

Child speech database research project attracts international attention

By [Birgit Ottermann](#), issued by [North-West University \(NWU\)](#)

7 May 2024

A research project aimed at compiling a child speech database for the South African context, with a focus on speech samples of typically developing Afrikaans and Sesotho sa Leboa-speaking children, is garnering attention both nationally and internationally.



Funded by the South African Centre for Digital Language Resources (SADiLaR) and led by Juan Bornman, now professor in speech-language therapy at Stellenbosch University and research fellow at the University of Pretoria, the long-term objective of this collaborative project is to develop technical solutions that will assist speech-language therapists.

"A research database of child speech samples is needed to develop a solution that would be able to perform an automated transcription and linguistic analysis of recorded child speech samples in different South African languages," Prof Bornman explains. "This would assist speech-language therapists to optimise the use of language sample analysis (LSA), which is widely considered the gold standard for multilingual language assessment. It will build on and expand the resource base for South African languages to contribute to new knowledge on mono- and multilingual language acquisition and allow for the development of new assessments for early identification and improved interventions for children at risk in mono- and multilingual contexts."

Starting with Afrikaans and Sesotho sa Leboa-speaking children, Bornman and her team collected over 700 minutes of speech data in each language (either in the home or clinic context) in interaction with a trained speech-language therapist who used standardised material (such as toys and books) adapted to the specific age group.

The research team comprises Bornman (formerly of the University of Pretoria), Prof Jeannie van der Linde (University of Pretoria), Dr Febe de Wet (formerly Stellenbosch University, now industry), Ms Petria Winter (née Liebenberg, formerly Masters student at the University of Pretoria, now PhD student), as well as international collaboration partners Prof Ulrike Lüdtké, Prof Jörn Ostermann, Dr Hanna Ehler and their PhD students (Lars Rumberg, Christopher Gebauer, Edith

Beaulac) at Leibnitz University Hannover (LUH) in Germany.

International exposure

To date, they have published five academic articles on their research findings in international peer-reviewed journals, one book chapter (for an international book, edited by Dr Melissa Bortz of St John's University in New York), and given two presentations at conferences in the Netherlands (at the 19th Experimental Methods in Language Acquisition Conference in Utrecht) and the US (at the 2023 American Speech and Hearing Association Convention in Boston). They also hosted a very well-attended webinar in May 2023 (with 293 registered participants) to discuss tips and traps of orthographic and phonetic transcription.

"We are excited that we have been invited to give two more presentations on our research at the 16th International Association for the Study of Child Language (IASCL) Congress, from 15-19 July 2024, in Prague, the Czech Republic," Bornman says.

Code-switching in monolingual children

Asked about their research findings, Bornman highlights the phenomenon of code-switching in multilingual South Africa where seemingly 'monolingual' children spontaneously include words of another language (such as English).

"We have found that code-switching to English is prevalent in monolingual Afrikaans-speaking children during spontaneous language samples. The data also showed that these children primarily inserted English nouns in the matrix language utterances (in this case, Afrikaans) by means of intrasentential [inside a sentence] code-switching. Ninety percent of the code-switching in the dataset were examples of intrasentential code-switching. Across all the age groups, the percentage of English words that were used (in terms of the number of different words) was less than 10% overall. The Afrikaans-speaking children's use of code-switching may indicate second language acquisition of English or that they have already acquired English as a second language. The current study makes an important contribution to the existing literature as these results may prompt that more specific boundaries for the definitions of monolingual and multilingual individuals should be described to classify these participants and the code-switching phenomenon," she says.

When it comes to obtaining speech samples of Sesotho sa Leboa-speaking children, Bornman mentions that they needed to hold a special workshop with experts to decide on the best possible location to obtain the most standard version of Sesotho sa Leboa, since the influence of South Africa's other languages is so prevalent. "We also hosted a transcription webinar, as we found it hard to find linguists to assist with the orthographic and phonetic transcription of the Sesotho sa Leboa data."

Growing early career researchers from master's to PhD

Another interesting finding was that a 30-minute language sample provides a representative language sample for Afrikaans-speaking children who are between three-and-a-half and nine-and-a-half years old with comparable results obtained from a 60-minute language sample.

"This has significant time implications for speech-language therapists in terms of reducing the amount of time needed for language sample analysis."

Bornman is also excited about how this collaborative research project created an opportunity for early career researchers to get involved in a project and further their career.

"Besides the involvement of international PhD students from Leibnitz University Hannover, we have the wonderful success story of Petria Winter (née Liebenberg), who started the project as a master's student and graduated from the project (*cum laude*) and enrolled for a PhD to continue this research."

The research team is now in the final six months of the project, finalising the language samples (and their orthographic

transcriptions) collected from 60 Afrikaans- and 30 Sesotho sa Leboa-speaking children. SADiLaR funded the data collection of 30 Afrikaans-speaking children, and a similar data set (also with 30 Afrikaans-speaking children) was created during a related project that was funded by the South African Academy for Science and Art.

"SADiLaR has enabled the project to make significant progress," Bornman says. "With this support, the project has gained momentum and new projects stemming from this are currently underway. Thank you, SADiLaR."

- " **Prof Bismark awarded National Order Of Merit by France** 27 May 2024
- " **Former world and national Morganti Ju-Jitsu champion bags a PhD** 23 May 2024
- " **GoAllOut student chapter breaks Guinness World Record** 17 May 2024
- " **Seven Eagles to soar at the 2024 Olympic Games in France** 15 May 2024
- " **Child speech database research project attracts international attention** 7 May 2024

[North-West University \(NWU\)](#)



The North-West University (NWU) is one of South Africa's top five universities; that offers superior academic excellence, cutting-edge research and innovation and teaching and learning. It all starts here.

[Profile](#) | [News](#) | [Contact](#) | [Twitter](#) | [Facebook](#) | [RSS Feed](#)

For more, visit: <https://www.bizcommunity.com>