

Towards Figurative Language Generation in Afrikaans

Imke van Heerden

Dept. of Comparative Literature
College of Social Sciences and Humanities
Koç University, Istanbul, Turkey
ivanheerden@ku.edu.tr

Anil Bas

Dept. of Computer Engineering
Faculty of Technology
Marmara University, Istanbul, Turkey
anil.bas@marmara.edu.tr

Introduction This paper presents an LSTM-based approach to figurative language generation, which is an important step towards creative text generation in Afrikaans. Due to scarcity of resources (in comparison to resource-rich languages), we train the proposed network on a single literary novel. This follows the same approach as [Van Heerden and Bas \(2021\)](#), however we explicitly focus and expand on fully automatic text generation, centring on figurative language in particular. The proposed model generates phrases that contain compellingly novel figures of speech such as metaphor, simile and personification.

Afrikaans NLP In [V et al. \(2020a,b\)](#) and [Adelani et al. \(2021\)](#), Masakhane shows that NLP research in African languages is under-represented. This is true for Afrikaans as well. Although some studies ([Van Zaanen and Van Huyssteen, 2003](#); [Sanby et al., 2016](#); [Ziering and Van der Plas, 2016](#); [Dirix et al., 2017](#)) and datasets ([Eiselen and Puttkammer, 2014](#); [Augustinus et al., 2016](#); [Roux, 2016](#)) are available, text generation in Afrikaans is very limited.

Related Work Studies in figurative text generation include simile ([Harmon, 2015](#); [Chakrabarty et al., 2020](#)), slogan ([Alnajjar et al., 2018](#)) and metaphor ([Gero and Chilton, 2019](#); [Brooks and Youssef, 2020](#); [Chakrabarty et al., 2021](#)) generation. These computational approaches involve style transfer and word masking but also, more traditionally, non-computational theories of metaphor creation, e.g. the tenor-vehicle model ([Richards, 1936](#)). Moreover, related work on figurative language (including the aforementioned) tends to focus on English and other resource-rich languages, using knowledge bases, graphs, pretrained networks and datasets.

Model We use a two-layer vanilla LSTM architecture ([Hochreiter and Schmidhuber, 1997](#)). The proposed model consists of two LSTM layers with

dropout layers, a fully connected layer and a softmax layer. The network trained on the South African novel *Die Biblioteek aan die Einde van die Wêreld* (*The Library at the End of the World*) ([Van Heerden, 2019](#)). The book consists of approximately 210K words and is well suited to the task given its near-poetic use of figurative language and imagery.

Approach Instead of following relational rules and constraining the model, we adopt Black’s interaction theory of metaphor ([1962](#)). This is relevant to the study given its emphasis on the generative function of figurative language. In this view, meaning is not static; metaphor does not draw on pre-existing likeness but instead *creates* new, often surprising, likeness between concepts ([Indurkha, 1992](#); [Veale et al., 2016](#)). Through metaphor, dissimilar domains *interact*, giving rise to novel connections and perspectives ([Way, 1991](#)). It is challenging to measure originality in generated text ([Klebanov and Madnani, 2020](#)). However, if figurative meaning does in fact rely on “mismatch” between domains ([Fogelin, 2011](#)), one might argue: the greater the mismatch, the greater the novelty.

Results Table 1 provides examples of generated phrases and sentences containing figures of speech such as metaphor and personification. (Note that punctuation and capitalisation were removed in some instances.) We did not set out to generate figurative language per se, meaning that there are no specific components to enforce rules or constraints on the model. Similar to the trained data, the network outputs unique descriptive formulations. Given our interest in creative text generation, we believe that the success of the results is not determined by the amount of similarities shared between referents, as explained by [Giles et al. \(1991\)](#). Instead, we prioritise unexpected associations between disparate concepts, e.g. “democracy” as a “roof garden” (See Table 1, line 19).

Original (Afrikaans)	Translation (English)
<p>ons biblioteek by die werkwoord geakaap die wêreld sê ek met boeke wêreldletterkunde in armoede saggies soos 'n spokerigheid in die vlug van papier sy vingers draai om haar gevoel ek het 'n gloeiende noordgrens woede is jou mond brand my in die oggendlug die wind stoppelbaard vorentoe sy kyk verras op, sy oë verlate verandering speel as foto's van die wind die petrolbomme wat nie vertel nie sukkel is hulle kuns onbeskermd skittering in die woord my rug se wit greep ek is geld want niks kan bloei nie aarselend weerskante van die staar demokrasie was 'n daktuin begin die sonsopkoms voor die dak van my gesig jou uitgespoel is 'n onderstebo losgewoel gesprekke vir die oomblik skoongeskraap bleek lewe in hierdie nuwe hande waar ek algoritmië kuier</p>	<p>our library hijacked at the verb the world I say with books world literature in poverty softly like a ghostliness in the flight of paper his fingers wrap around her feeling I have a glowing northern border anger is your mouth burn me in the morning air the wind stubbles forth she looks up in surprise, his eyes deserted change plays as photos of the wind the petrol bombs that do not tell struggling is their art unprotected brightness in the word my back's white grip I am money because nothing can bleed hesitant on either side of the stare democracy was a roof garden begins the sunrise before the roof of my face your rinsed-out is an upside-down tossed-loose conversations momentarily clean-scraped pale life in these new hands where I socialise algorithmically</p>

Table 1: Example results of figurative language generation. The translations might not do justice to the original given the distinctiveness of some of the formulations as well as the compounding nature of Afrikaans.

Discussion Defined as “language that is more expressive and/or poetic than referential in its linguistic function” (Chandler and Munday, 2011), figurative language comprises metaphor, simile, personification and various other figures of speech. It is a vital resource of creative writing (Baldick, 1996) often associated with originality. In fact, original metaphor is considered “the controlling element in all creative language” (Newmark, 1988). Because it is open to interpretation, metaphorical language invites the reader to participate in the process of meaning-making (White, 1996), thus facilitating connection between reader and text (Cohen, 1978). Compared to metaphors in everyday speech, unusual and unexpected figures stand out (Kovecses, 2010) and, in a literary context, capture readers’ attention (Steen, 1994). Furthermore, Gibbs et al. suggest that original metaphors “communicate more emotional intensity than conventional metaphor” (2002). Therefore, novel metaphor is not only related to creativity but also to emotion (Fainsilber and Ortony, 1987; Fussell and Moss, 1998). Future work could explore how this may lend emotional depth to generated text.

Acknowledgments

This paper has been produced benefiting from the 2232 International Fellowship for Outstanding Researchers Program of TÜBİTAK (Project No: 118C285). However, the entire responsibility of the paper belongs to the owner of the paper. The financial support received from TÜBİTAK does not mean that the content of the publication is approved in a scientific sense by TÜBİTAK.

The authors are grateful to Etienne van Heerden for sharing the original manuscript of his novel.

References

- David Ifeoluwa Adelani, Jade Abbott, Graham Neubig, et al. 2021. MasakhaNER: Named entity recognition for African languages. In *Proc EACL Workshop on AfricaNLP*.
- Khalid Alnajjar, Hadytullah Kundi, Hannu Toivonen, et al. 2018. “Talent, skill and support”: A method for automatic creation of slogans. In *Proc ICCS*.
- Liesbeth Augustinus, Peter Dirix, Daniel Van Niekerk, et al. 2016. AfriBooms: An online treebank for Afrikaans. In *Proc LREC*, pages 677–682.

- Chris Baldick. 1996. *The Concise Oxford Dictionary of Literary Terms*. Oxford University Press.
- Max Black. 1962. *Models and Metaphors: Studies in Language and Philosophy*. Cornell University Press.
- Jennifer Brooks and Abdou Youssef. 2020. Discriminative pattern mining for natural language metaphor generation. In *Proc Big Data*, pages 4276–4283.
- Tuhin Chakrabarty, Smaranda Muresan, and Nanyun Peng. 2020. Generating similes effortlessly like a pro: A style transfer approach for simile generation. In *Proc EMNLP*, pages 6455–6469.
- Tuhin Chakrabarty, Xurui Zhang, Smaranda Muresan, and Nanyun Peng. 2021. Mermaid: Metaphor generation with symbolism and discriminative decoding. In *Proc NAACL*.
- Daniel Chandler and Rod Munday. 2011. *A Dictionary of Media and Communication*. Oxford University Press.
- Ted Cohen. 1978. Metaphor and the cultivation of intimacy. *Critical Inquiry*, 5(1):3–12.
- Peter Dirix, Liesbeth Augustinus, Daniel Van Niekerk, and Frank Van Eynde. 2017. Universal dependencies for Afrikaans. In *Proc NoDaLiDa*, pages 38–47.
- Roald Eiselen and Martin Puttkammer. 2014. Developing text resources for ten South African languages. In *Proc LREC*, pages 3698–3703.
- Lynn Fainsilber and Andrew Ortony. 1987. Metaphorical uses of language in the expression of emotions. *Metaphor and Symbol*, 2(4):239–250.
- Robert J Fogelin. 2011. *Figuratively Speaking: Revised Edition*. Oxford University Press.
- ∇ et al. 2020a. Masakhane–machine translation for Africa. In *Proc ICLR Workshop*.
- ∇ et al. 2020b. Participatory research for low-resourced machine translation: A case study in African languages. In *Proc EMNLP*, pages 2144–2160.
- Susan R Fussell and Mallie M Moss. 1998. Figurative language in emotional communication. *Social and Cognitive Approaches to Interpersonal Communication*, pages 113–141.
- Katy Ilonka Gero and Lydia B Chilton. 2019. Metaphoria: An algorithmic companion for metaphor creation. In *Proc CHI*, pages 1–12.
- Raymond W Gibbs, John S Leggett, and Elizabeth A Turner. 2002. What’s special about figurative language in emotional communication. *The Verbal Communication of Emotions: Interdisciplinary Perspectives*, pages 125–149.
- Andrea Giles, Joseph F Kess, and Chiharu Uda. 1991. Metaphor as the creative origin of lexical ambiguity in English and Japanese. *Working Papers of the Linguistics Circle*, 10(1):49–62.
- Sarah Harmon. 2015. Figure8: A novel system for generating and evaluating figurative language. In *Proc ICCD*, pages 71–77.
- Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.
- Bipin Indurkha. 1992. *Metaphor and Cognition: An Interactionist Approach*. Kluwer Academic.
- Beata Beigman Klebanov and Nitin Madhani. 2020. Automated evaluation of writing – 50 years and counting. In *Proc ACL*, pages 7796–7810.
- Zoltan Kovecses. 2010. *Metaphor: A Practical Introduction*. Oxford University Press.
- Peter Newmark. 1988. *A Textbook of Translation*, volume 66. Prentice Hall.
- Ivor Armstrong Richards. 1936. *The Philosophy of Rhetoric*. Oxford University Press.
- Justus Roux. 2016. South African National Centre for Digital Language Resources. In *Proc LREC*, pages 2467–2470.
- Lauren Sanby, Ion Todd, and Maria C Keet. 2016. Comparing the template-based approach to GF: the case of Afrikaans. In *Proc WebNLG*, pages 50–53.
- Gerard Steen. 1994. *Understanding Metaphor in Literature: An Empirical Approach*. Longman.
- Etienne Van Heerden. 2019. *Die Biblioteek aan die Einde van die Wêreld*. NB-Uitgewers.
- Imke Van Heerden and Anil Bas. 2021. AfriKI: Machine-in-the-loop Afrikaans poetry generation. In *Proc EACL Workshop on Bridging Human–Computer Interaction and Natural Language Processing*, pages 74–80.
- Menno Van Zaanen and Gerhard Van Huyssteen. 2003. Improving a spelling checker for Afrikaans. In *Proc CLIN*, pages 143–156.
- Tony Veale, Ekaterina Shutova, and Beata Beigman Klebanov. 2016. *Metaphor: A Computational Perspective*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.
- E Cornell Way. 1991. *Knowledge Representation and Metaphor*. Springer Science & Business Media.
- Roger M White. 1996. *The Structure of Metaphor: The Way the Language of Metaphor Works*. Wiley.
- Patrick Ziering and Lonneke Van der Plas. 2016. Towards unsupervised and language-independent compound splitting using inflectional morphological transformations. In *Proc NAACL*, pages 644–653.