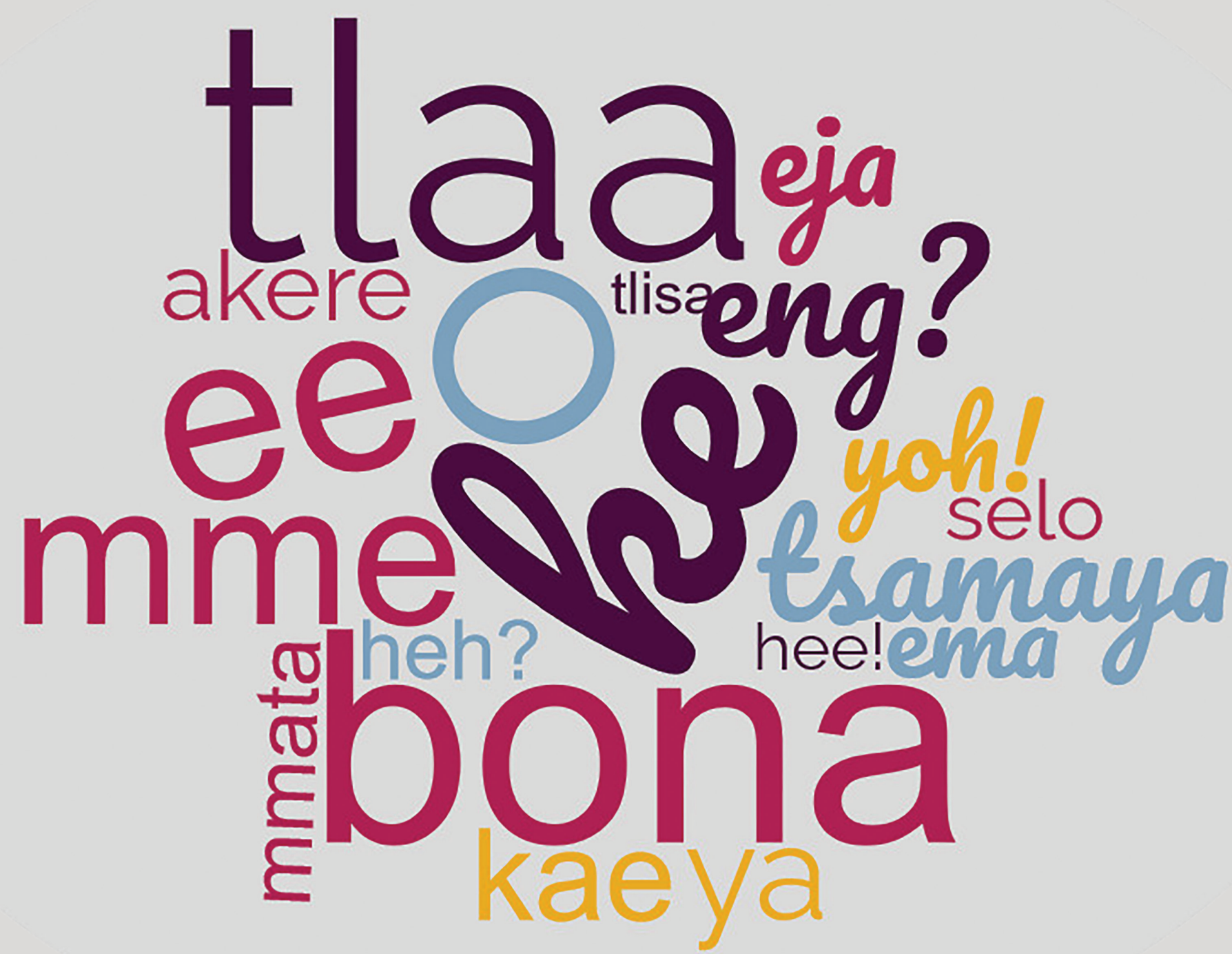# Determining the core vocabulary of Setswana-speaking Grade R learners as used during school activities

Morwesi G. Mogatusi (MA AAC) Supervisor: Prof Kerstin Tönsing
Co-supervisor: Dr Refilwe Morwane

Main contribution

- **A Setswana core vocabulary list was established as a resource to support vocabulary selection for AAC systems that can be used by children from Setswana language backgrounds who need AAC.**
- **This list can help service providers to develop linguistically robust and culturally relevant AAC systems for this population.**

Introduction

- Core vocabulary lists are often used to select vocabulary for picture-based AAC systems.
- A core vocabulary comprises of the words most frequently and commonly used by speaking individuals.
- An estimated 18 000 children with limited speech in South Africa could benefit from a Setswana AAC system.
- This study aimed to develop a core vocabulary list in Setswana as a vocabulary resource for Setswana AAC systems.

Methods

- A descriptive observational design was employed for this study.
- Six preschoolers from three different schools participated.
- Participants were fitted with body worn audio recording devices and their spontaneous speech was recorded during preschool activities over a period of 3 days.
- The spontaneous speech recordings of these learners were transcribed, coded and analysed using Microsoft Word™ and Excel™.

Results

- A total of 1 112 different words were found.
- In order for a word to be classified as core, it needed to:
  - Occur in the language sample once or more per 2 000 words
  - Occur in the samples of at least three of the six participants.
- There were 249 core words and 863 fringe words in the sample. See Figure 1.
- Overall, core words were used 86% of the time while fringe words were used only 14% of the time. See Figure 2.
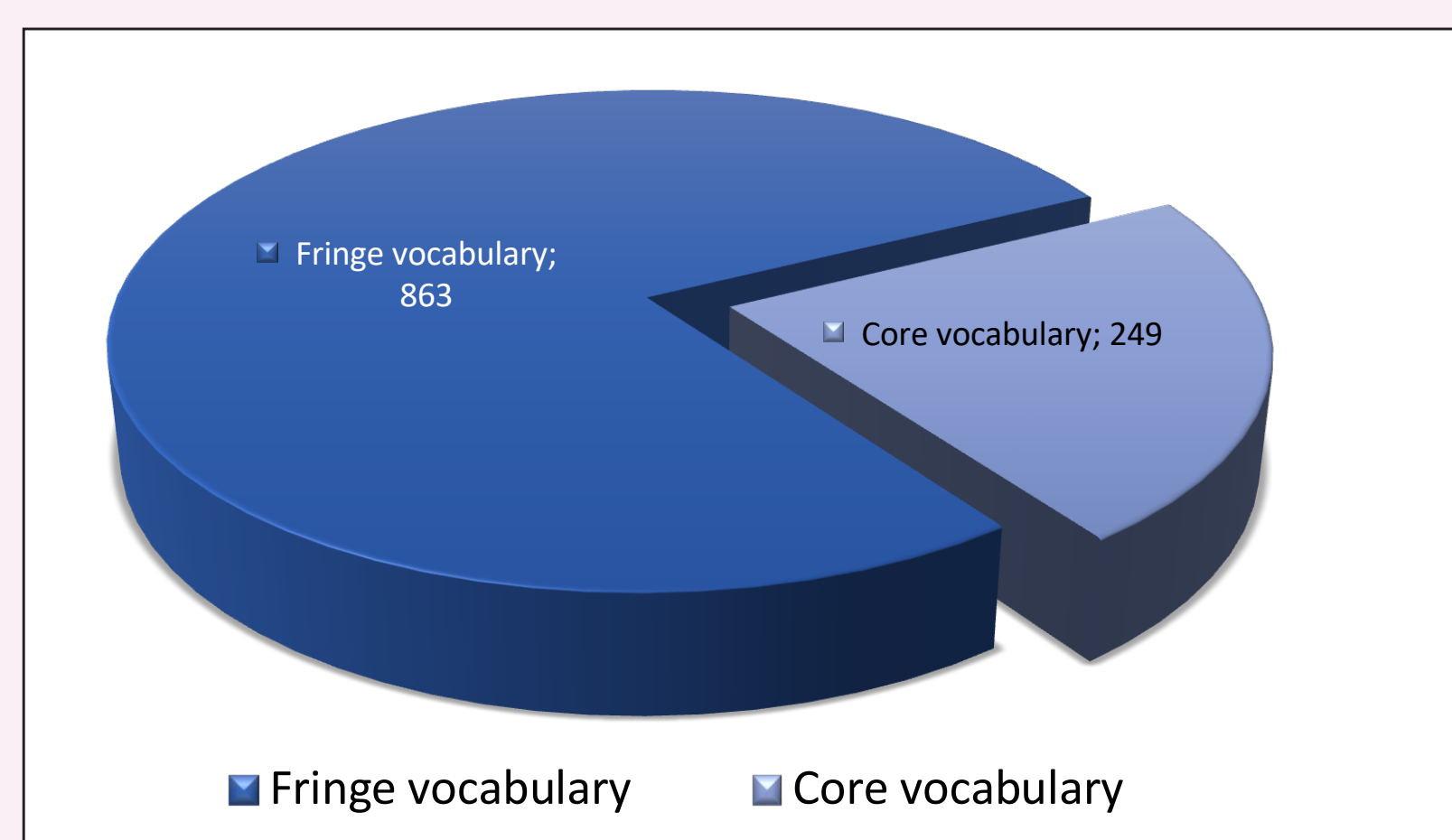

Figure 1. The number of different words designated as core versus fringe vocabulary
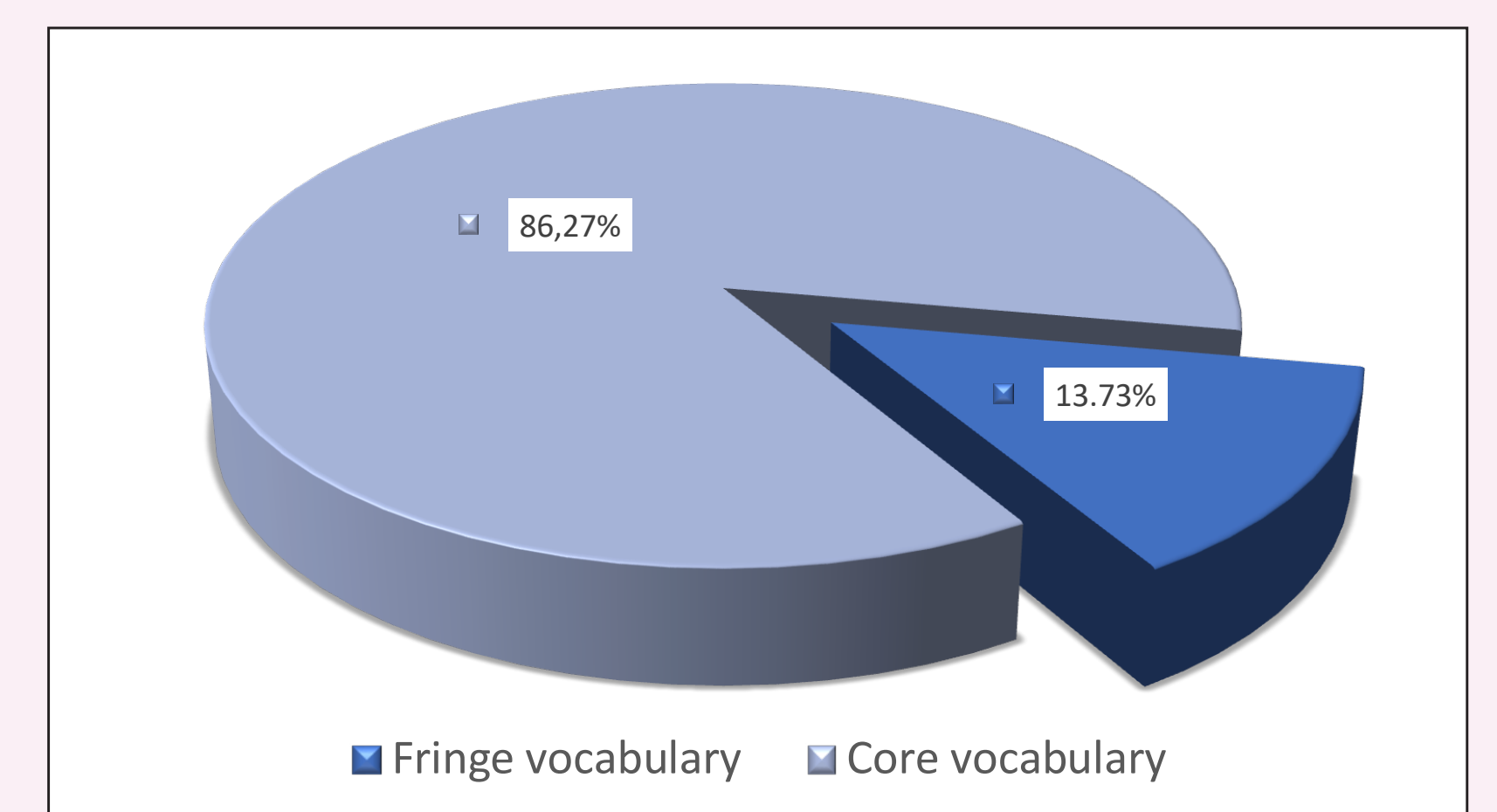

Figure 2. Coverage of core and fringe vocabulary

- The Setswana core vocabulary was analysed and classified into **content words** (can be used in isolation and have meaning) and **structure words** (syntactical words that do not carry meaning but give meaning to sentences).
- It was found that two-thirds of the core vocabulary list comprised of content words and one-third of structure words. See Figure 3.
- A further analysis was conducted to classify the core words into different parts of speech. See Figure 4.
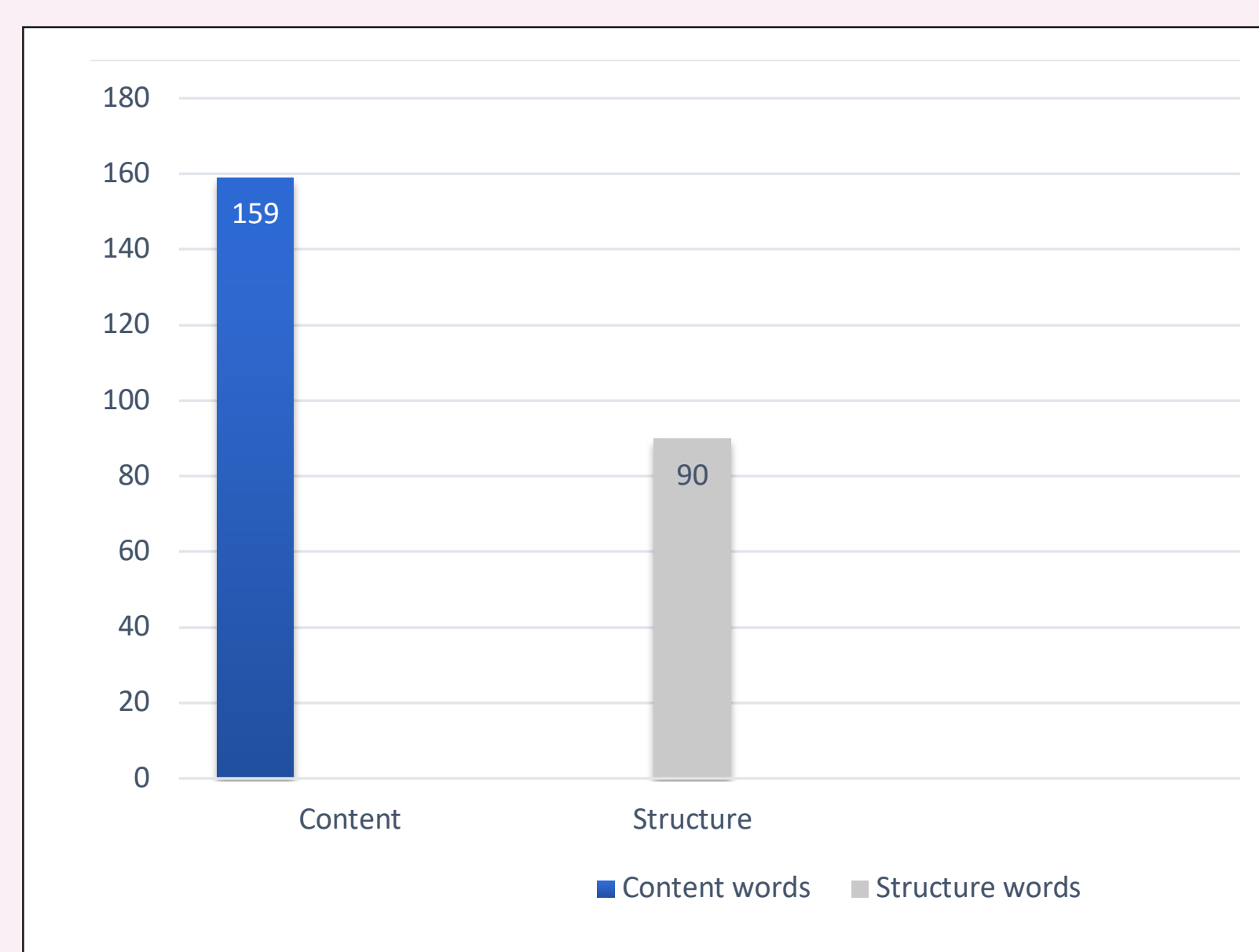

Figure 3. Number of content and structure words found in the core vocabulary
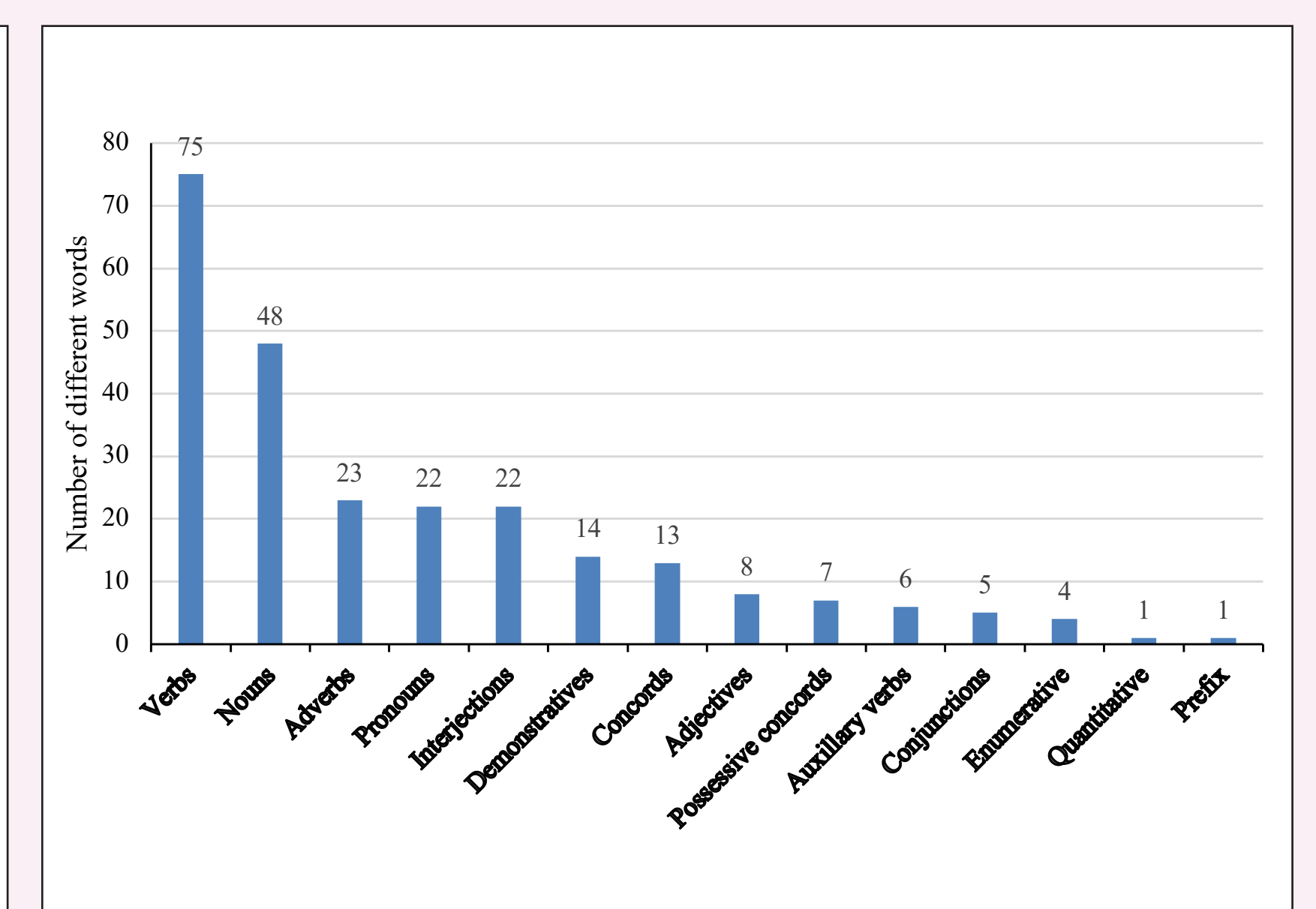

Figure 4. Number of different words of core vocabulary according to parts of speech

Conclusion

- The Setswana core vocabulary list can be used as a resource for vocabulary selection for Setswana AAC systems for preliterate children.
- This core list must be used in combination with other vocabulary sources such as informants or environmental inventories. This will ensure a vocabulary that is appropriately individualised while still allowing the formulation of novel sentences.

The Setswana and other South African core vocabulary lists can be obtained here:

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA