

SELECTED AI-POWERED CHATBOTS' PATTERNS OF IDENTITY CONSTRUCTION OF THREE MAJOR ETHNIC GROUPS IN NIGERIA

Ugoala Bibian ¹ 

Department of English

Faculty of Arts; National Open University of Nigeria, Abuja, Nigeria

bugoala@noun.edu.ng

Abstract: The proliferation of chatbots in online interaction raises critical questions on how identity is constructed in human-machine discourse. This study aims to integrate insights from the Social Identity Theory and Cognitive Load Theory to examine how selected chatbots (Copilot, Gemini, and ChatGPT,) construct identity about three major ethnic groups in Nigeria. Three prompts in relation to identity construction about the three major ethnic groups in Nigeria were fed into each large language model, the responses they gave served as the data for the study. The specific objectives are: to unravel how chatbots through their responses in relation to (name, food, and dance) identify cultural traits of a particular ethnic group; to describe the pattern of “be” verb in chatbots’ responses to cultural identification of a particular ethnic group, and to investigate the pattern of chatbots’ responses to “wh” questions in relation to cultural identification of particular ethnic group. Chatbots generate responses from large data sets, however, this study argues that large language models’ responses to prompts does not truly represent the fluidity and naturalness which characterize human interactions. The study found that in some cases, the chatbots fail to properly distinguish the cultural identity of a particular region. They also in some cases in their responses, fail to respond in relation to the tense in which a prompt is structured. As chatbots gradually permeate all sectors, the study recommends that when seeking the true identity of an ethnic group, chatbots are not a reliable option, they lack the originality with which human interaction is known for.

Keywords: Artificial intelligence; Chatbots; ChatGPT; Copilot; Gemini

How to cite this article:

Cite as: Ugoala, B. (2026). Selected AI-powered Chatbots’ Pattern of Identity Construction of Three Major Ethnic Groups in Nigeria, *Journal of Studies in Language, Culture, and Society, (JSLCS)* 9(2), PP. 176-187

¹ Corresponding author :Ugoala Bibian ORCID ID: <https://orcid.org/0000-0002-6628-4400>

1. Introduction

The 21st century is remarkable in advancing technology in the different sectors of the society. Part of this advancement is societal shifts as a result of globalization and increased internet connectivity of people from diverse cultures and backgrounds. The period is defined by the upgrade of smartphones, skills like adaptability to traverse prompt change, critical thinking, innovation, and artificial intelligence. In all these, artificial intelligence (AI) seems to be one of the tools making more waves in human affairs, as it enables machines to learn, organize and solve human problems much like humans would do by using intelligent algorithms and large datasets generated online. In terms of human interactions, AI can comprehend and decode both spoken and written language. AI-powered machines now interact like humans, thus, raising critical questions on how human cultural identity is presented on these chatbots. Large language models like ChatGPT, Microsoft Copilot, and Google Gemini, are AI-powered Chatbots which enhance human competencies in performing tasks in faster by gathering responses to prompts through large datasets online. The study aims to examine how chatbots through the prompts fed into them portray cultural identity of three major ethnic groups in Nigeria. Dou & Rim (2025, p. 66) note that “the art of crafting precise instructions to elicit desired responses from these complex AI systems is known as ‘prompt engineering’”. The specific objectives of the study are: to unravel how chatbots through their responses in relation to (name, food, and dance) identify cultural traits of a particular ethnic group; to describe the pattern of “be” verb in chatbots’ responses to cultural identification of a particular ethnic group; and to investigate the pattern of chatbots’ responses to “wh” questions in relation to cultural identification of particular ethnic group. The location of the current study is Nigeria. Nigeria is located in West Africa. It borders Benin Republic to the west, Cameroon to the east, Chad and Niger to the north, and the Gulf of Guinea to the South. It is the most densely populated country in Africa. It is a country with the most ethnic groups in Africa. Oluwadele et al. (2023, p. 62) note that Nigeria is said to have over 250 ethnic groups. Eyiuche et al. (2020, p. 2) point out that the three major ethnic groups in Nigeria are Hausa, Igbo and Yoruba due to the fact that they occupy a wider geographical region and has many speakers.

1.1 Large Language models (LLMs)

Large language models (LLMs) are deep learning systems crafted and designed to process language; Large language models unlike language are not natural. Language is natural vocal meaningful sound made by humans for interaction. It is the foundation of human communication and plays important roles in shaping human exchanges with the world. Large language models are open to large assemblages of data and texts, making them able to understand and generate natural language and other constructions contents in order to perform a wide range of tasks. Their training objective is to effectively forecast the next word or missing word in a phrase, clause, sentence or paragraph. This objective is known as “language modeling”. They study the distributions of words, the syntactic structures of a language, learning how and what sentences are like. Large language models learn sequences in texts and generate a language that follows those patterns; they copy how utterances behave. Natural language processing (NLP) and machine learning research helped in the emergence of LLMs; also the development of Natural language processing (NLP) and machine learning research is mainly responsible for the emergence and advancement of artificial intelligence during the late 2010s and 2020s.

Large language models are increasingly being applied in so many disciplines like Chemistry, linguistics, computer science, and the humanities. They stand out as groundbreaking advances, leading a transformation with their remarkable language understanding and generation capabilities. However, while the deployment of LLM to tasks

comes with some advantages, it also comes with some challenges. Zhu et al. (2025) investigate the possible problems of large language models in the peer review process, using Claude 2.0 to generate citation requests, peer review reports etc. Artificial intelligence recognition tools like GPTzero, and zeroGPT were deployed to know if the reviews were noticed as large language models' generated. The study found that the large language model-generated reviews were somehow similar to human reviews, but lacked depth in that they lacked detailed critique. The study concludes that large language models can be easily be misused by users to weaken the peer review process by producing subjective, difficult-to-detect content, and confusing output, this poses serious threat to academic integrity. The study further recommends that detection tools and guidelines are necessary to make sure that large language models improve rather than undermine the peer review process.

Zhang et al. (2025) note that in spite of the advanced language generation and comprehension capabilities of large language models, they still face major challenges such as security, safety, privacy issues that weaken their reliability and effectiveness. As a result of these challenges, Zhang et al. (2025) provide an all-inclusive overview of the weaknesses and security mechanisms that can help booster large language models deployment in diverse fields.

Large language models are accessible to the public through chatbots like Microsoft's Copilot, Open AI's ChatGPT, Google's Gemini assistant, Meta Llama models, and Anthropic's Claude. LLMs process language and generate replies, but they cannot on their own maintain continuous conversation or recall past interactions. In spite of the advances by LLM in human spheres, this paper interrogates critical concerns of humanity, like cultural issues as responded to by Large language model-powered chatbots. In this study, the data is got from the first three mentioned (Microsoft's Copilot, Open AI's ChatGPT, and Google's Gemini assistant). These chatbots deploy advanced AI to understand and generate human-like text.

1.2 Chatbots

Chatbots are artificial intelligence crafted to mimic human-like conversations. Chatbots are designed to handle continuous dialogues, and also remember previous conversations; this allows them to generate comprehensible context-aware replies. Large language models like GPT-4o, Claude 3.5, Mistral-7B, and Llama-3 are powerful engines driving chatbots.

Chat Generative Pre-Trained Transformer (ChatGPT) is a generative artificial intelligence chatbot developed by OpenAI; it was launched in November 2022. It is a deep learning model that deploys generative artificial intelligence algorithms or set of rules to make available personalized answers to individual end user. ChatGPT uses GPT-5 which is a generative pre-trained converter to generate speech, images, and text in response to individual prompts. It leverages natural language processing (NLP) systems and recollects pieces of information from earlier interactions to make human-like exchanges that can be deployed for diverse purposes. According to Almarie et al. (2023), ChatGPT is one of the most widely used large language model that has gained global attention. Large language models like ChatGPT is a technological advancement of the 21st century changing different phases of society through improving human abilities and capabilities in tasks like research, and writing, at the same time introducing novel challenges in ethics, systems, economies, education, and daily lives. ChatGPT can generate text output like writing creative content and responding to questions in a wide-ranging and instructive way, in natural language understanding, it can analyse and understand the semantics of text as well as identify entities and relationships in sentences. Based on an inputted question, ChatGPT can answer questions, and provide answers to spontaneous questions.

Google Gemini is a multimodal Artificial Intelligence that functions as an advanced chatbot. It is Google's innovative conglomerate of diverse semiotic resources artificial intelligence models and an assistant powering diverse Google products. Google Gemini deploys its own personal of multimodal large language models known as Gemini. It is able to comprehend and generate codes, videos, images, audio, and text integrating with Google apps for individual assistance. Google Gemini is successor to Google Assistant, however with greater intimate depth. It summarizes bulky document. The multimodal capabilities of Google Gemini give it the leverage to process and combine diverse data types in a single interface, allowing the user to make enquiries about pictures, or generate content from an audio text. Google Gemini connects with individual's Google bionetwork to find pieces of information, and summarise content. Gemini has been trained on large corpus of multimodal and multilingual data sets. It deploys a converter model and a neural network design that Google introduced in 2017. Gemini family of multimodal AI models comes in diverse forms (like Gemini 1.5 Flash, and Gemini 1.5 Pro), each form is improved for different devices and tasks. Gemini AI can work across programming languages to comprehend, explain, and generate code. Gemini can be used to mine text from images and caption images. It can analyse non-linguistic elements like figures, diagrams and charts.

In late August to early September, 2024 Google began releasing a new feature known as “Gems” that allows users adapt the Gemini chatbot to perform any task or ask questions on any topic. Users gained access to “Gems in September, 2025. The new feature helps to break difficult topics for easier comprehension. However, Google Gemini has its own shortfall just like other large language models in that sometimes it gives inaccurate historical figures.

Microsoft Copilot is a generative Artificial Intelligence chatbot developed by Microsoft AI. It uses a combination of potent models, mainly OpenAI's GPT series like GPT-4o and advanced GPT-5. Microsoft Copilot was launched in 2023 under the name Bing Chat, as a built-in feature for Microsoft Edge and Microsoft Bing. Throughout 2023, Microsoft unified the Copilot branding across it numerous chatbot products, thus bolstering the Copilot correlation. Microsoft Copilot is built on and deploys large language models to provide context-aware AI assistance for output; it uses advanced large language models trained on large datasets to comprehend and generate human-like text. Microsoft Copilot is an AI-powered assistant used of numerous tasks such as boosting productivity, generating creativity, and assisting with general information through web searches. It combines with Microsoft 365 apps like Excel, Word, Outlook, Powerpoint, and Team to assist work-related tasks like drafting emails, summarizing documents and creating presentations deploying natural language. The broader Copilot can be used for content generation and all-purpose questions. Copilot is powered by aa amalgam of Microsoft technologies and OpenAI.

2. Literature Review

The proliferation of Large language models and chatbots in different fields of human endeavours have sparked off interesting researches. For example, Raiaan, et al. (2024) discuss the central concepts of Large Language Models and provide an overview of the existing works on Large Language Models, including their history, transformers, resources, training methods, applications, impacts, challenges, and so on. The study amongst things illustrates how Large Language Models can impact positively on the society, and how AI can be used to solve problems. The study also examines issues and challenges that the deployment of AI to real-world issues can generate. Through these, the authors envisage that their study will help

researchers, practitioners, experts and so on to truly understand the progression of Large Language Models, their application, challenges and future aims.

Caroll & Borycz (2024) examine the scope of students' use of Artificial intelligence tools in research-related assignments and whether information literacy training could impact their awareness of these tools. Students' responses to Likert-scale questions show that so many students lack extensive understanding using large language models for research-related purposes before the information literacy sessions. Though after the students were exposed to a didactic lecture with an engineering librarian that explained how to use the tools, many students reported that the tools will be useful for future tasks. This shows that librarian-led information training can help students in developing more refined comprehension of the challenges and use cases for AI inquiry-based assignments.

Naveed et al. (2024) review article provide an overview of the literature on a broad range of LLM-related concepts. The article provides a systematic survey, quick, and comprehensive reference for the researchers and practitioners working on LLM to draw insights from wide, instructive reviews of the prevailing works to progress the LLM research.

Cao et al. (2025) note that the general intelligence characteristics of LLMs are mainly reflected in three aspects: technologically, they exhibit common sense, deep reasoning, strong generalization, and natural human-computer interaction; in terms of intelligence, they demonstrate memory-driven core features, powerful data-driven learning capabilities, and exceptional generalization abilities; in terms of thought, they possess highly humanlike cognitive traits such as contextual understanding, analogy, and intuitive reasoning.

Shahzad et al. (2025) explore LLMs' types, history, and training processes, alongside their application in education, including digital and higher education settings. A novel theoretical framework is proposed to guide the integration of LLMs into education, addressing key challenges such as personalization, ethical concerns, and adaptability. The study presents practical case studies and solutions to barriers, such as data privacy and bias, offering insights into their role in enhancing the teaching–learning process. The study advances knowledge through highlighting the potential of LLM in transforming education.

Labadze et al. (2023) examine the benefits of chatbots in pedagogy, their prospects, pitfalls, and concerns in deploying artificial intelligence in educational settings. After conducting a wide search across several databases, and applying some criteria, 67 relevant studies were chosen for review. The study came up with the findings of the numerous benefits of incorporating AI in education as perceived from educators' and learners' perspectives. Learners gain from AI-powered chatbots in the area of development of numerous skills; homework and study assistance; and an individualized learning experience. The study discovered that AI-powered chatbots assist educators with time-saving assistance and improved teaching methods. The study pointed out that in spite of the advantages of AI-powered chatbots, educators need to take into consideration issues like ethical considerations, accuracy, and reliability, in AI applications.

Alva & Caceres (2025) examine how human characteristics of chatbots like competence and warmth, influence customer experience in the banking section of Lima, Peru. Due to the uncreasing deployment of artificial intelligence in the financial sector, the study set to understand chatbots influence on users' loyalty and satisfaction. Deploying a non-experimental quantitative method, the study evaluated if the social presence of chatbots boosts users' perception especially among youngsters. Findings show that though users value chatbots' human characteristics, but neutral responses persist showing the absences of emotional connection in some cases. This highlights the significance of optimizing chatbots adoption to sustain competitiveness in the digital ear. The study foregrounds the importance

of adopting artificial intelligence solutions to customer expectations in order to promote a more positive and individualized experience.

2.1. Social Identity Theory

The theories guiding this study are Tajfel & Turner (1986) Social Identity Theory (SIT), and Sweller (1988) Cognitive Load Theory. Tajfel & Turner (1986) posit that people have cultural traits that unite them or set them apart. Group membership unites members of a group, and they favour one another because they share same characteristics like naming, food pattern, dressing, dance etc. Through these traits, a group is identified. People in the same group generally maintain positive social image of their group, leading to a shared identity. For example, identity is not just personal “I am Nkechi”, but also social “I am Igbo”, “I am a student”. A group strives to show positive in-group preference through social identification, social categorization, and social comparison, and out-group prejudice. Social identification showing how individuals embrace the identity of their in-group connecting their self-esteem to the noble and worthy status of the group. By Social categorization Tajfel & Turner (1986) show that people share themselves and others into social groups into in-groups, and out-groups in order to comprehend the social world. People place themselves and others into social groups such as “fan”, “professional”, or “entrepreneur”. Social comparison shows that groups equate themselves to out-groups often making them to sustain an affirmative social identity and self-prestige. The theory provides a framework for comprehending how group dynamics shape our identity and interfaces in the world at large. Social Identity Theory explains how group affiliation such as nationality, religion, gender, ethnicity and profession are constructed and signaled through the use of language.

4.2. Cognitive Load Theory

Humans are not robots that receive continuous prompts without resting, or receive excessive information. Human cognitive ability has limits at a particular time, context, or situation. This perhaps informs Sweller (1988) Cognitive Load Theory which is based on the idea that humans have limited working memory, this makes them to have limited understanding, and ability to learn effectively. The theory is based on the idea that human working memory is limited in capacity and duration, that learning fails when the memory is overloaded.

Cognitive load theory is a framework that aims at enhancing teaching and learning by taking into consideration the role and hindrances of working memory especially in relation to the cognitive load experienced during assignments and tasks. The theory is significant in its postulations of the significance of organizing information in bits in order to reduce cognitive overload and improve task performance. Researchers such as Cowan (2010) notes that human working memory can handle only small bits of information at once, if one exceeds this limit, then concentration wanes, learning is impeded, and misunderstanding comes up. Learning breaks down when the brain is overloaded. The theory lays emphasis on three types of load: intrinsic cognitive load refers to the mental energy required by the complex nature of the information or task that is to be learnt. Extraneous cognitive load refers to the needless work caused by how the information is presented, such as bad presentations, unnecessary distractions like noise, confusing instructions etc. Extraneous cognitive load leads to information overload, in that instead of an individual learning new things, the excess information becomes a hindrance due to the fact that an individual's processing capacity is exceeded. All forms of excess pieces of information more than required is categorized as extraneous cognitive load, which are not important to the main task or message. This is often as the information is presented either in the form of excessive details, unwanted details, or poor organization. Germane load is the productive energy wasted processing information to

form mental schemas for long-term memory; such as actively linking new concepts to already existing knowledge. Cognitive load theory is significant to the present study in that it helps to explain what information overload as noticed in the chatbots does to the receivers.

3. Methodology

3.1 Context

The study is carried out in Nigeria, and it adopts the descriptive method of enquiry to interrogate the linguistic items in chatbots responses about name, food, and dance about the three major ethnic groups in Nigeria (Igbo, Yoruba, and Hausa). Name, food, and dance are some of the items through the cultural traits of ethnic groups are identified.

3.2 Participants

Names are vital to the identity of a people (Lieberson, 2000; Hakala et al., 2020; Chi-Yue Chiu et al. 2010; Fahlbusch & Heuser, 2015; Richard, 2018; Reszegi, 2020; Völkel, 2023). Richard (2018) sees name as important element in cultural identification, also, Reszegi (2020) notes that names are primarily linguistic elements with an identifying role; and further avers “that names that exhibit linguistic and cultural constraints also play an important role in this identity formation” (Reszegi, 2020, p. 154). Völkel (2023, p. 151) submits that names are given to entities of cultural importance and of significance for humans in social life; name usage and name-giving practices reflect social beliefs and cultural values; and their meaning is strongly shaped by cultural and historical contexts. Naming of humans is so vital to identity of a group this is evident in Fahlbusch & Heuser (2015: 18) assertion that “... no society is known in which humans are not given names”. Scholars have also shown that dance is related to the culture of ethnic groups. Aulia (2023), notes that “Dance often reflects cultural traditions, values, and stories contained in its movements, costumes, music, and historical background”. Zhang (2023, p 67) notes that “... dance is an art form that conveys emotions through the language of the body while also depicting the living conditions of various ethnic groups.” Bakka et al. (2025) note that dance is a unique element that shows the culture of a people.

Monin & Szczurek (2014, p. 155) on the relation of food to culture, point out that “Given the paramount importance of food in daily life, food can provide a valuable window through which to explore and understand cultural contexts”. Food is also part of the culture of an ethnic group, Savannah (2023) says that “Food is considered a part of intangible cultural heritage, a way of life that is passed down from generation to generation”. Food is an essential part of cultural identity.

3.3 Procedures

The study is qualitative as it focuses on detailed descriptions and interpretations of the data set based on the research questions. Three sets of questions for each ethnic group about name, food, and dance in relation to identity construction about the three major ethnic groups in Nigeria (Igbo, Hausa, and Yoruba) were fed into Copilot, Gemini, and ChatGPT, the responses they gave served as the data for the study. The questions were: What tribe in Nigeria bears the name Amaka? What tribe in Nigeria bears the name Olusegun? What tribe in Nigeria bears the name Danladi? Which native food is special to Igbo? Which native food is special to Yoruba? Which native food is special to Hausa? Can you tell me the name of Igbo’s special dance? Can you tell me the name of Yoruba’s special dance? Can you tell me the name of Hausa’s special dance? Each of the names (Amaka, Olusegun, and Danladi) are names unique to each of the ethnic group. For the analysis, insights from the Social Identity Theory of Tajfel & Turner (1986) and Cognitive Load Theory by Sweller (1988) were applied

to the chatbots' responses to examine how selected chatbots (Copilot, Gemini, and ChatGPT,) construct identity about three ethnic groups.

4. Results

The first question fed into Copilot: "What tribe in Nigeria bears the name Amaka?", shows that Copilot was able to identify the name Amaka as being borne by the Igbo tribe of Nigeria. It further gives the meaning of the name, the gender association of the name, and popularity of the name and asks further question: *Want to explore other beautiful Igbo names or their meanings?* For Yoruba, Copilot identifies Olusegun as a Yoruba name, it gives the meaning, origin, cultural significance, and a notable bearer of the name as a former president of Nigeria. It also asks the question: *If you're curious about other Yoruba names and their meanings—or how naming traditions differ across Nigerian tribes—I'd be happy to dive deeper!* For Hausa, Copilot identified Danladi as a Hausa name, goes on to elaborate by narrating the origin of the name and uses, asks a question: *Want to explore other Hausa names and their meanings?*

The next chatbot chosen for the study, Gemini identifies the name Amaka as from the Igbo people of southeastern Nigeria, gives the longer versions of the name as "Chiamaka", and explains the name. For Yoruba, Gemini identifies the name Olusegun as a Yoruba name. It gave another spelling of the name as "Oluwasegun", it explains the name, and its significance. Gemini was also able to identify the name "Danladi" as of Hausa origin; it went further to explain the name.

ChatGPT identifies Amaka as Igbo name borne by the Igbo people of southeastern Nigeria. Like Gemini, ChatGPT gave the full version of the name as "Chiamaka", and explains it. It asks the question: *Would you like me to also give you a list of other common Igbo names and their meanings, similar to Amaka?* For Yoruba, ChatGPT, identifies the name "Olúṣégún" as of Yoruba origin, it breaks the name into morphemes and explains each in English; it ends up asking the question: *Would you also like me to compare this with Igbo names that carry similar meanings of victory or triumph?* ChatGPT identifies Danladi as a Hausa name, it ended up asking the question: *Would you like me to also explain other Hausa day-names like Ladi, Tanko, Gomina, etc., which are given based on the day or circumstance of birth?*

In respect of the second prompt which is "which native food is special to Igbo, (Yoruba), (Hausa)". Copilot gave eight types of food as special to Igbo: Osikapa (Rice Dishes), Yam with Goat Meat Pepper Soup, Abacha (African Salad), Nkwobi, Ogbono Soup, Egusi Soup, Ofe Onugbu (Bitter Leaf Soup), Ofe Oha (Oha Soup), in this responses, we can see overelaboration and information overload. The prompt has singular subject, and singular verb, but Copilot did not observe this in its response. After listing the eight types of food, goes on to ask: *Would you like a recipe for one of these dishes or tips on how to prepare them at home?* For Yoruba, Copilot begins with an introductory statement: Here are some of the most iconic Yoruba foods: and lists eight kinds of food for Yoruba: Amala, Ewedu Soup, Gbegiri, Efo Riro, Ofada Rice with Ayamase Sauce, Asun, Akara and Moi Moi, and asks: *Want to dive into how one of these is prepared or explore its cultural significance?* For Hausa, Copilot begins with and an introductory statement like it did for Yoruba: Here are some standout Hausa delicacies, and lists five types of foods for Hausa: Tuwo Shinkafa, Koko and Kosai, Masa (Waina), Miyan Kuka, Dambu, and ends by asking the question: *Want to try making one at home or learn about the cultural stories behind them?* The determiner "some" is not in sync with the singular subject and singular verb in the prompt.

In Gemini's response to the second prompt, seven kinds of food were listed for Igbo: Ukwa (African Breadfruit), Oha Soup, Nkwobi, Abacha (African Salad), Ofe Owerri, Ofe Onugbu (Bitter leaf Soup), Ofe Egusi (Melon Seed Soup), and Isi Ewu (Goat Head). Gemini gives excess information through not adhering to the singular subject and singular verb in the prompt. For Yoruba, Gemini listed eight kinds of food: Àmàlà, Ewédú Soup, Gbègìrì Soup, Èfò Rírò, Òfàdà Rice and Àyàmàsè, Mói Mói, and Iyan. For Hausa, Gemini was specific in line with the verb of the prompt using "is": The native food that's most special to the Hausa people of northern Nigeria *is* Tuwo Shinkafa.

ChatGPT listed six types of food as special to the Igbos: Ji (Yam), Ukodo (Yam Pepper Soup), Nkwobi, Abacha (African Salad), Ofe Oha (Oha Soup), Ofe Onugbu (Bitter leaf Soup), Fufu (Akpu). Of the listed food, Ukodo (Yam Pepper Soup) is not an Igbo delicacy—but Urhobo (a tribe in Delta state, Nigeria). After listing the foods, ChatGPT, further asks: *Would you like me to highlight the ceremonial foods like those used in festivals, marriages, and funerals or the **everyday delicacies** that are most loved?* For Yoruba, ChatGPT listed six types of food: Èwà Agoyin (Mashed Beans with Sauce), Àmàlà, Èfò Riro (Vegetable Soup), Gbegiri (Bean Soup), Òbè Egú sí (Egusi Soup), Ofada Rice and Ayamase (Designer Stew). Like Copilot, ChatGPT asked: *Would you like me to compare the Yam (Igbo) and Amala (Yoruba) traditions side by side, since both foods hold cultural pride for their people?* For Hausa, ChatGPT mentioned several dishes: Tuwo Shinkafa, Miyan Kuka (Baobab Leaf Soup), Kilishi, Suya, Fura da Nono, Waina (Masa). It asks: *Would you like me to create a comparison of the "special native foods" of the three major tribes (Igbo, Yoruba, Hausa) side by side, so you can see their cultural uniqueness?*

The third prompt which centres on dance, has Copilot identifying the dance as "Atilogwu" as unique to Igbo; it elaborately explained Atilogwu stating its origin, meaning, music and costume, occasions, and style. Thereafter asked: *Want to dive deeper into its history or see how it's performed today?* For Yoruba, Copilot mentioned five types of dance: Bata Dance, Egungun Dance, Gelede Dance, Eyo Dance, and Sakara Dance. It asked: *Would you like to explore the music or costumes that go with one of these dances?* Copilot listed "Koroso dance" as unique to Hausa, and further asked: *Want to see how it compares to other Nigerian dances like Igbo's Atilogwu or Yoruba's Bata?*

Dissimilar to Copilot, Gemini mentioned five types of dance for Igbo: Atilogwu, Ikorodo, Mmanwu (Masquerade Dance), Egedege, and Ikpirikpi-ogu. Gemini mentioned seven dance types for Yoruba: Bàtá dance, Èyò dance, Ègúngún dance, Àpàlà dance, Gèlèdé dance, Dùndún dance, and Ògbè dance. For Hausa, Gemini, like Copilot named Koroso dance, went ahead to over inform stating the origin, style, and performance. It further mentioned the Sharo dance as unique to Hausa.

ChatGPT identifies "Atilogwu" as unique to Igbo, and asked: *Would you like me to also mention other notable Igbo dances apart from Atilogwu?* ChatGPT responds for Yoruba by listing the Bàtá dance, Ìjálá (hunters' dance), Apala dance, and Àgbékè (a women's dance), and asked: *Would you like me to also give you a comparison between the Igbo's Atilogwu and the Yoruba's Bàtá so you can clearly see the cultural differences?* ChatGPT identified "Koroso." as Hausa's special dance, and also named two other minor dances "Rara" and "Hawan Sallah". It further asked: *Would you like me to also tell you the special dances of the Igbo and Yoruba, so you can see how each major tribe's dance reflects its culture?*

5. Discussion

Chatbots—Copilot, Gemini, and ChatGPT in their responses to name, food, and dance to identify the cultural traits of a particular ethnic group overelaborated in their responses leading to information overload, showing some level of uncertainty of the question asked. In order to want to get the answer right, the chatbots gave so many responses in various forms. This shows that chatbots are not very exact in their responses to questions or prompts, this way, they cannot be seen as authentic sources of cultural identification of a particular ethnic group. Cognitive Load Theory by Sweller (1988) submits that due to the fact that working memory is limited in scope, instructional materials and learning skills should be limited to working memory load, this will enhance efficient schema formation. This means that, the brain can only hold some information at a time, and needs not to be over loaded with too much information. Seufert, Jänen, and Brünken (2007) corroborate this as they found that learners with low prior knowledge are particularly vulnerable to overload. This contrasts with the chatbots' form of response to questions or prompts, as they over load the learner with too much information. The analysis show that chatbots are also not very certain about food items for each ethnic group, this is seen in ChatGPT's response for the special food for the Igbos. It used a kind of food "Ukodo" that is for another ethnic group (the Urhobos in Delta state of Nigeria) as an Igbo food. The word "Ukodo" is an Urhobo word for local yam porridge spiced with spices.

The second objective of the study which is to describe the pattern of "be" verb in chatbots' responses to cultural identification of a particular ethnic group shows that chatbots are not very specific all the time in response to the verb form in questions fed into them. This is the case with Gemini; its response for food for Igbo and Yoruba did not adhere to the "be" singular form of the verb in the question: "which native food is special to Igbo, Yoruba, Hausa". The singular verb "is", in the question requires a singular answer, but Geminin gave the plural form, mentioning more than what the verb required. For Hausa, it adhered to the singular verb through identifying only one kind of food for Hausa in response to the singular verb "is", showing inconsistency in its response. Chatbots, apart from not adhering to the verb form in questions, also do not tailor their responses to adhere to "wh questions". This is observable in the questions fed into the chatbots beginning with "What tribe in Nigeria bears ...", and "Which native food is ...". In response to these "wh questions", Copilot listed eight types of food for Igbo and Yoruba, and lists five for Hausa, showing that it could not identify questions requiring just a single response.

The study also discovered that some chatbots follow the question-response-question pattern in response to a user's question. This means that, when a user feeds chatbots like Gemini, and ChatGPT with a question, the chatbot responds, and asks the user an end-of-response question as a form of closing. Copilot in its end-of-response question uses the structure: "Want to explore..." to cue the user into asking further questions. ChatGPT in its end-of-response question uses the modal auxiliary "would you want to ..." in some cases, and omits the subject "you" in some other instances in its end-of-response questions "want to ..." instead of "Do you want to ...". The current study agrees with Alva & Caceres (2025) who found that chatbots though chatbots' have some humanlike characteristics, but they lack emotions which characterise human interactions.

6. Conclusion

For now, chatbots cannot be seen as authentic tools in the cultural identification of particular ethnic groups. Chatbots leverage the large language models assess to millions of data stored online to provide "likely answers" to questions. Large language model is an algorithm for giving or producing the response that most sounds like a response to a question,

collated from a large pool of answers provided by humans previously. From the data, it shows that chatbots pool so many answers from online to come up with a conglomeration of answer that is accurate. When answers are not enough, it will sieve from the few answers available online to give an approximation that looks like a correct answer, but may not be very correct.

Artificial Intelligence Declaration

The author declares that no generative artificial intelligence (AI) tools were used in the conception, analysis, writing, or preparation of this manuscript.

Ethical Declaration

This study was conducted in accordance with the ethical principles for research involving gathering of secondary data from online chatbots for the study. All data sets were anonymised.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions

The author solely conceived and designed the study, wrote the abstract, the introduction, conducted the literature review, collected and analyzed the data, interpreted the findings, drafted and revised the manuscript, and approved the final version for publication.

References

- Almarie, B., Paulo E. P. Teixeira, Kevin, P., Carlos, Augusto, R., & Felipe, F. (2023). Editorial - The Use of Large Language Models in Science: Opportunities and Challenges. <https://doi.org/10.21801/ppcrj.2023.91.1>
- Alva, G., & Caceres, M. (2025). Human Characteristics in Chatbots Cause an Impact on Customer Experience in The Banking Sector. *Journal of Technology Management and Innovation*, 20(2), 52–66. <https://doi.org/10.4067/s0718-27242025000200052>
- Aulia, O. (2023). Interpreting Identify through Dance, Cultural Imaging and Social Meaning. <https://medium.com/@oniaulia/interpreting-identity-through-dance-cultural-imaging-and-social-meaning-20cc7077e6f2>
- Bakka, Egil, Kunej, Rebeka, Sarkar Munsu, Urmimala, Savage, Patrick E., and López-Yáñez, María Gabriela. 2025. “Has Dance Become Definable?” *Traditions of Music and Dance* 57 (1), 115–137, <https://doi.org/10.1017/tmd.2025.10026>
- Blank, I. A. (2023). What are large language models supposed to model? *Trends in Cognitive Sciences*, 27(11), 987-989. <https://doi.org/10.1016/j.tics.2023.08.006>
- Cao, Z., Zhang, X. & Zeng, D.D. (2025). Large language models: Technology, intelligence, and thought. *Frontiers of Engineering Management*, 12, 710–715. <https://doi.org/10.1007/s42524-025-5004-3>
- Caroll, J.A. & Borycz, J. (2024). Integrating large language models and generative artificial intelligence tools into information literacy instruction. *The Journal of Academic Librarianship* 102899, 50(4), 1-10 <https://doi.org/10.1016/j.acalib.2024.102899>
- Coates, R. (2018). Meaningfulness in literary naming within the framework of The Pragmatic Theory of Proper hood (TPTP). *Onomastica Uralica* 14 (Volume 5 of the Proceedings of the 26th International Congress of Onomastic Sciences [ICOS-26], Debrecen, Hungary, (<http://mnytud.arts.unideb.hu/onomural/kotetek/ou14/15.pdf>))
- Contreras, J. & Mabel, G. A. (2005). Alimentación y Cultura. Perspectivas Antropológicas. *Investigaciones Sociales*, 11(19):387-392 <https://doi.org/10.15381/is.v11i19.8073>
- Dou, A., & Rim, K. (2025). Measuring Teachers’ AI Literacy through ChatGPT Prompt Engineering. *Atras Journal*, 6(2), 64-75. <https://doi.org/10.70091/Atras/vol06no02.4>

- Eyiuche, D. E., Polit, U. Y., Amaechi, I. M., & Okoye, H. C. (2020). Prevalence of bleeding symptoms in Nigerian population: A cross-sectional study of the three major ethnic groups, *Thrombosis Update*, 1, 1-6. <https://doi.org/10.1016/j.tru.2020.100016>.
- García, M. (2025). Food, religion, and gender in medieval and early modern Iberia: A theoretical and conceptual framework. *Food and Foodways*, 33(2), 150–170. <https://doi.org/10.1080/07409710.2025.2482485>
- Labadze, L., Grigolia, M. & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal Education Technology High Education* 20, 56. <https://doi.org/10.1186/s41239-023-00426-1>
- Monin, B., & Szczurek, L. M. (2014). Food cultures. In A. B. Cohen (Ed.), *Culture reexamined: Broadening our understanding of social and evolutionary influences* (pp. 155–190). American Psychological Association. <https://doi.org/10.1037/14274-009>
- Oluwadele, L. B., Adedapo, Adediran. & Sunkanmi, M. O. (2023). Ethnic and cultural diversity in Nigeria: A panacea for national development or decapitation? *Integrity Journal of Arts and Humanities* 4(4), 62-70. <https://doi.org/10.31248/IJAH2023.076>
- Raiaan, M. A. K, Md. Saddam, H. M., Kaniz, F., Nur, M. F., Sadman, S., Most, M. J. M., Jubaer A. Mohammed, E. A., & Sami A., (2024), A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges, *IEEE Access Multidisciplinary Rapid Review*, 12, 26839-26874. <https://doi.org/10.1109/ACCESS.2024.3365742>
- Reszegi, K. (2020) Culture – Identity. The Semantic Structure of Names. In: Caiazzo, Luisa–Coates, Richard–Azaryahu, Maoz (eds.), *Naming, Identity and Tourism*. Newcastle upon Tyne, Cambridge Scholars Publishing, 154 –167.
- Savannah, E. (2023). Food as Culture, Food Customs, and Cultural Identity. <https://slofoodbank.org/en/food-as-culture/>
- Seufert, T., Jänen, I., & Brünken, R. (2007). The impact of intrinsic cognitive load on the effectiveness of graphical help for coherence formation. *Computers in Human Behavior*, 23(3), 1055–1071. <https://doi.org/10.1016/j.chb.2006.10.002>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Ugoala, B. (2017). Communicative and Cohesive Devices in President Buhari's Directive to the Service Chiefs on the Urgency of Fighting Boko Haram Terrorism in Nigeria. *Journal of Multimedia Technology & Communication Studies*, 3 (1), 187 - 211.
- Völke, S. (2023). Named entities, naming practices, and their meanings – linguistic types and cultural contexts. *International Journal of Language and Culture*, 10(2), 151–175. <https://doi.org/10.1075/ijolc.00048.vol>
- Zhang, C. (2023). Exploration of Ethnic Dance Education and the Development Pathways for Dance Culture Inheritance, *Frontiers in Art Research*, 5(17), 67-71, <https://doi.org/10.25236/FAR.2023.051712>
- Zhang, R. Hong-Wei Li, Xin-Yuan Qian, Wen-Bo Jiang, Han-Xiao Chen, (2025). On large language models safety, security, and privacy: A survey. *Journal of Electronic Science and Technology*, 23(1), 1-16. <https://doi.org/10.1016/j.jnlest.2025.100301>
- Zhu L, Lai Y, Xie J, et al. (2025). Evaluating the potential risks of employing large language models in peer review. *Clin Transl Disc*. 5(4), 1-10 <https://doi.org/10.1002/ctd2.70067>