Available: DOI: https://doi.org/10.54117/gjpas.v2i2.100 Research article

Gadau Journal of Pure and Allied Sciences
Gadau J Pure Alli Sci, 2(2): 140-145 (2023)
ISSN: 2955-1722
DOI: https://doi.org/10.54117/gjpas.v2i2.100

Automated conversion of numeral to words in Hausa language

Rabi Mustapha*, Falalu Ibrahim Lawal and Muhammad Aminu Ahmad

Department of Computer Science, Faculty of Science, Kaduna State University, P.M.B 2339, Kaduna State, Nigeria.

*Correspondence: rabichubu@kasu.edu.ng

Abstract

Working with cardinal numbers is undoubtedly an integral part of our everyday life, which tells us about quantity. They are used for instance, in numeracy lessons, on receipts, slips, tellers in banks, to represent figures on documents, to tell time and are often represented inconsistently. However, the words form representations are often inconsistent with the standard Hausa. Standard Hausa as compared to other languages such as Arabic and British English, the Numeric System is characterized by a multitude of concatenations as well as dialect variations which causes misinterpretations and confusion. Also, initial findings show that no such automation published work has been done specifically in the numerical system of Standard Hausa language. In order to fill the gap, this study proposed an algorithm for converting numerals to words for the numeric system of Hausa language automatically in standard practice. The aim of the study is to examine the effect of iteration and mapping structures algorithm for converting numeral to cardinal form in standard Hausa language. The algorithm has been tested on approximately one hundred million sequences of numbers produced accurate results. The results showed that our method is successful in automating conversion of numerals to words with a high precision.

Article History

Received: 27/03/2023
Accepted: 25/05/2023
Published: 21/08/2023

Keywords
Algorithm;
Automation;
Conversion;
Hausa Language;
Numerical System

How to cite this paper: Mustapha, R., Lawal, F.I. and Ahmad, M.A. (2023). Automated conversion of numeral to words in Hausa language. *Gadau J Pure Alli Sci, 2(2):140-145*. https://doi.org/10.54117/gjpas.v2i2.100.

1.0 Introduction

It is estimated that there are about 7,139 spoken languages today in the World and 40% of which are endangered (Eberhard *et al*., 2021). Hausa language on the other hand is the widest spoken language in sub-Saharan Africa. Several factors such as; war, trade, search for livelihood, quest for knowledge and migration are attributed to the spread of Hausa language. However, the most indispensable aspect in the treasury of human language, the numeral counting system of Hausa language are also affected (Bunza, 2018). Hausa is thought to be one of the most unusual and sophisticated numerical systems of any natural language in the world (Bala, 2015; Garba, 2018). Introduction of the Arabic language, literary tradition extending back several centuries before the invasion of the colonial masters, has dispossessed many of the purely older names. Words such as '*gomiya'* and '*hauya'* (used for expressing multiples of ten, twenty respectively) for counting and many were borrowed from Arabic (Migeod, 1914). Doing with numbers is inherent in today's digital world where information is ubiquitous. Numbers are inconsistently written or printed in books, newspapers, etc. and on cheques, tellers, and papers etc., due to the evolutionary dynamic of languages, influence of the Arabic language, and a number of geographical dialects. One can notice variations in the dialects in the area of Kano, Zaria, Bauchi, Daura, Sokoto, Gobir, and northwards into Niger. Standard Hausa is coined based on '*Kananci'* the dialect of Kano and has been recognized as the norm for the written language as contained in books and mass communication (Africa: linguistik-und-sprachen/african-languages/hausa, 2015).

Undoubtedly in human history, numbers and numerals share a common root of origin and be that it may for Hausa method originated from either of the following (Bunza, 2018):

Journal of the Faculty of Science, Bauchi State University Gadau, Nigeria
Computer Sciences

1. **Fingers and toes**:

The system of counting using fingers has since been and still in use today in different parts of the world. People often use it to fill the communication gaps especially in the market places (Maikanti, 2014). Traditionally, Hausa people use fingers to count, mostly from right hand to left one - as the case may be.

Figure 1 illustrates how numbers from one to ten are demonstrated. This counting method uses a decimal system - *base ten*, if the counting terminates at *ten*, the two arms would be held together to translate *ten*. In the case of *twenty* the hands would be clapped twice, likewise *thirty* up to a *hundred*.

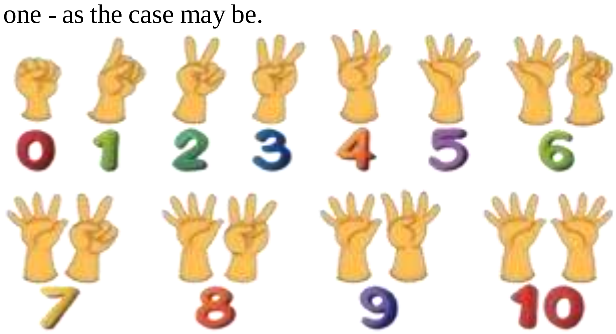

Figure 1: Counting using fingers

2. **Drawing/writing**:

Historically, ancient Hausa culture, counting is represented, see Table 1 by either drawing lines on the ground/sand or putting dots equal to the numbers

Table 1: Drawing/Writing method of counting

Numerals	1	2	3	4	5	6	7	8	9	10
Dots	.	..	...	.. ..	. .. ..	... ...	. ...	....	. ...	....

Following British colonization, which forcefully intruded Roman alphabet although, had to be learned from scratch into newly created education system in 1909 to slow down already established *Ajami* - the adapted Hausa language Arabic script (Abdalla, 2005). Today, this representation of the language has been superseded for most purposes by the Roman script (Africa: linguistik-und-sprachen/african-languages/hausa, 2015).

Phonemic inventory of Hausa

There are total of 34 consonants in Standard Hausa. The vowels comprising of 5 short vowels and 5 corresponding long vowels and 3 diphthongs making a total of 13. Some consonants are not found in English. Most common of these are usually the hooked letters i.e. *ɓ*, *ɗ*, *ƙ* and the semi vowel *`y*, which are entirely different from the corresponding plain letters b, d, k and y. Table 2 represents *Sakkwatanci* dialect, standard Hausa and the English equivalence of counting system from 1-10.

Table 2. SEQ Table * ARABIC 2: Cardinal

Num.	**Sakkwatanci Dialect**	**Phonetic Representation**	**Standard Hausa**
1	ɗăi/ɗăyaá	/ɗăi/ɗăjá/	ɗăyá
2	bíyú	/bíjú/	bíyú
3	úkkù	/úkkʷù/	úkù
4	húɗû	/húɗû/	húɗû
5	bìyát	/bìját/	bìyár
6	shíddà	/ʃíddà/	shídà
7	bákwài	/bákwài/	bákwài
8	tákwàs	/tákwàs/	tákwàs
9	tárà	/tárà/	tárà
10	goómà	/gómà/	gómà

The counting system from eleven changes with additional '*sha'* before the next unit with optionally additional word '*goma'*. The '*sha'* stands for 'swallow'. In this view, '(*goma*) *sha'* is to be repeated in counting up to nineteen '19' as represented in Table 3.

Table 3. SEQ Table * ARABIC 3: Cardinal representation of numbers 11-19

Num.	'Sakkwatanci' Dialect	Phonetic Representation	Standard Hausa
11	(gómà) shâ ɗãi	/(gómà) shâ ɗãi/ ɗãjá/	(gómà) shâ ɗãyá
12	(gómà) shâ bíyú	/(gómà) shâ bíjú/	(gómà) shâ bíyú
13	(gómà) shâ úkkù	/(gómà) shâ úkkwù/	(gómà) shâ úkù
14	(gómà) shâ húɗũ	/(gómà) shâ húɗũ/	(gómà) shâ húɗũ
15	(gómà) shâ bìyát	/(gómà) shâ bìját/	(gómà) shâ bìyár
16	(gómà) shâ shíddà	/(gómà) shâʃíddà/	(gómà) shâ shídà
17	(gómà) shâ bákwài	/(gómà) shâ bákwài/	(gómà) shâ bákwài
18	(gómà) shâ tákwàs	/(gómà) shâ tákwàs/	(gómà) shâ tákwàs
19	(gómà) shâ tárà	/(gómà) shâ tárà/	(gómà) shâ tárà

Compound numbers from twenty and greater, the numbers will then, be derived by adding '*da'* morpheme in between the two numbers 20 and the primary numbers (1-9) till it gets to 29 before the number would change to a unique number which is 30 see Table 4 for the examples.

Table 4. SEQ Table * ARABIC 4: Cardinal

Num.	'Sakkwatanci' Dialect	Phonetic Representation	Standard Hausa
21	àshìrín dà ɗãi	/àʃìrín dà ɗãi/ɗãjá/	àshìrín dà ɗãyá
22	àshìrín dà bíyú	/àʃìrín dà bíjú/	àshìrín dà bíyú
23	àshìrín dà úkkù	/àʃìrín dà úkkʷù/	àshìrín dà úkù
24	àshìrín dà húɗũ	/àʃìrín dà húɗũ/	àshìrín dà húɗũ
25	àshìrín dà bìyát	/àʃìrín dà bìját/	àshìrín dà bìyár
26	àshìrín dà shíddà	/àʃìrín dà síddà/	àshìrín dà shídà
27	àshìrín dà bákwài	/àʃìrín dà bákwài/	àshìrín dà bákwài
28	àshìrín dà tákwàs	/àʃìrín dà tákwàs/	àshìrín dà tákwàs
29	àshìrín dà tárà	/àʃìrín dà tárà/	àshìrín dà tárà
30	tàlàtín	/tàlàtín/	tàlàtín

Hundreds are formed starting with the word for hundred '*ɗari'*, followed by the multiplier unit,for example '*ɗari biyu'* i.e 200, '*ɗari uku*' i.e. *300*, '*ɗari huɗu*' i.e. 400 and so on. Thousands on the other hand, are formed starting with the word for thousand '*dubu*', followed by the multiplier unit. 2,000 is represented as '*dubu biyu*', 3,000 as '*dubu uku*', 4,000 as '*dubu huɗu*' and so on. The equivalent word for million is '*miliyan'*, for billion is biliyan, for trillion is '*tiriliyan*'.

Automatic conversion of numbers has attracted much attention due to its broad applications in day to day activities. Several algorithms have been proposed to automatically convert a number to its textual equivalent. Object-oriented approach in C++ to convert numeric values into corresponding words in the Uzbek language (Utkir, Mukhriddin, Bakhtiyor, and Zarmasov, 2020). A web-based English to Yorùbá numeral translation system implemented using Google Web App Engine with support for python and a Python package for natural language processing support NLTK, translates English numbers both in figure and text to its standard Yorùbá form has been proposed in (Agbeyangi, Eludiora, and Popoola, 2016; Olúgbénga and Ọdẹ, 2014). Another algorithm has demonstrated conversion of numbers to words, however the range is limited from 1 to 1000 and displays the result onscreen. For example, 235 would be "Two Hundred and Thirty-five." (Ben, 2017). It is undoubtedly that, unlike other languages, number forms in Tamil have recursive properties and details of their nuances in dealing with 90s, 900s, 9000s numbers, to word conversion using base-10 has been proposed in (Muthiah and Sathia, 2020). The algorithm works for parsing integral and floating point non-negative numbers; it parses text forms of numbers in both Indian and American standard (i.e. using millions, billions).

With all these developments, however, the methods were perfectly doing well when applied to a particular language, and do not apply to other languages, not even the dialects within a language. Although previous research showed successful outcomes, they for unknown reasons failed to discuss the feasibility of automating number conversion in Hausa language. With the complexity of the number system of Hausa language pointed out by Bala (2015), , the researchers however, considered the feasibility of automating the conversion of numbers into words in Hausa language.

2.0 Materials and methods

This section focuses on the design of a rule based algorithm with lookup functions using pseudocode. It was later implemented using Python Programming language. With the complexity of the number system of Hausa language pointed out by Bala (2015), However, we looked for the feasibility of automating the conversion of numbers into words in Hausa language.

3.0 Results and discussion

The algorithm was tested on approximately ten billion numbers, Figure 3 - 4 and produced accurate results with a high precision, Figure 5. This shows how powerful algorithms can be in automating seemingly difficult tasks. The result has also shown a rule-based system can be powerful and consistent in its task.

```
class Kalma
    procedure Kalma()
        alifiyya[""," dubu "," miliyan "," biliyan "," tiriliyan "]
        gomiya = ["",""," ashirin"," talatin"," arba'in"," hamsin",
                    " sittin"," saba'in"," tamanin"," casa'in"]
        daidai = [""," daya"," biyu"," uku"," hudu"," biyar",
                    " shida"," bakwai"," takwas"," tara"]
        sha = [" goma", " goma sha-daya"," goma sha-biyu",
                    " goma sha-uku"," goma sha-hudu"," goma sha-biyar",
                    " goma sha-shidda"," goma sha-bakwai"," goma sha-takwas",
                    " goma sha-tara"]
    endprocedure

    procedure fassara(lamba):
        if lamba == 0:
            return "sifiri"
        jumla = ""
        alif = 0
        while True:
            saura = int(lamba % 1000)
            if saura != 0:
                str = daruruwa(saura)
                majoni = ""
                if alif > 0 :
                    majoni = " da "
                endif
                jumla = alifiyya[alif] +str + majoni + jumla
            endif
            alif = alif + 1
            lamba = int(lamba/1000)
            if lamba <= 0:
                break
            endif
        return jumla
        endwhile
    endprocedure

    procedure daruruwa(lamba):
        jumla = ""
        saura = int(lamba%100)
        if saura < 10:
            jumla = jumla + daidai[saura]
        elif saura < 20:
            jumla = jumla + sha[int(saura%10)]
        else:
            majoni = ""
            if lamba%10 > 0 :
                majoni = " da"
            endif
            jumla = gomiya[int(saura/10)] + majoni + daidai[int(saura%10)];
        endif

        majoni = ""
        if lamba%100 > 0 :
            majoni = " da"
        endif
        if int(lamba/100) > 0:
            jumla = "dari" + daidai[int(lamba/100)] + majoni + jumla
        endif
        return jumla
    endprocedure
endclass
```

Figure 2. SEQ Figure * ARABIC 5: preview translated_Hausa.csv with Pandas

The algorithm has been subjected to a sequence of integer numbers from 1 to 9,999,999 generated by Python Figure 3:

```
for number in range(1,10,000,000,000):
```

Figure 3: Python generated numbers

Figure 4 shows the codes for

```
obj = kalma()
with open("translated_Hausa.csv", "w", encoding="utf-8") as myfile:
    for number in range(1,10000000):
        myfile.writelines(f"{number},{obj.fassara(number)}\n")
```

```
import pandas as pd
df = pd.read_csv(r'translated_Hausa.csv')
df.head
```

```
<bound method NDFrame.head of              1
daya
0         2                                biyu
1         3                                 uku
2         4                                hudu
3         5                               biyar
4         6                               shida
...     ...                                 ...
9999993 9999995  miliyan tara da dubu dari tara da casa'in d...
9999994 9999996  miliyan tara da dubu dari tara da casa'in d...
9999995 9999997  miliyan tara da dubu dari tara da casa'in d...
9999996 9999998  miliyan tara da dubu dari tara da casa'in d...
9999997 9999999  miliyan tara da dubu dari tara da casa'in d...
```

Figure 3. SEQ Figure * ARABIC 2: Pseudo code of the algorithm

1. Algorithm

The algorithm is divided into three different parts, *the Kalma* () method - a constructor that initiates the lookup arrays up on the creation of objects in the class. Another method *fassara* () takes a positive integer as an argument is called to make the translation task. The last method *daruruwa* () serves as a helper to the *fassara* () method. The algorithm converts integral numbers received to generate corresponding words for that number. Rule based systems (algorithms) that automatically translate numeral representations into word representations are of great interest to us.

4.0 Conclusion

We have successfully designed, implemented and experimented an algorithm, *Kalma*, and consistently performs well.

In future work, we plan to expand the algorithm to also convert floating numbers.

Declarations

Ethics approval and consent to participate

Not Applicable

Consent for publication

All authors have read and consented to the submission of the manuscript.

Availability of data and material

Not Applicable.

Competing interests

All authors declare no competing interests.

Funding

There was no funding for the current report.

References

Abdalla, U. A. (2005). Divergent Similarities: Culture, Globalization and Hausa Creative and Performing Arts. *International Conference on Literature in Northern Nigeria*, 4-5.

Africa: linguistik-und-sprachen/african-languages/hausa. (2015, February 10). Retrieved from Humboldt University of Berlin Web site: https://www.iaaw.hu-berlin.de

Agbeyangi, A. O., Eludiora, S. I., and Popoola, O. A. (2016, December). Web-Based Yorùbá Numeral Translation System. *IAES International Journal of Artificial Intelligence (IJ-AI), Vol. 5*(No. 4), pp. 127 – 134.

Bala, H. S. (2015). *A Comparatve Study of Numeric Systems of English and Hausa Languages.* Retrieved from Usmanu DanFodio University, Sokoto Website: oer.udusok.edu.ng

Ben, T. (2017). *GameMaker: Studio 100 Programming Challenges.* West Sussex: Apress.

Bunza, A. M. (2018). *Mathematical Heritage In Hausa Number System: (A Proposal for TeachingMathematics using Nigerian Languages).* Federal University, Gusau, Languages and Cultures. Gusau: University Research Center, Federal.

Eberhard, David, M., Gary, F. S., and Charles, D. F. (2021, October 20). *Guides: how-many-languages*, 24. (E. L. World, Producer, and Dallas, Texas: SIL International) Retrieved October 20, 2021, from SIL International Web site: http://www.ethnologue.com

Garba, F. A. (2018). Endangered Language In Nigeria: A Case Study Of Gera Language Of Bauchi State. Kenyatta University, Department of English and Linguistics. Kenyatta: Kenyatta University Libary. Retrieved October 21, 2021

Maikanti, S. (2014). Counting System In 'Sakkwatanci' Dialect Of Hausa Language. In *Counting System Of Some Dialects Of Igbo, Hausa And Yorùbá* (p. Department of Hausa). Ososami Road, Oke-Ado, Ibadan: Masterprint Publishers.

Mbah, E. E., Ndubuisi, O. A., and Ayegba, M. (2014). A Comparative Analysis of the Numerals and Counting Systems of Igbo and Igala. *International Institute for Science, Technology and Education (IISTE)*, 61-65.

Migeod, F. W. (1914). *A Grammar of the Hausa Language.* London: KEGAN PAUL, TRENCH, TRUBNER CO . , LTD.

Muthiah, A., and Sathia, M. (2020). Generation and Parsing of Number to Words in Tamil. *Tamil Internet Conference* (pp. 1-16). Tamil: INFITT (International Forum for Information Technology in Tamil).

Olúgbénga, O. A., and Ọdẹ, ́. A. (2014, August 13). Computational modelling of Yorùbá numerals in a number-to-text conversion system. *Journal of Language Modelling*, pp. 167–211.

Utkir, K., Mukhriddin, M., Bakhtiyor, A., and Zarmasov, E. (2020). A Novel Algorithm of Numbers to Text Conversion for Uzbek Language TTS Synthesizer. *2020 International Conference on Information Science and Communications Technologies (ICISCT).* Tashkent: IEEE Xplore. doi:10.1109/ICISCT50599.2020.9351434.