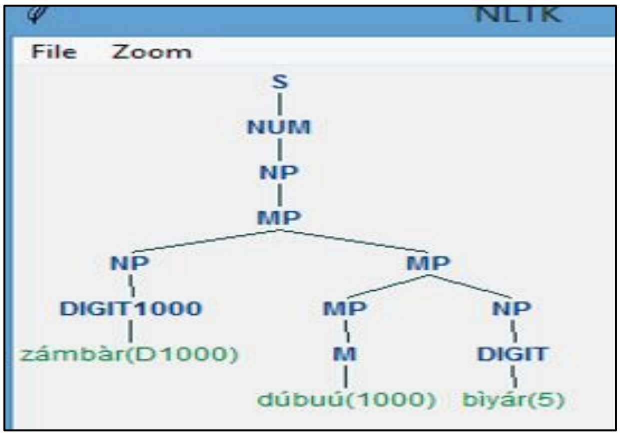


Figure 7. Parse tree for 5,000,000 using *zámbàr*

3.2.2 Use-Case diagram

Use-Case diagrams are used in capturing the (functional) requirement of a system, which includes actors (human beings interacting with the system) relationship to important system processes and also the relationships between (different) Use-Cases (Bell, 2003). The system Use-Case diagram is as shown in Figure 8. In this work, the system (Fig. 8) has an actor (the user interacting with the system) and three use-cases (load application, type an Arabic number and click convert, and view the converted number). Therefore, the system allows the user to load the software application, type an Arabic number and click convert for the system to transcribe the inputted number, and finally view the converted number (transcribed *Hausa* number).

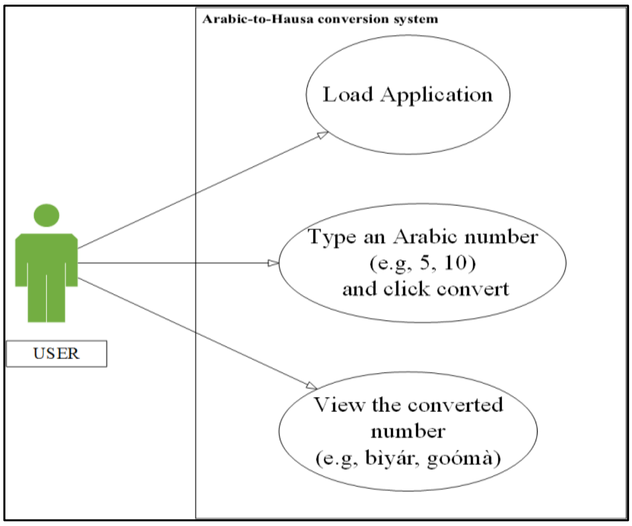


Figure 8. Use-Case diagram of the system

3.2.3 Class diagram

A class diagram depicts a system static structure (Bell, 2003). It serves as a kind of system specification and also states the kind of objects that can exist, the data being encapsulated by the object, and the relationship between system objects (Priestley, 2000). The developed system class diagram (Figure 9) is explained below.

- MainWindow: The graphical user interface (GUI) is generated using this class, and this same class inherits the PyQt5 class.
- CFGparser: This class implements the designed *Hausa* numeral grammar. The parse trees for *Hausa* numerals are also generated by this class. The nltk class is inherited by this same class.
- ConvertToNumber: The transcription task is performed by this class and implemented using the following methods:
    - Translate method: This method combines the different parts (unit, tens, etc.) of a number to gives the final result of the transcription process.
    - Unit method: transcribed the unit part of a number.
    - Tens method: transcribed the tens part of a number.
    - Hundred method: transcribed the hundredth part of a number.
    - Thousand method: transcribed the thousandth part of a number.
    - Million method: transcribed the millionth part of a number.

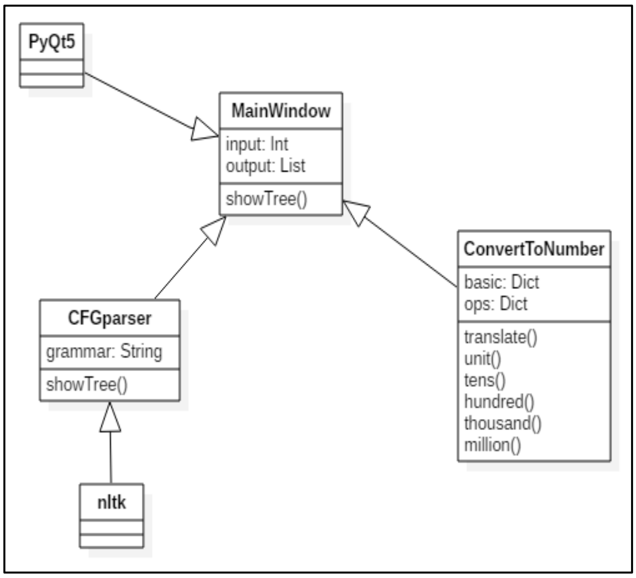
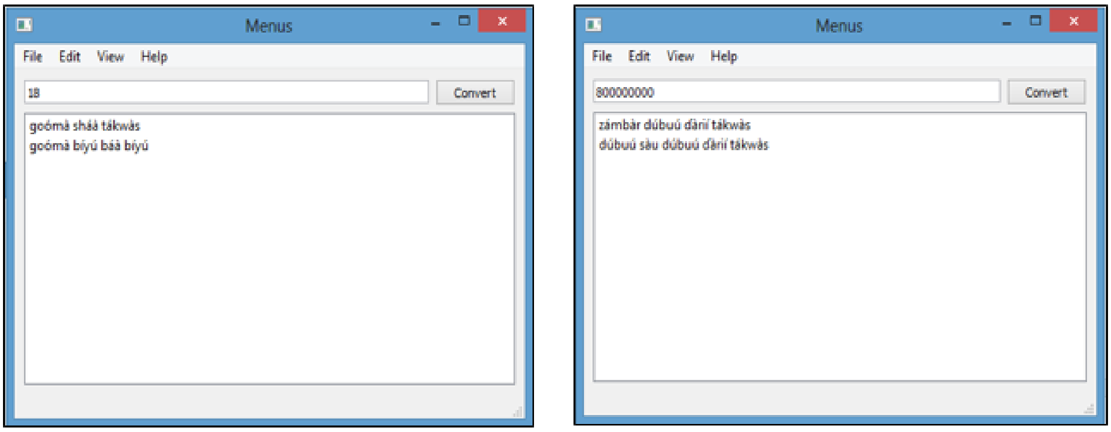


Figure 9. Class diagram of the system

3.3 System Implementation

The system is a window application (desktop) implemented using Python programming language and its supporting modules, the natural language toolkit (NLTK) and PyQt. While NLTK was used in the implementation of the designed *Hausa* numeral grammar as well as the generation of the numbers' parse trees, the PyQt is used in the graphical user interface (GUI) implementation. Figure 10 shows the developed system screenshots displaying outputs for the numbers 18 and 800 million.



a) 18 b) 800,000,000

Figure 10. System outputs

3.4 Evaluation

The evaluation involves questionnaire design and administration.

3.4.1 Questionnaire Design and Administration

The questionnaire was designed to have questions on two group of numbers: Numbers with Single Representation (NSR - 5, 12, 50, 75, 223, 500, 2000, 80000 and 500000), and Numbers with Multiple Representation (NMR - 18, 79, 458, 2000000 and 500000000). Both the NSR and NMR contain basic and higher numbers designed to test the experimental respondents' understanding of these numbers' representation in traditional *Hausa* counting system. The questionnaires were administered to fifteen (15) literate *Hausa* speakers (expert included). The expert is a Professor of linguistics, and the experimental respondents' are of educational level ranging from undergraduate to postgraduate with ages above twenty-five years. All the respondents' are from core Northern states of Nigeria, with a majority having the knowledge of *Hausa* orthography ranging from adequate to excellent.

3.4.2 System Evaluation

Syllable accuracy, orthography accuracy and syntax accuracy are the three criteria used in the system evaluation. The experimental respondents' conversion accuracy was compared with that of the developed system, and human expert (professional *Hausa* speaker). The use of appropriate tone marks on number name is what is meant by syllable accuracy. Each word in a number name is evaluated for its syllable accuracy and scored accordingly. A typical example is seen in the two-syllable *Hausa* number name *bìyár* (5), where *bì* is one-syllable with low tone on *i*, and *yár* is another syllable with a high tone on *a*. The use of correct spellings that include appropriate vowel length on number name is meant by Orthography accuracy. Each word in a number name is evaluated for its Orthography accuracy and scored accordingly. A typical example is illustrated in the correctly spelled number, *gooma* (10), along with its long vowel (*oo*) length clearly written, as opposed to the wrongly spelled one *goma*. The syntax accuracy was evaluated in accordance to the syntactic structure of the *Hausa* number names. The position of a word in a number name is considered, and a well-positioned word in a number name attracts a good score. A typical example is also seen in the number *ɗàrií bíyú* (200), which is (100 x 2), and not *bíyú ɗàrií* (2 x 100). So, the former would attract good score for each word, as they are well-positioned, and the latter would have zero for each word as they are not well-positioned.

4. Results and Discussion

The evaluation results for both NSR, and NMR based on syllable, orthography, and syntax accuracy are presented and discussed in this section.

4.1 Numbers with Single Representation (NSR)

The numbers with single representation are 5, 12, 50, 75, 223, 500, 2000, 80000 and 500000.

4.1.1 NSR Syllable Accuracy

The syllable accuracy result for the numbers with a single representation (NSR) is shown in Figure 11. The result shows that the respondents have an average score of 0%, while that of the system and the Expert are 100%. The respondents' performance is due to their inability to tone-marked any of these numbers (not even wrongly) as it was observed in their responses. This is perhaps due to lack of using tone marks in *Hausa* ordinary writings as opposed to scholarly works.

4.1.2 NSR Orthography Accuracy

The Orthography accuracy result for the numbers with a single representation (NSR) is shown in Figure 12. The result shows that the respondents have an average score of 40.1%, while that of the system and the Expert are 100%. The respondents' performance is due to their inability to write most parts of the number names in their traditional *Hausa* form as well as the inability to show the long vowel length where appropriate or necessary by all of them. It was also observed that most of the respondents used the present-day *Hausa* counting system in the representation of numbers in this category. Thus, the caused for the low scores on most of the numbers, especially 50 and 80000, as seen in the result presented in Figure 12. This affirmed the statement of both Amfani (2016) and Zaslavsky (1973) that the traditional *Hausa* counting system is influenced by both English and Arabic languages.

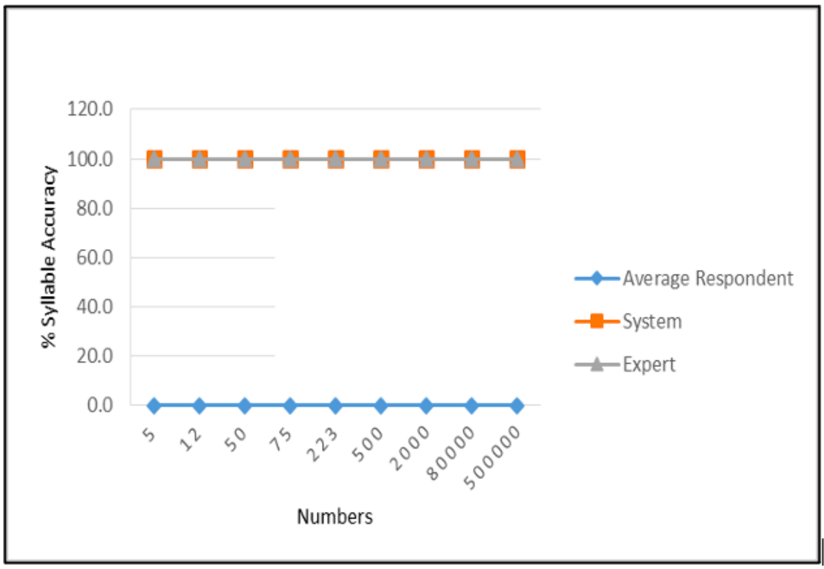


Figure 11. Syllable Accuracy

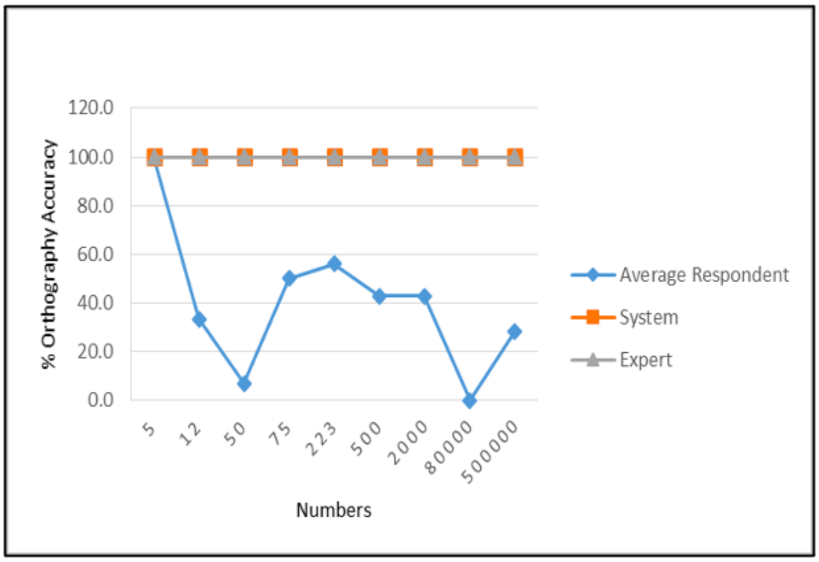


Figure 12. Orthography Accuracy

4.1.3 NSR Syntax Accuracy

The Syntax accuracy result for the numbers with a single representation (NSR) is shown in Figure 13. The result shows that the respondents have an average score of 62.8%, while that of the system and the Expert are 100%. It was observed that the respondents' have difficulties in writing some parts of the number names in their correct position due to their usage of the present-day *Hausa* counting in representing most of these numbers. This effect is seen much in the numbers 50, 75 and 80000, as shown in Figure 13.

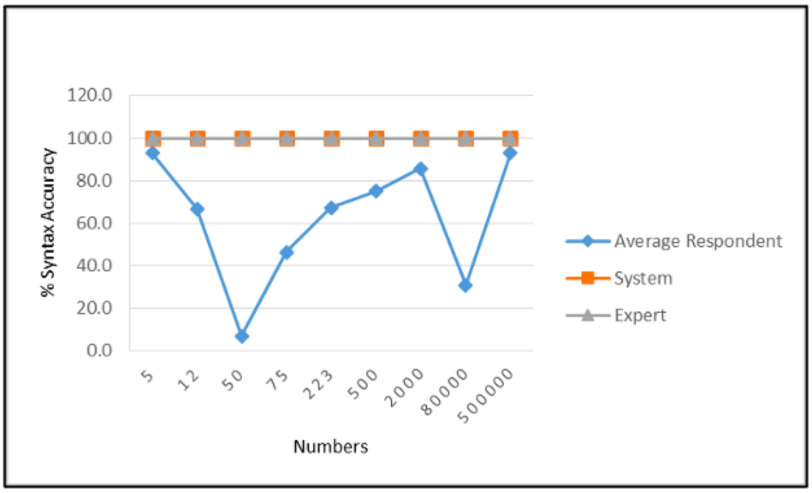


Figure 13. Syntax Accuracy

4.2 Numbers with Multiple Representation (NMR)

The numbers with multiple representations are 18, 79, 458, 2000000 and 500000000.

4.2.1 NMR Syllable Accuracy

The syllable accuracy result for the numbers with multiple representations (NMR) is shown in Figure 14. The result shows that the respondents have an average score of 0%, while that of the system and the Expert are 100%. The respondents' performance is due to their inability to tone-marked any of these numbers (not even wrongly), as it was observed in their responses. This is perhaps due to none used of tone marks in *Hausa* ordinary writings as opposed to scholarly works.

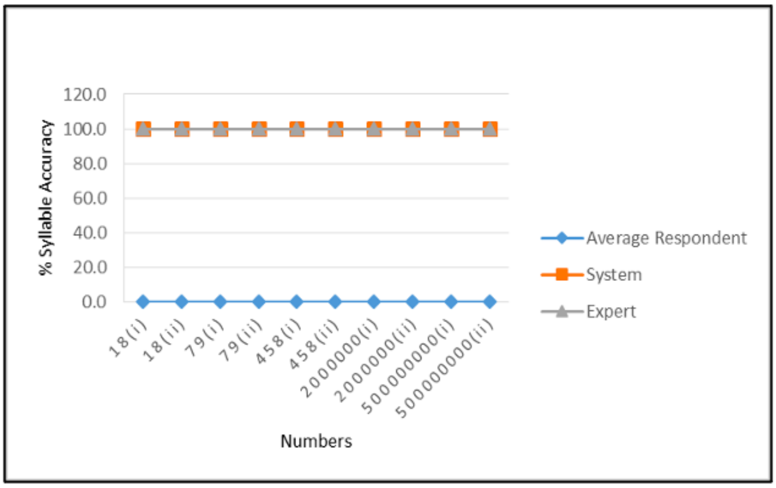


Figure 14. Syllable Accuracy

4.2.2 NMR Orthography Accuracy

The Orthography accuracy result for the numbers with multiple representations (NMR) is shown in Figure 15. The result shows that the respondents have an average score of 21.4%, while that of the system and the Expert are 100%. The respondents' performance is due to their inability to write most parts of the number names in their traditional *Hausa* form as well as the inability to show the long vowel length where appropriate or necessary by all of them. It was also observed that most of the respondents used the present-day *Hausa* counting system in the representation of these numbers, and could not represent the numbers 18, 79, and 458 using the subtraction method as indicated in Figure 15 [18(ii), 79(ii), and 458(ii)]. Also observed was the inability of the respondents to correctly represent the higher numerals 2000000 and 500000000 using the word *zámbàr* as also indicated in Figure 15 [2000000(ii) and 500000000(ii)]. This also affirmed the statement of both Amfani (2016) and Zaslavsky (1973) that the traditional *Hausa* counting system is influenced by both English and Arabic languages.

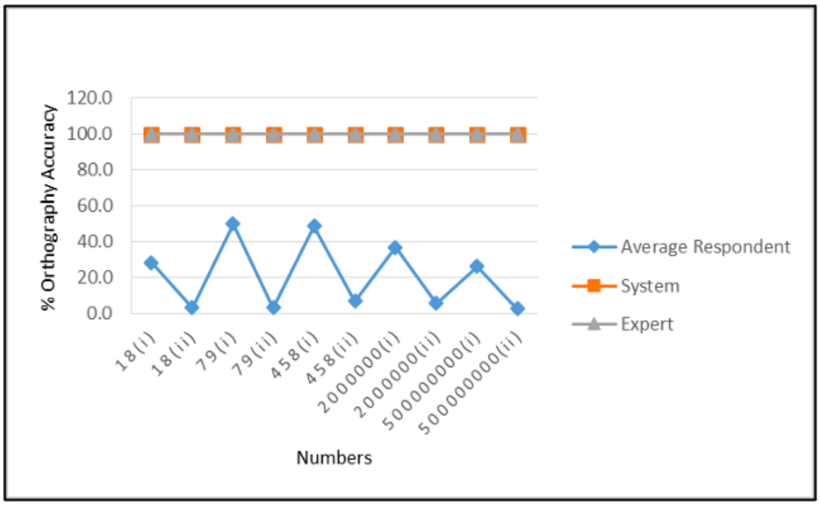


Figure 15. Orthography Accuracy

4.2.3 NMR Syntax Accuracy

The Syntax accuracy result for the numbers with multiple representations (NMR) is shown in Figure 16. The result shows that the respondents have an average score of 31.7%, while that of the system and the Expert are 100%. It was observed that the respondents' have difficulties in writing most parts of the number names in their correct position due to the users of the present-day *Hausa* counting in representing most of these numbers. Furthermore, is their inability to represent the numbers 18, 79, and 458 using the subtraction method as indicated in Figure 16 [18(ii), 79(ii), and 458(ii)] as well as the higher numerals 2000000 and 500000000 using the word *zámbàr* as also indicated in Figure 16 [2000000(ii) and 500000000(ii)]. The complete evaluation result for both NSR, and NMR is summarized in Table 3.

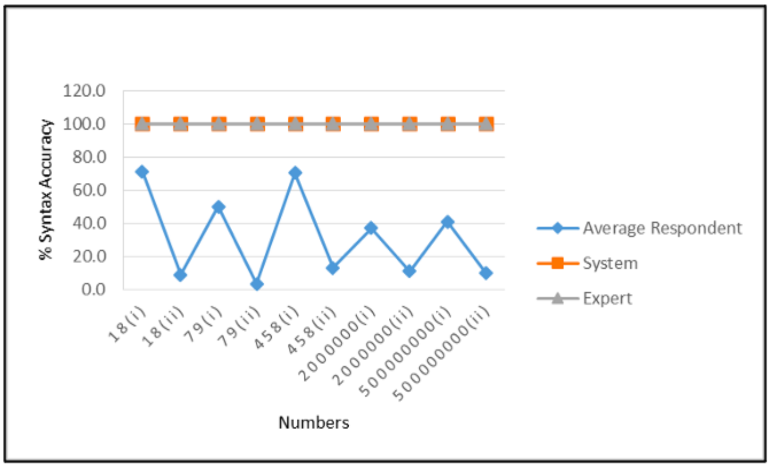


Figure 16. Syntax Accuracy

Table 3. Evaluation Result Summary

	Metrics (Accuracy)	**Respondents' (%)**	**System (%)**	**Expert (%)**
NSR	Syllable	0	100	100
	Orthography	40.1	100	100
	Syntax	62.8	100	100
NMR	Syllable	0	100	100
	Orthography	21.4	100	100
	Syntax	31.7	100	100

5. Conclusion and Future Work

This study presented the development of an Arabic-to-*Hausa* number conversion system. The system transcribes Hindu-Arabic cardinal numbers from 1 to 1-billion to their equivalent traditional *Hausa* number names. The study shows that *Hausa* language, specifically its traditional numeral system is likely to be a victim of language endangerment due to the influence of both Arabic and English languages. It is evident that the majority of the experimental respondents are not familiar and even never heard of many of the traditional *Hausa* numerals. Of more concerns is the fact that the respondents are not even aware of their ignorant knowledge of the traditional *Hausa* numerals. Thus, most of them claiming to have adequate to excellent knowledge of the traditional *Hausa* number system orthography, despite being learned (mostly at postgraduate level) and having age greater than 30 years. However, with the expert response, it is very clear that the (developed) system accuracy outperformed the other respondents' response. Therefore, this study's likes have to be encouraged to familiarize *Hausa* speakers of this generation the traditional *Hausa* numeral system. Else, our vital environmental complement, language, could be endangered, and finally overcome by attrition and death at the extreme as reported by Fabunmi & Salawu (2005). However, this system could promote the use of *Hausa* language, specifically its traditional counting system in speech processing and also help to increase its functional load. The developed system could also be very useful in effectively teaching and learning *Hausa* language, specifically its traditional numeral system (including its structure). It is hoped that future researchers would give necessary attention to these areas such as contextual *Hausa* numeral system analysis, and *Hausa* numeral textual-form conversion to numbers.

Acknowledgement

The authors appreciate the support of the Nigerian Petroleum Technology Development Fund (PTDF) scholarship board. Also appreciated is Prof. A. H. Amfani of the department of Nigerian Languages, Usmanu Dan-Fodio University, Sokoto for his unrelenting guidance.

References

Adelani, D. I., Hedderich, M. A., Zhu, D., Berg, E. V. D., & Klakow, D. (2020). Distant Supervision and Noisy Label Learning for Low Resource Named Entity Recognition: A Study on Hausa and Yoruba. arXiv preprint arXiv:2003.08370.

Agbeyangi, A. O., Eludiora, S. I., & Popoola, O. A. (2016). Web-Based Yorùbá Numeral Translation System. *IAES International Journal of Artificial Intelligence (IJ-AI), 5*(4), 127-134.

Akinade, O. O., & Odejobi, O. A. (2014). Computational modelling of Yoruba numerals in a number-to-textconversion system. *Journal of Language Modelling* , 2(1), 167-211.

Amfani, A. H. (2016). An Outline of the Hausa Numeral System. In Ozo-mekuri, & E. S. L., *The Numeral Systems of Nigerian Languages* (Chan ed., pp. 239-244). Port Harcourt, Nigeria: M and J Grand Orbit Communications Ltd.

Babarinde, O. (2013). The Linguistic Analysis of the Structure of the Yoruba Numerals. *Journal of Literature, Languages and Linguistics, 1*, 78-88.

Bashir, M., Rozaimee, A., & Isa, W. M. W. (2017). Automatic Hausa language text summarization based on feature extraction using Naive Bayes Model. World Applied Science Journal, 35(9), 2074-2080.

Bell, D. (2003). An Introduction to the Unified Modeling Language. Retrieved March 30, 2020, from *http://www.developer.ibm.com/technologies/web-development/articles/an-introduction-to-uml/*

Bunza, A. M. (2018, December 28). *Mathematical Heritage in Hausa Number System: (A Proposal for Teaching Mathematics using Nigerian Languages)* . Retrieved May 10, 2019 from Prof. Aliyu Muhammad Bunza - A Professor of African Cultures (Hausa): *https://www.alibunza.com/2018/12/mathematical-heritage-in-hausa-number.html*

Eludiora, S. I., Afolabi, B. O., & Sanusi, A. A. (2016). Development of English to Hausa machine Translation System. *Research Journal of Engineering and Technology,* 2(1), 12-32.

Eludiora, S., & Nwokocha, R. (2017). Development of an Interactive Hausa Language Learning System. *International Journal of Computer Applications (0975 – 8887),* 176(5), 13-22.

Eludiora, S., & Odejobi, O. (2016). Development of an English to Yoruba Machine Translator. I.J. Modern Education and Computer Science, 11, 8-19. doi:10.5815/ijmecs.2016.11.02

Eludiora, S., & Oluwa, H. (2017). Development of a Yorùbá Arithmetic Multimedia Learning System. *Universal Journal of Educational Research, 5*(5), 862-873. doi:10.13189/ujer.2017.050518

Fabunmi, F. A., & Salawu, A. S. (2005). Is Yorùbá an Endangered Language? *Nordic Journal of African Studies*, 14(3), 391–408.

Hess, A. L. (1970). A Critical Review of the Hindu-Arabic Numeration System. *The Arithmetic Teacher,* 17(6), 493-497.

Inuwa-Dutse, I. (2021). The First Large Scale Collection of Diverse Hausa Language Datasets. arXiv preprint arXiv:2102.06991.

Longman (2009). Longman Dictionary of American English. England: Pearson Education Limited.

Mbah, E. E., Ndubuisi, O. A., & Monday, A. (2014). A Comparative Analysis of the Numerals and Counting Systems of Igbo and Igala. Research on Humanities and Social Sciences, 4(26), 61-65.

Newman, P. (2007). *A Hausa-English Dictionary.* United State of America: Yale University Press.

Newman, R. M., & Newman, P. (2001). The Hausa Lexicographic Tradition. *Lexikos 11 (AFRILEX-reeks/series), 11*, 263-286.

Ninan, O., Iyanda, A., Elesemoyo, I., & Obasa, E. (2017). Computational Analysis of Igbo Numerals in a Number-to-text Conversion System. *Journal of Computer and Education Research,* 5(10), 241-254. doi:10.18009/jcer.325804

Okafor, K., Obayi, I., Ugwoke, F., & Ogbu, I. (2014). NigLT Ver 1.0: An Indigenous Language Translator Application for Major Ethnic Groups in Nigeria. *International Journal of Advanced Scientific and Technical Research,* 4(4), 882-890.

Priestley, M. (2000). *Practical Object-Oriented Design with UML*. Berkshire, England: McGRAW-Hill Publishing Company.

Shuaibu, G. (2014). Mathematics in Hausa Culture: Some Examples from Kano State-Nigeria. *IOSR Journal of Mathematics (IOSR-JM)*, 10(2), 167-171.

Waziri, M. Y., Saidu, I., & Musa, H. (2010). A Mathematical Approach on Solving Hausa Puzzles in Northern Nigeria. *International Conference on Mathematics Education Research (ICMER 2010). 8*, pp. 694-699. Procedia Social and Behavioral Sciences.

Wolff, H. E. (2013). *Hausa language*. Retrieved April 21, 2019, from Encyclopaedia Britannica.

Zaslavsky, C. (1973). *Africa Counts: Number and Pattern in African Cultures* (3rd ed.). USA: Lawrence Hill Books.