Restraining Cultural Stereotyping in Computational Linguistics through Computational Ethics

Ikechukwu Monday Osebor[✉]
Department of Philosophy, University of Delta
Carol C. Ohen[✉]
Department of Languages, University of Delta

(Received: 18th December 2024; Revised: 14th April 2025; Accepted: 21st April 2025)

Abstract

Computational linguistics is one of the achievements of science and technology in the 21st century. It has the ability to enable machines to understand, analyze and process human language with the aid of algorithms. Computational linguistics can in advertently perpetuate cultural stereotypes if not carefully considered in the development of language processing algorithms and models. It is important for computation al linguists to be aware of the potential biases in their work and strive to create inclusive and culturally sensitive tools and resources. Computational ethics, can promote diversity and inclusivity in computational linguistics, we can help mitigate the impact of cultural stereotypes and contribute to a more equitable and respectful society. Using a philosophical method of analysis, this study finds that cultural stereotypes can result from the misrepresentation and misunderstanding of cultural nuances, privacy violations, and many others. How can these moral issues be addressed? The study concludes that the implementation of computational ethics in the development of algorithms which recognize linguistic diversities can promote fairness, transparency, and respect for human rights.

Keywords: ethics, computational analysis, technology, linguistics, AI

✉ ikechukwu.osebor@unidel.edu.ng
Corresponding Author*
✉ carol.ohen@unidel.edu.ng

Introduction

Technology evolves and influences various aspects of human life, including communication, privacy, bias, and employment (Osebor, & Onyekpe, 2024). Computational linguistics (CL), the interdisciplinary field that uses computational methods to process and analyze natural language, has profound moral implications (Frankenreiter & Livermore, 2020). Understanding these moral dimensions is essential for ensuring the development and application of computational linguistics in language analysis. One of the primary moral implications of computational linguistics is cultural stereotyping that is tied to the marginalization of less commonly spoken languages in computational analysis (Mendelsohnet al., 2020).

In many Nigerian schools, cultural stereotyping exists, with the teaching and learning of major ethnic languages. This will lead to a decline in linguistic diversity and the extinction of minority languages. The forms of cultural stereotyping include surveillance or the increase of corporate involvement in deploying sophisticated language processing tools to monitor conversations, predict behaviors, and profile individuals without their knowledge (D'Arcy & Bender, 2023).

In Nigeria, government allocates resources for the development of computational tools (e.g., translation software and speech recognition). These resources are often allocated to widely spoken languages, such as Igbo, Hausa and Yoruba leaving less common languages without adequate technological support (Weidingeret al., 2021). This poses significant ethical challenges regarding the balance between security and individual freedom. The moral responsibility lies in creating regulations and frameworks that protect individuals from unjust surveillance while allowing beneficial uses of these technologies (Irabor, & Monday (2021; Vidgen, & Derczynski, 2020). This is problematic, especially the application of automated content without moderation, where biased algorithms can unfairly target certain cultural or ethnic groups, leading to discrimination and to the perpetuation and amplification of bias and uneven distribution of technology (Vidgen, & Derczynski, 2020 ; Hudley, et al., 2020). This can exacerbate existing inequalities, as marginalized groups may not benefit from advancements in language technology. Machine learning models can also be manipulated to increase cultural stereotypes. This includes the use of trained language models to perpetrate gender, racial, or cultural biases, which can reproduce these biases in the outputs (Hovy, & Prabhumoye, 2021). This can lead to unfair or discriminatory outcomes, such as biased hiring algorithms or biased content moderation systems. This study argues that ethical computational linguistics must strive for fairness and inclusivity (Chandrabose & Chakravarthi, 2021).

Cultural stereotyping has led to the emergence of a deepfake generation, which can create realistic but fake audio and video content that poses risks concerning misinformation, fraud, and defamation (Weidinger, et al., 2021). These tools can be used to manipulate public opinion, interfere with elections, and perpetrate scams, leading to severe societal harm (Irabor, & Osebor, 2022). Despite these challenges, computational linguistics plays a pivotal role in shaping the future landscape of technology-driven communication and understanding. Advancements in computational linguistics have revolutionized natural language processing, enabling machines to comprehend, generate, and interact with human language at unprecedented and increasingly interconnected levels. The ability to develop multilingual and cross-lingual models fosters inclusivity and facilitates seamless communication among diverse linguistic communities (Norpolatova, 2023).

Researchers and AI developers have a moral duty to recognize and mitigate biases in their models (CL) by curating balanced and representative datasets, developing techniques to detect and correct bias, and continuously evaluating the impact of these technologies on different demographic groups (Kimera, & Choi, 2024). For example, when training a model for machine translation, it's essential to ensure that the corpus includes diverse linguistic variations and cultural expressions. Ensuring fairness is not only a technical challenge but a moral imperative to promote equality and prevent harm (Karpouzis, 2024). The moral issue here centers on the societal responsibility to manage cultural stereotyping in work places. It has led to employment displacement in Nigeria (Köhler, et al., 2019). Additionally, companies benefiting from cultural stereotyping should consider the broader societal impact and contribute to the social safety net, ensuring a just transition for those affected by technological unemployment (Köhler, et al., 2019: Weidinger, et al., 2021). The aim of this study is to recommend the role of computational ethics in addressing cultural stereotyping in computational linguistics. The study suggests that policymakers around the world should implement computational ethics in the development of AI algorithms to enhance cultural diversity in language analysis.

Methodology

The study presents an in-depth analysis to gain an understanding of computational linguistics and its moral implications. It addresses the ethical issues of computational linguistics and the study suggests the adoption and implementation of computational ethics by policy makers to enhance justice, fairness, equality, and privacy protection in computational linguistics. To achieve the aforementioned purposes, the study employs a philosophical method of analysis. Conceptual analysis involves breaking down complex concepts into simpler components to better understand their meaning and interrelations. This method will help clarify the use of terms and concepts such as artificial intelligence, linguistics, computational ethics, and many others in different contexts, ensuring precise communication and understanding.

Computational linguistics

Computational linguistics emerged in the mid-20th century, motivated by the practical needs of machine translation during the Cold War era (Gaspari, 2024). Early pioneers, such as Warren Weaver and Noam Chomsky, laid the groundwork by proposing models for how languages could be systematically analyzed and processed by machines (Léon, 2021). Chomsky's generative grammar provided a formal framework that heavily influenced early computational approaches, emphasizing the rule-based nature of syntax (Yang et al., 2017).One of the core philosophical questions in computational linguistics is the nature of language itself in cultural understanding. Is language a formal system that can be entirely captured by rules and algorithms, or is it inherently fuzzy and context-dependent? Early rule-based approaches treated language as a set of formal, syntactic rules that could be programmed into a computer (Sidorov, 2013). However, this perspective has been challenged by more recent probabilistic and data-driven models, which suggest that language is better understood through statistical patterns and large datasets (Himanen et al., 2019).

The debate between formalism and empiricism in computational linguistics mirrors broader philosophical debates. Formalists argue that the structure of language can be precisely described by formal grammars and logical systems (Newmeyer, 2000). Empiricists, on the other hand, argue that language understanding arises from exposure to large amounts of data and that statistical methods are more effective in capturing the nuances of natural

language (Chater et al., 2015). Jackendoff developed the theory of conceptual semantics, which seeks to explain the relationship between linguistic expressions and their meanings (Jackendoff, 1988). He argues that understanding language requires understanding the conceptual structures that influence computational approaches to language that seek to integrate multiple levels of linguistic analysis (Jackendoff, 1996). Alan Turing's work on computation and artificial intelligence provided the necessary theoretical background for understanding and creating language-processing algorithms while the Chinese Room argument and emphasis on intentionality challenge the field to consider the depth of semantic understanding beyond mere symbol manipulation (Yang, 2012: Searle, 1982).). Ferdinand de Saussure's structuralism emphasizes the relational aspect of language elements, crucial for developing computational models that deal with syntax and semantics (Seuren, 2015).

Jürgen Habermas's focus on communicative action and universal pragmatics underscores the importance of understanding language in context, relevant for dialogue systems and discourse analysis (Cooke, 1994). Together, these thinkers have shaped the field of computational linguistics, driving forward the quest to understand and process human language through computational means. The field of computational linguistics, which involves the use of computers to process and analyze human language, raises several moral and ethical implications. These issues, stemming from the potential consequences of language technology, have led to cultural stereotyping. Addressing these implications requires employing a computational ethical approach, involving policy, technology design, and ethical standards.

The Moral Implications of Computational Linguistics

Computational language models can perpetuate harmful stereotypes and biases present in training data, leading to unfair outcomes, especially related to race, gender, or nationality (Gallegos et al., 2024). This presents moral concerns about privacy violations, the potential for unwarranted surveillance, unethical language use, demographic bias in language use, and political propaganda incorporated in the datasets, among others (Van den Hoven, 2017). Computational algorithms, such as chatbots and recommendation engines, could manipulate users' choices or beliefs without their full awareness, undermining personal autonomy (Prunkl, 2024). Issues related to cultural sensitivity and representation may lead to misrepresentation or misunderstandings and the marginalizing of certain linguistic and cultural groups. The implication here is that as AI becomes more autonomous, it becomes difficult to determine who is responsible for mistakes in cultural stereotyping, raising concerns about accountability for harmful or biased outputs. To address these issues, it is crucial to adopt computational ethics in the development of AI algorithms to ensure inclusivity in computational linguistics. This will help promote ethical guidelines of fairness, transparency, cultural sensitivity, privacy protection, and inclusivity in the development and use of computational linguistics technologies.

Practical Suggestions for Computational Ethics

Given the rapid pace of technological advancement, it is crucial to consider computational ethics in the development of AI algorithms to avoid unintended consequences in language analysis (Anindya & Lusiyani, 2024). Technologies should be developed and deployed responsibly to prevent the marginalization of the minority languages. The deployment of computational ethics will address moral imperatives such as fairness, transparency, and inclusivity in the field of computational linguistics to harness the power of language technology for the benefit of society while minimizing harm, recognizing that

different ethnic groups have their languages (Monday, 2020c; Monday, 2020d). In an automatic text generation model, the computational ethical framework could mandate AI developers to disclose the origin of training data and algorithms to ensure they are free from biased or stereotypical content. This would make it easier to track and address problematic outputs through interdisciplinary dialogue that works together to navigate the ethical landscape of computational linguistics, ensuring that its advancements contribute positively to the human condition.

Computational ethics, a subfield of artificial intelligence (AI) addresses the implementation of ethical decision-making in machines language. As technology systems become more integrated in language analysis, the ability to make decisions that align with human values becomes crucial (Gabriel, 2020: Monday, 2020a). It seeks to design, implement, and evaluate algorithms that can make ethical decisions or assist humans in making such decisions (Duan, et al., 2019). Computational ethics is concerned with the principles that govern technology language analysis, focusing on concepts like justice, rights, and duties to stakeholders (Awad,et al., 2022; Monday,2020b). This includes implementing safeguards, promoting digital literacy to help people recognize and resist manipulated content, and working with policymakers to create laws that deter malicious use of these technologies (Hua, & Jiang, 2024).

Computational ethics has the potential to significantly enhance accessibility and inclusivity (Dowell et al., 2019). For instance, speech recognition and text-to-speech technologies can assist people with disabilities, such as those who are visually impaired or have speech impairments (Zaki, & Ahmed, 2024). Ensuring that these technologies are truly inclusive and accessible to all is a moral obligation. Developers should focus on creating tools that accommodate diverse languages, dialects, and accents. This involves recognizing and addressing the needs of minority language speakers and ensuring that technological advancements do not further marginalize already disadvantaged groups. For example, sentiment analysis data or models that wrongly associate certain words with negative sentiments due to cultural stereotypes should be adjusted by the algorithm to treat all cultural contexts equally, ensuring that no group is unfairly represented. Computational ethics can be a powerful force for social good, promoting inclusivity and equal opportunities for all (Warschauer, 2004). This involves conducting research transparently, sharing findings openly, and engaging with a diverse range of stakeholders, including ethicists, sociologists, and representatives from affected communities (Pratt, 2019). Moreover, researchers should be mindful of the long-term impacts of their work. The precautionary principle, which advocates for caution in the face of uncertainty, is particularly relevant in computational linguistics. It ensures the use of technologies in language to predefined ethical rules, regardless of the outcome. It bases ethical decisions on the outcomes, aiming to maximize overall good or minimize harm (Anderson, 2008). This approach can be effective but is limited by the quality and bias of the training data. Systems learn ethical behavior through trial and error, receiving feedback from their environment. For example, in virtual assistants or Chabot systems, where cultural context heavily influences language use, ethical guidelines could ensure that the assistant avoids assuming that stereotypes about how people from different cultures speak or behave. For instance, it would avoid assuming all speakers of a particular language have the same mannerisms or preferences. The adoption of computational ethics could handle complex scenarios but may require extensive training. It is a rule-based and learning-based method that can leverage the strengths and make clear-cut decisions in ambiguous situations (Anderson, 2008).

Ethics can be subjective and culturally dependent, making it challenging to define universal norms for the use of technologies (Osebor, 2024a). Implementing transparency and accountability mechanisms can help ensure that decisions made by algorithms are explainable and can be reviewed by humans. Ensuring that a technology system's ethical framework is appropriate for diverse contexts is a significant challenge, as AI systems can inherit biases present in their training data, leading to unfair or discriminatory outcomes (Schwartz, et al., 2022). Ensuring fairness and mitigating bias is critical, especially in computational linguistics. Computational ethics must be implemented in data analysis. The implementation of ethical guidelines specific to CL can help address cultural stereotyping (Osebor, 2024b).

Conclusion

This study aims to investigate the moral implications of computational linguistics. The moral implications of computational linguistics are complex and multifaceted, requiring coordinated efforts across different domains. Implementing computational ethics would involve robust privacy protections, addressing bias and fairness, mitigating its impact on employment, combating misinformation, and establishing strong ethical frameworks. It would help to harness the benefits of language technologies while minimizing their potential harms. This objective seek to create a balanced that respects individual rights and promotes societal well-being in computational linguistics.

REFERENCES

Anderson, S. L. (2008). Asimov's "three laws of robotics" and machine metaethics. *Ai & Society, 22*, 477-493.

Anindya, W. D. & Lusiyani, R. (2024). The use of ICT in implementation of autonomous learning: students' perception. *Social Sciences, Humanities and Education Journal (SHE Journal), 5*(1), 15 – 22.

Awad, E., Levine, S., Anderson, M., Anderson, S. L., Conitzer, V., Crockett, M. J., & Tenenbaum, J. B. (2022). Computational ethics. *Trends in Cognitive Sciences, 26*(5), 388-405.

Chandrabose, A., & Chakravarthi, B. R. (2021, April). An overview of fairness in data–illuminating the bias in data pipeline. In Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion (pp. 34-45).

Chater, N., Clark, A., Goldsmith, J. A., & Perfors, A. (2015). Empiricism and language learnability. OUP Oxford.

Cooke, M. (1994). Language and reason: A study of Habermas's pragmatics. MIT Press.

D'Arcy, A., & Bender, E. M. (2023). Ethics in linguistics. *Annual Review of Linguistics, 9*, 49-69.

Dowell, N., Lin, Y., Godfrey, A., & Brooks, C. (2019). Promoting inclusivity through time-dynamic discourse analysis in digitally-mediated collaborative learning. In Artificial Intelligence in Education: 20th International Conference, AIED 2019, Chicago, IL, USA, June 25-29, 2019, Proceedings, Part I 20 (pp. 207-219). Springer International Publishing.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International journal of information management, 48*, 63-71.

Frankenreiter, J., & Livermore, M. A. (2020). Computational methods in legal analysis. *Annual Review of Law and Social Science, 16*, 39-57.

Gabriel, I. (2020). Artificial intelligence, values, and alignment. *Minds and machines, 30*(3), 411-437.

Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M. M., Kim, S., Dernoncourt, F., & Ahmed, N. K. (2024). Bias and fairness in large language models: A survey. *Computational Linguistics, 50*(3), 1097-1179.

Gaspari, F. (2024). The History of Translation Technologies. In The Routledge Handbook of the History of Translation Studies (pp. 324-338). Routledge.

Himanen, L., Geurts, A., Foster, A. S., & Rinke, P. (2019). Data-driven materials science: status, challenges, and perspectives. *Advanced Science, 6*(21), 1900808.

Hovy, D., & Prabhumoye, S. (2021). Five sources of bias in natural language processing. *Language and linguistics compass, 15*(8), e12432.

Hua, S., Jin, S., & Jiang, S. (2024). The limitations and ethical considerations of ChatGPT. *Data Intelligence, 6*(1), 201-239.

Hudley, A. H. C., Mallinson, C., & Bucholtz, M. (2020). Toward racial justice in linguistics: Interdisciplinary insights into theorizing race in the discipline and diversifying the profession. *Language, 96*(4), e200-e235.

Irabor, B. P. & Monday, O. I. (2021). A Heideggerian Reflection on the Psycho-Moral Consequences of Cyber-Bullying. *Journal of Ethics and Legal Technologies, 3*(2), 139-150.

Irabor, B. P., & Osebor, I. M. (2022) The Moral Implications of Cyberbullying Vis-À-Vis Parental Concerns. *Abraka Humanities Review, 12*(1), 162 – 169.

Jackendoff, R. (1996). Conceptual semantics and cognitive linguistics. https://www.degruyterbrill.com/document/doi/10.1515/cogl.1996.7.1.93/html?lang=en

Karpouzis, K. (2024). Plato's Shadows in the Digital Cave: Controlling Cultural Bias in Generative AI. *Electronics, 13*(8), 1457.

Kimera, R., Kim, Y. S., & Choi, H. (2024). Advancing AI with Integrity: Ethical Challenges and Solutions in Neural Machine Translation. arXiv preprint arXiv:2404.01070.

Köhler, J., Geels, F. W., Kern, F., Markard, J., Onsongo, E., Wieczorek, A., ...& Wells, P. (2019). An agenda for sustainability transitions research: State of the art and future directions. *Environmental innovation and societal transitions, 31*, 1-32.

Léon, J. (2021). Automating linguistics. Springer.

Mendelsohn, J., Tsvetkov, Y., & Jurafsky, D. (2020). A framework for the computational linguistic analysis of dehumanization. *Frontiers in artificial intelligence, 3*, 55.

Monday, O. I. (2020b). Is altruism always sufficient for organ donation? vroom's expectancy theory, for expanding the organ donor pool. *Saudi Journal of Kidney Diseases and Transplantation 31*(2), 503-507. https://doi.org/10.4103/1319-2442.284026.

Monday, O. I. (2020c). The Ethics of Organ Farming. *Türkiye Biyoetik Dergisi 7*(4), 183-186. https://philpapers.org/rec/MONTEO-29.

Monday, O. I. (2020d). The Relevance of deep ecological principles in Aquatic Crisis: A philosophical Analysis. *Bangladesh Journal of Bioethics 11*(2), 42-48. https://doi.org/10.3329/bioethics.v11i2.50392.

Monday, O.I. (2020a). Ethics of cognitive restructuring: A rehabilitation of rape victims and offenders. *Voice in Bioethics, 6*. https://doi.org/10.7916/vib.v6i.7223

Newmeyer, F. J. (2000). Language form and language function. MIT press.

Norpolatova, N. (2023). The Importance of Advancing Computational Linguistics, *Fundamental And Applied Research Conference Proceedings Volume.11*(5) (pp.128-130). https://orientalpublication.com/index.php/iscrc/article/view/1289

Osebor, I. M, & Benson P. I. (2023). The Role of Metaphysics in The Development of Native-Centric Education. *Oracle of Wisdom Journal of Philosophy and Public Affairs, 7*(3), 78-85. https://acjol.org/index.php/owijoppa/article/view/4199/4114.

Osebor, I. M. (2023a). Is Pet-Keeping A Violation of Animal Rights? A Consequentialist Standpoint. *AQUINO| Journal of Philosophy 3*(2), 129-137. https://acjol.org/index.php/aquino/article/view/4243/4155.

Osebor, I. M. (2024b). The Role of Structural Ethics in Restraining 'Japa' among Clinicians. Ramon Llull *Journal of Applied Ethics, 15*. https://doi.org/10.60940/rljaev1n15Id428072.

Osebor, I. M., & IwegbuOnyekpe, H. (2024). Interdisciplinary Dialogue: A Moral Standard for The Development of Transhumanism. *Oracle of Wisdom Journal of Philosophy and Public Affairs (OWIJOPPA), 8*(1).

Pratt, B. (2019). Inclusion of marginalized groups and communities in global health research priority-setting. *Journal of Empirical Research on Human Research Ethics, 14*(2), 169-181.

Prunkl, C. (2024). Human autonomy at risk? An analysis of the challenges from AI. *Minds and Machines, 34*(3), 26.

Sayers, D., Sousa-Silva, R., Höhn, S., Ahmedi, L., Allkivi-Metsoja, K., Anastasiou, D., ...&Yayilgan, S. Y. (2021). The Dawn of the Human-Machine Era: A forecast of new and emerging language technologies.

Schwartz, R., Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). Towards a standard for identifying and managing bias in artificial intelligence (Vol. 3, p. 00).US Department of Commerce, National Institute of Standards and Technology.

Searle, J. R. (1982). The Chinese room revisited. *Behavioral and brain sciences, 5*(2), 345-348.

Seuren, P. A. (2015). Prestructuralist and structuralist approaches to syntax. In Syntax—theory and analysis: An international handbook (pp. 134-157). Mouton de Gruyter.

Sidorov, G. (2013). Syntactic dependency based n-grams in rule based automatic English as second language grammar correction. Int. J. Comput. *Linguistics Appl., 4*(2), 169-188.

Van den Hoven, J. (2017). Privacy and the varieties of informational wrongdoing. In Computer ethics (pp. 317-330). Routledge.

Vidgen, B., & Derczynski, L. (2020). Directions in abusive language training data, a systematic review: Garbage in, garbage out. *Plos one, 15*(12), e0243300.

Warschauer, M. (2004). Technology and social inclusion: Rethinking the digital divide. MIT press.

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P. S., & Gabriel, I. (2021). Ethical and social risks of harm from language models. arXiv preprint arXiv:2112.04359.

Yang, C., Crain, S., Berwick, R. C., Chomsky, N., & Bolhuis, J. J. (2017). The growth of language: Universal Grammar, experience, and principles of computation. *Neuroscience &Biobehavioral Reviews, 81*, 103-119.

Yang, X. S. (Ed.). (2012). Artificial intelligence, evolutionary computing and metaheuristics: in the footsteps of Alan Turing (Vol. 427). Springer.

Zaki, M. Z., & Ahmed, U. (2024). Bridging Linguistic Divides: The Impact of AI-powered Translation Systems on Communication Equity and Inclusion. *Journal of Translation and Language Studies, 5*(2), 20-30.