

## Research Article

# Describing the Spoken Language Skills of Typically Developing Afrikaans-Speaking Children Using Language Sample Analysis: A Pilot Study

Petria Liebenberg,<sup>a</sup> Jeannie van der Linde,<sup>a</sup> Isabella Schimper,<sup>a</sup> Febe de Wet,<sup>b</sup> Marien Graham,<sup>c</sup> and Juan Bormman<sup>d</sup>

<sup>a</sup>Department of Speech-Language Pathology and Audiology, University of Pretoria, South Africa <sup>b</sup>Department of Electrical and Electronic Engineering, Stellenbosch University, South Africa <sup>c</sup>Department of Science, Mathematics and Technology Education, University of Pretoria, South Africa <sup>d</sup>Centre for Augmentative and Alternative Communication, University of Pretoria, South Africa

## ARTICLE INFO

## Article History:

Received May 15, 2022

Revision received August 26, 2022

Accepted November 15, 2022

Editor-in-Chief: Amanda J. Owen Van Horne

Editor: Katie Squires

[https://doi.org/10.1044/2022\\_LSHSS-22-00077](https://doi.org/10.1044/2022_LSHSS-22-00077)

## ABSTRACT

**Purpose:** Language sample analysis is widely regarded as the gold standard of language assessment. However, the uncertainty regarding the optimal length of sample and the limited availability of developmental language data for nonmainstream languages such as Afrikaans complicate reliable use of the method. The study aimed to provide guidelines on representative length of sample and concurrently provide a preliminary description of the spoken language skills of Afrikaans-speaking children.

**Method:** The study involved 30 typically developing Afrikaans-speaking children aged between 3;6 and 9;6 (years;months). A descriptive research design was used to transcribe and analyze 1-hr interactions collected in natural environments of participants who were recruited using referral sampling. Video and audio recordings of the samples were transcribed using adapted Sampling Utterances and Grammatical Analysis Revised analysis procedures.

**Results:** Results indicated that mean length of utterance in words per minute, number of different words per minute, and total number of words per minute stabilized at 30 min and no significant differences were found between 30 min and longer time segments. Morphology results concur with existing developmental findings in Afrikaans. Lexical diversity results correlated with the findings of the lexical specificity and accuracy in the Prutting and Kirchner Pragmatic Protocol (Prutting & Kirchner, 1987). The developmental trajectories for pragmatic and phonological development were consistent with existing guidelines.

**Conclusions:** The study concluded that a 30-min interaction segment provides a representative language sample for Afrikaans-speaking children who are between 3;6 and 9;6. It provides promising preliminary developmental data and clinical guidelines, confirming the potential of language sample analysis (LSA) as a reliable component of language assessment in Afrikaans.

The effects of globalization on speech-language pathology (SLP) practices are evident in the increasingly diverse caseloads worldwide (van Dulm & Southwood, 2014). In South Africa, challenges in language assessment are compounded by the 11 official languages (isiZulu,

isiXhosa, Afrikaans, English, Sesotho sa Leboa, Setswana, Sesotho, Xitsonga, siSwati, Tshivenda, and isiNdebele) and diversity regarding dialects, socioeconomic status, and culture. Afrikaans is the third most widely spoken official language, around 13.5% (Statistics South Africa, 2012), and is the second most prevalent language of learning and teaching in single medium schools in South Africa (South African Department of Basic Education, 2010; Wildsmith-Cromarty & Balfour, 2019). It is disconcerting, therefore, that only a few language assessment instruments

Correspondence to Jeannie van der Linde: [jeannie.vanderlinde@up.ac.za](mailto:jeannie.vanderlinde@up.ac.za).

**Disclosure:** The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

and methods have been developed for Afrikaans-speaking children (Southwood, 2013). This limited availability of language assessment instruments in some world languages poses an increasing challenge internationally as the incidence of multilingual individuals is rising faster than language assessment tools are being developed (Ebert, 2020). Despite the obvious pitfalls and dangers of using inappropriate language assessment measures (Barratt et al., 2012; Bornman et al., 2018; Pascoe et al., 2013; Southwood & Van Dulm, 2015; Verdon et al., 2015), clinicians assess Afrikaans-speaking preschool children's spoken language skills using unvalidated measures in the absence of reliable and valid alternatives (Southwood, 2005). As in the case of all other South African languages and nonmainstream languages across the world, there is an urgent need for culturally and linguistically appropriate language assessment measures (Southwood, 2013; Southwood & Van Dulm, 2015).

## Language Sample Analysis

Reliable language assessment in nonmainstream languages is challenging (Barratt et al., 2012; Bornman et al., 2018; Pascoe et al., 2013; Southwood & Van Dulm, 2015; Verdon et al., 2015). Language sample analysis (LSA) may provide a reliable alternative to traditional clinical language assessment as a more linguistically and culturally appropriate language measure (Bowles et al., 2020; Govindarajan & Paradis, 2019; Hux et al., 1997).

### Strengths of LSA

LSA provides a way to assess the form (*morphology*, *syntax*, and *phonology*), content (*semantics*), and use of language (*pragmatics*) as well as a child's ability to integrate these domains to communicate in everyday conversations (Bowles et al., 2020). For the purpose of this study, LSA was used to assess children's language in the context of familiar, natural, and linguistic activities, such as in conversations or narratives (Bowles et al., 2020; Channell et al., 2018). LSA therefore also provides a unique opportunity to be used as a method to describe the development of spoken language. Meaningful information about a child's functional and social use of language, an aspect often overlooked in traditional language assessment measures, can be obtained with LSA (Bowles et al., 2020; Gentileau-Lambin et al., 2019; Spencer et al., 2020).

## Language Measures to Describe Spoken Language Skills Using LSA

### Form of Language: Morphosyntactic Skills

Mean length of utterance (MLU) is a measure that indicates the average length of an utterance and can be used with reference to words or morphemes to provide an

overarching reflection of a child's morphosyntactic skills (Pezold et al., 2020). For this study, an alternative method of MLU in words (MLU-w) was used to ensure the highest possible reliability and generalizability of the findings for use in clinical settings (Oosthuizen & Southwood, 2009).

### Form of Language: Phonology

For phonological analysis, phonetic transcription and analysis can be used to accurately describe and ultimately assess phonological developmental trajectories (Geertsema, 2016). The phonetic transcription of samples in this study was beyond the scope of the current research. However, the limited availability of such data in Afrikaans (Geertsema, 2016) underlines the usefulness of LSA as a method to describe phonological development.

### Content of Language: Semantics

Calculating number of different words (NDWs) in a language sample is common practice among researchers in the field of LSA of narratives and discourse to determine lexical diversity (Charest et al., 2020; Ebert, 2020; Ebert & Scott, 2014; Imgrund et al., 2019; Pavelko & Owens, 2017). NDW has higher test-retest reliability in larger samples than the total number of words (TNW) used and is useful to indicate the diversity of vocabulary used (Pezold et al., 2020). This measure can be used to describe the lexical diversity as well as to determine differences in lexical diversity between different age groups.

### Use of Language: Pragmatics

The recognized Prutting and Kirchner (1987) consists of descriptions of 30 behaviors subdivided into three categories, namely, verbal aspects (e.g., topics; turn taking), paralinguistic aspects (e.g., intelligibility; prosodics), and nonverbal aspects (e.g., kinesics; proxemics) to guide pragmatic skill observation in naturalistic tasks (Adams, 2002; Prutting & Kirchner, 1987). Within these subcategories, relevant information may be gained regarding narrative pragmatic skills development (Gentileau-Lambin et al., 2019). LSA, which represents a naturalistic linguistic task, can be used to describe the use of language as well; this protocol provides a valid and reliable guideline to describe and observe pragmatics.

## Barriers to the Use of LSA

Despite the benefits of LSA, divergent and often conflicting descriptions of the context as well as the elicitation and analysis methods for LSA have been put forward, a situation that detracts from its perceived reliability (Alvesson & Kärreman, 2011; Bliss et al., 1998; Farahani, 2013). The most prominent barriers to efficient LSA across the globe, especially in low and middle-income

countries (LMICs), include (a) resource constraints in terms of time and manpower, (b) limited guidelines regarding the length of an adequate sample, and (c) the limited availability of developmental data for nonmainstream languages (Pavelko et al., 2016), specifically South African languages, which includes Afrikaans.

Collecting, transcribing, and analyzing language samples are time-consuming processes (Pavelko et al., 2016). In South Africa, where SLPs have extremely large caseloads with 1:25,000 SLP-to-population ratio (Kathard & Pillay, 2013), time constraints are a pressing issue (Moonsamy & Kathard, 2015). In comparison with countries such as the United States, United Kingdom, Canada, and Australia where the SLP-to-population ratio ranges between 1:2500 and 1:4700 (Kathard & Pillay, 2013), the time constraints and undersourcing in South Africa are particularly menacing. It is evident that improving LSA guidelines, such as the parameters of what constitutes a representative language sample in terms of time or the number of words needed to ensure efficient language assessment in practice, is an urgent matter.

The literature provides inconsistent guidelines for the length of sample (in either number of utterances or in minutes) needed to provide a holistic view of a child's natural use of language (Pezold et al., 2020; Tommerdahl & Kilpatrick, 2014). Suggested sample length in terms of number of utterances ranges from a minimum of 50 utterances up to 175-utterance samples (Gavin & Giles, 1996; Pavelko & Owens, 2017; Pavelko et al., 2016; Shipley & McAfee, 2016). Some studies recommend measuring samples according to the duration of the sample in minutes, with recommended duration ranging from 1 to 2 min to more than 11 min (Heilmann et al., 2010; Pavelko et al., 2016; Tilstra & McMaster, 2007). Research investigating the reliability and representativeness of shorter samples (1- and 3-min samples) compared to longer (7 min) conversational and narrative samples found that the shorter samples could be regarded as dependable, because there were no significant differences between these samples (Heilmann et al., 2010). Although there is no conclusive evidence regarding the representative length of samples, researchers agree that length of sample influences the outcome of the assessment and that longer samples do tend to yield more reliable results (Heilmann et al., 2010). The undefined length of sample (Pavelko & Owens, 2017) and the limited uniformity regarding sample collection procedures (Hux et al., 1997) may influence the validity of LSA as a criterion-referenced language assessment measure. This could explain why SLPs often view language sampling and LSA as time consuming and unstructured and therefore opt not to use it in clinical practice despite acknowledging the valuable information that it may provide (Casby, 2011; Heilmann et al., 2010; Pavelko et al., 2016).

In addition, despite research in monolingual and bilingual children that provides norms for narrative development (e.g., the database available in SALT), there is lack of developmental language data for the different South African languages. This fact, together with the lack of consensus on the length or duration of samples (in terms of minutes or number of utterances), makes the use of LSA in South African languages challenging (Bedford et al., 2013; Brothers et al., 2008). Providing developmental language guidelines requires a relatively homogenous standardization group (De Lamo White & Jin, 2011) with equivalence in terms of culture, socioeconomic status, linguistic background, age, gender, and ethnicity (Saenz & Huer, 2003). The diverse population of multilingual and multicultural South Africa, therefore, poses specific challenges for obtaining developmental language data. While LSA is a well-suited approach within the South African context and other diverse contexts, the lack of developmental language data for nonmainstream languages limits the use of LSA across languages (Van Dulm & Southwood, 2014). In this context, the use of criterion-referenced measures is usually regarded as more appropriate (Shipley & McAfee, 2016). Criterion-referenced measures compare the individual's level of performance on a specific skill with a clinical expectation or predetermined performance criterion, which is based on developmental language data or language descriptions (De Lamo White & Jin, 2011).

The lack of information about the typical language skills of children who speak nonmainstream languages as a first language hampers the process of appropriate assessment (Southwood & Van Dulm, 2015). This unsatisfactory situation prompted the investigation of LSA as an alternative criterion-referenced measure in Afrikaans-language assessment, which would require describing the typical spoken language skills of Afrikaans-speaking children. A preliminary report can yield developmental language data validating LSA as a criterion-referenced measure for practice.

Although LSA could provide a viable alternative to norm-referenced standardized measures, a way to address the lack of appropriate resources to evaluate the content, form, and use of language for comprehensive language assessment in nonmainstream languages, challenges such as the lack of developmental language data (De Lamo White & Jin, 2011; Southwood & Van Dulm, 2015), contradicting evidence regarding adequate length of sample (Pezold et al., 2020), and time constraints (Pavelko et al., 2016) prevent SLPs across the globe from effectively using LSA when conducting language assessment. If research could show that the representative sample length was actually attainable within a reasonable time frame, it would be a boon to SLPs and greatly to the advantage of their clients. This study therefore intended to answer the

following questions: (a) What length of language sample is representative of spoken language skill when using LSA (in number of words or minutes), and (b) what are the typical spoken language skills of Afrikaans-speaking children?

## Research Aims

This pilot study aimed to propose guidelines for a representative language sample by investigating and describing length-of-sample results for Afrikaans-language samples. It further sought to provide a preliminary description of the spoken language skills of typically developing Afrikaans-speaking children between 3;6 (years;months) and 9;6, using LSA.

## Method

### Research Design

A descriptive, cross-sectional, and quantitative design (Leedy & Ormrod, 2020) was used to collect, transcribe, and analyze audio- and video-recorded language samples of the spoken language of typically developing Afrikaans-speaking children. The audio recording ensured clear and high-quality language recordings for reliable transcription, while video recordings assisted in the analysis of language use.

### Participants

Referral sampling (Chambers et al., 2020) was used to recruit a stratified sample in terms of gender and age. Inclusion criteria were the following: (a) Afrikaans first language speakers; (b) between the ages of 3;6 and 9;6; (c) typically developing; (d) middle to high socioeconomic status (SES), that is, tax-paying and living in what is considered middle-to-high-class neighborhoods; and (e) normal hearing status. The latter three criteria were included as delayed development, low SES, and hearing loss are commonly associated with language delays (Bowles, 2018; Tomblin et al., 2014).

Although only Afrikaans first language (“monolingual”) speakers were included, the majority of the participants ( $n = 24$ ; 80%) had also been exposed to a second language, as is to be expected in a multilingual country such as South Africa. The language of learning and teaching for the majority of the participants ( $n = 29$ ; 96.67%) was Afrikaans, whereas one participant’s language of learning and teaching was both Afrikaans and English.

Code-switching occurred throughout the samples due to English language exposure. These observations led to a separate in-depth paper (Liebenberg et al., in press) about this phenomenon, as it was an unexpected result

given the fact that the children were regarded as “monolingual” despite living in a multilingual country.

For socioeconomic inclusion, all the participants’ ( $N = 30$ ) caregivers earned a collective household income of more than R 7,000 per month and thereby were tax-paying citizens.

The age range was stratified into five subcategories, namely, 3;6–3;11 ( $M = 3.72$ ,  $SD = 0.13$ ), 4;0–4;11 ( $M = 4.22$ ,  $SD = 0.13$ ), 5;0–5;11 ( $M = 5.47$ ,  $SD = 0.31$ ), 6;0–6;11 ( $M = 6.33$ ,  $SD = 0.43$ ), and 7;0–9;6 ( $M = 8.48$ ,  $SD = 0.73$ ), with three male and three female participants per category. These age categories were based on the age ranges most often used in language assessment. This resulted in a total sample of 30 participants.

## Procedure

Institutional review board approval was obtained before commencing with data collection [HUM001/1220 (Amendment)]. Participation in the study was voluntary, and informed consent was obtained from the participants’ legal guardians as well as age-appropriate assent from the participants themselves. Confidentiality was maintained throughout the study by assigning alphanumeric codes to each participant, omitting all personally relevant information from the transcriptions after analysis, and by excluding this information when reporting individual findings.

## Equipment and Materials

*Materials for candidacy.* A custom-designed biographic questionnaire was used. It included questions regarding family involvement, parental education, and SES in order to determine inclusion into the sample. Specific questions regarding developmental and scholastic concerns (the latter focused on the older participants) were also included in the questionnaire. For the younger participants ( $< 7;11$ ), developmental information was obtained using the Parent’s Evaluation of Developmental Status: Developmental Milestones (PEDS:DM) screening tool (Glascoe, 2013). The PEDS:DM is only suitable for children from birth to 7;11. The validity, reliability, and accuracy of this developmental screener have been established (Bedford et al., 2013; Brothers et al., 2008), and it has often been used in the South African context (e.g., van der Linde, 2015; van der Merwe et al., 2017). The PEDS:DM, an English tool that has been translated into five South African languages, including Afrikaans, has been widely used for developmental surveillance in South Africa and for research purposes (van der Linde, 2015; van der Merwe et al., 2017). For inclusion in this study, individuals had to have no concerns listed on the translated PEDS:DM.

In the case of the older participants ( $> 7;11$ ), their scholastic performance records were used as the basis for

developmental inclusion, the requirements being that they had never repeated a school year nor had ever received speech therapy; information also obtained from the biographic questionnaire. The hearScreen Mobile App (hearX Group) was used to detect possible hearing loss. This method has been shown to have high accuracy when compared with traditional screening measures (Mahomed-Asmail et al., 2016).

*Materials for elicitation.* A qualified SLP elicited discourse and narrative interactions, using age-appropriate pictures and wordless books as well as age-appropriate, gender-neutral toys, and games. The intention was that the elicitation materials should not create noise that would mask the speech signal, but it was challenging to avoid noise generated during natural play as well as normal background noise in home environments and residential neighborhoods.<sup>1</sup> All materials were culturally relevant to ensure narratives and materials that would be familiar in the linguistic, cultural, and socioeconomic context of participants (Southwood & Russell, 2004). The materials were selected by a practicing, licensed Afrikaans-speaking SLP who is knowledgeable in child language.

Three different sets of elicitation material were used for data collection. All the toys, pictures, and wordless books were selected to be developmentally appropriate for, respectively, 3;6–4;11 years, 5;0–6;11 years, and 7;0–9;6 years. Each set included materials for drawing as the introductory activity. For participants from 3;6 to 4;11, the SLP started by requesting, “Draw a picture of yourself.” For the 5;0–6;11 age category, the introductory activity prompt was extended to, “Draw a picture of everyone you live with.” The prompt for the drawing activity in the 7;0–9;6 age category was, “Draw a picture of yourself with your friends.” The toys for the youngest age group included a toy kitchen and doctor or vet toys to be used for the elicitation of symbolic play, personal event narratives, and familiar routines. For the 5;0–6;11 age group, toys such as wild animals and sea creatures were included for story generation and elicitation of discourse. Some familiar routines that were included were shopping and fast-food scenes to encourage symbolic play and personal event narratives. The materials for the oldest group included games with dice and questions that required some verbal reasoning, such as, “If you had to go to the moon, what would you take with and why?” The activities focused on emotions, verbal reasoning, and

problem-solving, and participants were prompted to convey what the people in the wordless books were thinking.

To obtain the best possible quality audio recording, an appropriate microphone (Zoom H1n Handy Recorder) was used; its technical specifications and portability made it suitable to use. It was augmented by a mobile phone (Samsung Galaxy S7 Edge) for the video recordings. For the transcription and analysis of the samples, a personal computer was used according to the procedures prescribed by Sampling Utterances and Grammatical Analysis Revised (SUGAR; Pavelko & Owens, 2017) with noted adaptations in the Appendix for Afrikaans (Oosthuizen & Southwood, 2009).

### Data Collection

After the parents of the participants had provided their informed consent, they completed the biographic questionnaire. The interaction with the participants took place in their natural home environments, with the exception of one participant who came to the Department of Speech-Language Pathology clinic at the University of Pretoria for the interaction. The elicitation of the samples typically took place with only the child and the SLP present, although a family member also joined in some cases. The recorded interaction lasted 1 hr in each case. Considering setup time, obtaining assent, and hearing screening in addition to the interaction itself, however, the SLP spent approximately 75 min with each participant.

### Transcription Procedures

Three suitably qualified raters (two SLPs and one linguist) transcribed the samples using their personal computers and a free downloadable software program, ELAN (ELAN v. 6.1, 2021), designed for the transcription of audio or video recordings. To ensure consistency, the researcher discussed transcription procedures (see the Appendix) with each rater and shared a document containing all the procedures with them. Five randomly selected interactions—one per age group—were transcribed by all three raters to measure interrater reliability.

The SUGAR procedures that were used included that the sample was transcribed by retyping only the child’s utterances, with spaces between each word. All words that the child directly imitated were omitted from the transcription. No utterance was changed in any way; therefore, no morphemes that the child omitted were added. The SUGAR procedures further prescribe that no fillers, such as “uhm,” should be transcribed and no disfluencies should be included. Repeated words were only included if they were used for emphasis. When more than two clauses were joined by “en” {and}, only the first “en” was transcribed and the rest were considered a run-on sentence. Onomatopoeia was omitted from the samples and therefore did not count for NDW and TNW measures.

<sup>1</sup>Although the home environment provides functional, naturalistic information about a child’s functioning, it is not always feasible in practice to visit clients (Kramer et al., 1979). This study was conducted during the COVID-19 pandemic and could therefore not accommodate all the participants in a clinical environment due to the government’s regulations at the time of data collection.

Transcription procedures are described in detail in the Appendix.

The transcription time for each interaction was approximately 4 hr. This study supports the notion that transcription is a meticulous and time-consuming process (Pavelko et al., 2016). Consequently, clear guidelines regarding the shortest possible sample length of interaction to ensure a representative sample is essential.

## Data Analysis

*Length-of-sample guidelines.* The full 1-hr interactions were analyzed to determine guidelines for length of representative sample. Each sample was divided into timed intervals of 0–5 min, 0–10 min, 0–20 min, 0–30 min, 0–40 min, and 0–60 min. The analysis was started at 0 min when the SLP initiated the interaction with the first request to produce a drawing. Measures such as NDW and TNW, which are directly proportional to time, will obviously increase as time increases. To ensure that the length-of-sample guidelines were based on comparable measures, therefore, the measures were calculated per minute for NDW, and TNW was calculated for each of the following interval values: 5, 10, 20, 30, 40, and 60 min. The MLU-w was calculated for the same intervals. The researchers compared the MLU-w, NDW per minute, and TNW per minute at each time interval in order to determine when these measures would stabilize to yield a representative sample.

Familiarity influences were considered to determine the length-of-sample guidelines. The “warm-up effect” suggests that the first part of a language sample may skew the results, as SLPs differ in their interaction styles when trying to overcome the unfamiliarity influence on the interaction (Heilmann et al., 2010).

To investigate the potential influence of this effect, the transcribed samples were analyzed twice, in sections of 30- and 40-min duration. The sample from 0 to 30 min

(Sample 30 A) was correlated to the sample from 10 to 40 min (Sample 30 B), and the same procedure was used to correlate the two samples of 40 min each (0–40 min [40 A] vs. 10–50 min [40 B]). No statistically significant differences between Samples A (from 0 min) and Samples B (from 10 min) were found for any of the measures.

## The Spoken Language Skills of Typically Developing Afrikaans-Speaking Children

LSA procedures were used to analyze each participant’s transcribed sample. The comprehensive capabilities of LSA include assessing all language domains, that is, form (*morphology, syntax, and phonology*), content (*semantics*), and use (*pragmatics*; Owens, 2016). LSA is also a naturalistic method to describe spoken language skills with the potential to describe multiple linguistic factors beyond the ones described here.

For this study, MLU-w, NDW, and the Prutting and Kirchner Pragmatic protocol (Prutting & Kirchner, 1987) scores were calculated from the transcribed and coded data to describe the spoken language skills of typically developing Afrikaans-speaking children. The procedures for the calculation of each of the abovementioned language measures can be seen in Table 1.

## Interrater Reliability

Interrater reliability was determined for the transcriptions as well as the Pragmatic Protocol annotations. Three raters (two SLPs and one linguist) transcribed the samples, whereas a separate group of three SLPs implemented the Pragmatic Protocol to rate the sample in terms of language use.

The interrater agreement measures for two of the variables, number of total utterances (NTU) and NDW,

**Table 1.** Measures for data analysis in each language domain.

Language domain and measure	Procedures and examples in Afrikaans
Morphology and syntax: <i>Mean length of utterance (MLU)</i>	The alternate method of MLU-w calculation has been adapted to still include utterances where the child completes the adult’s utterance and social utterances, such as <i>wat is dit?</i> “what is this?” <i>nee dankie</i> “no thanks,” <i>kyk hier</i> “look here,” <i>so</i> “like this.” These utterances from part of typical discourse and were therefore not excluded during transcription (Oosthuizen & Southwood, 2009).
Phonology: <i>Informal description</i>	To calculate MLU-w, the TNW was divided by the number of total utterances (NTU). Typical phonological errors noted in the Afrikaans samples included distortion of the /s/ and the alveolar trill /r/ sound in all sound positions. Further research is required to provide developmental phonological data in the spontaneous language samples of young Afrikaans-speaking children.
Semantics: <i>Number of different words (NDW)</i>	The NDW were calculated for each participant to compare each age category and determine the developmental semantic data as well as to identify age-related changes to the semantic skills of Afrikaans-speaking children.
Pragmatics: <i>Prutting and Kirchner Pragmatic Protocol (1987)</i>	The samples were video-recorded, which enabled the researchers to consider and investigate typically occurring pragmatic behaviors using an evidence-based checklist, namely, Prutting and Kirchner Pragmatic Protocol (Prutting & Kirchner, 1987). (Adams, 2002; Prutting & Kirchner, 1987).

were excellent (NTU = 0.963; NDW = 1.000; Cicchetti, 1994). For MLU-w and TNW, the interrater reliability was determined using nonparametric Spearman correlations, which revealed moderate to very strong correlations (MLU-w = .676; TNW = .977; Akoglu, 2018). For pragmatics scores, a strong positive correlation of .889 was found between Rater 1 and Rater 3, and the values of the two raters were averaged.

## Results

Extreme values (or outliers) may appear in a data set. When cleaning the data, it was inspected for any problematic outliers due to data entry or data capturing or errors, and it was found that there were no erroneous extreme values. All true non-erroneous outliers were kept, as many researchers have advocated for not simply blindly removing outliers (unless they are clearly erroneous and cannot be corrected), as these are legitimate values and not actual errors representing natural variation in the population (Nicklin & Plonsky, 2020; Streiner, 2018). Non-erroneous outliers may contain potentially useful and unexpected results, and the removal of these from the data set may influence the results (Streiner, 2018). In addition to this, Mowbray et al. (2019) caution against blindly deleting extreme values when the sample size is not large enough to withstand the deletion of any values, which is the case for this study.

Although inclusion of extreme values may have pitfalls such as skewing the data, these can be addressed by using more robust statistics that are less influenced by outliers, such as reporting on the median and interquartile range (IQR) rather than the mean and standard deviation (SD) and using nonparametric methods rather than parametric tests (for inferential statistics; Streiner, 2018). The recommendation to use robust nonparametric tests without outlier removal has been made by other researchers as well (Bakker & Wicherts, 2014; Wale et al., 2020), because nonparametric methods can tolerate outliers. Therefore, for this study, the median and IQR alongside the mean and SD were reported, and nonparametric methods were used for all variables that differed significantly from normality to address the possible undue influence of retaining the true non-erroneous outliers.

### Length-of-Sample Guidelines

To determine representative length-of-sample guidelines, Friedman's test and post hoc WSR tests (when Friedman's  $p$  value was less than 0.05) were used to determine whether there were statistically significant differences between the different intervals for the MLU-w measures and the per minute measures that were

calculated for number-of-word and total-number-of-words measures (NDW/m and TNW/m). Table 2 depicts all the significant and nonsignificant differences found during pairwise comparison between the different intervals.

No statistically significant differences were found between 20 min and the subsequent time segments for MLU-w and TNW/m measures across all five age groups (see Table 2). For NDW/m, the results of the Friedman test for the 3;6–3;11 and the 4;0–4;11 cohorts were not significant ( $p > .05$ ) and pairwise comparisons were not run. For the three older cohorts, however, the differences between 20 and 60 min were statistically significant ( $p < .05$ ). The NDW/m measure at 30 min versus 60 min for the 7;0 to 9;6 age cohort also differed significantly ( $p = .045$ ). In addition to the length-of-sample guideline in minutes, the average TNU and TNW measured at 30 min for each age category are shown in Table 3. This provides further guidelines in terms of average number of utterances and words obtained in timed samples.

## The Spoken Language Skills of Typically Developing Afrikaans-Speaking Children

The descriptive statistics relating to the measures for morphosyntactic skills (MLU-w), semantics (NDW), and pragmatics (Pragmatic Protocol [PP]) appear in Table 4.

### Morphosyntactic Skills

When comparing the youngest group (ages 3;6–3;11) to the oldest group (ages 7;0–9;6), there were significant differences in the measures for MLU-w ( $p = .009$ ), NDW ( $p = .002$ ), and TNW ( $p = .009$ ). Although the MLU results obtained for both the 5-year-olds and the 6-year-olds were higher than that of the 7- to 9-year-olds, the differences between the cohorts (5;0–5;11 vs. 7;0–9;6 [ $p = .937$ ] and 6;0–6;11 vs. 7;0–9;6 [ $p = .589$ ]) were not significant. Similarly, the 3-year-olds obtained higher mean MLU-w scores than the 4-year-olds, but the difference was not statistically significant ( $p = .310$ ).

### Phonology

An informal description of phonology could be obtained reliably after SLPs reached consensus regarding the observed error patterns from the data collected for LSA. The most common phonological error across all age categories was distortion and/or deletion of the /r/ sound ( $n = 8$ ). In the youngest group, /s/ distortion ( $n = 1$ ), weak syllable deletion ( $n = 2$ ), and final consonant deletion ( $n = 1$ ) were noted in one or two cases. The same errors (/s/ distortion  $n = 1$ ; weak syllable deletion  $n = 1$ ; final consonant deletion  $n = 2$ ) were noted in the 4;0–4;11 age category, but the final consonant deletion errors were much more inconsistent than in the younger group. One

**Table 2.** *p* values for the post hoc Friedman test for differences of MLU-w/m, NDW/m, and TNW/m between the different intervals.

Time	<i>p</i> values for:	3;6–3;11	4;0–4;11	5;0–5;11	6;0–6;11	7;0–9;6
5 min vs. 10 min	MLU-w	0.537	0.165	0.643	0.123	0.758
	NDW/m	—	—	0.465	0.465	0.855
	TNW/m	0.715	—	0.068	—	—
5 min vs. 20 min	MLU-w	0.165	0.217	0.165	*0.014	0.643
	NDW/m	—	—	*0.045	0.144	0.584
	TNW/m	0.201	—	0.068	—	—
5 min vs. 30 min	MLU-w	*0.045	*0.045	0.280	0.217	0.123
	NDW/m	—	—	*0.009	*0.014	0.280
	TNW/m	0.064	—	0.219	—	—
5 min vs. 40 min	MLU-w	*0.009	0.123	0.537	*0.003	0.355
	NDW/m	—	—	*0.002	*0.003	*0.028
	TNW/m	*0.018	—	*0.028	—	—
5 min vs. 60 min	MLU-w	*0.009	*0.031	*0.045	*0.045	0.643
	NDW/m	—	—	*0.000	*0.000	*0.003
	TNW/m	*0.018	—	*0.001	—	—
10 min vs. 20 min	MLU-w	0.440	0.877	0.355	0.355	0.877
	NDW/m	—	—	0.201	0.465	0.465
	TNW/m	0.361	—	1.000	—	—
10 min vs. 30 min	MLU-w	0.165	0.537	0.537	0.217	0.217
	NDW/m	—	—	*0.045	0.064	0.165
	TNW/m	0.123	—	0.563	—	—
10 min vs. 40 min	MLU-w	*0.045	0.877	0.877	0.165	0.537
	NDW/m	—	—	*0.018	*0.028	*0.018
	TNW/m	*0.045	—	0.715	—	—
10 min vs. 60 min	MLU-w	*0.045	0.440	0.123	0.643	0.877
	NDW/m	—	—	*0.001	*0.001	*0.002
	TNW/m	*0.045	—	0.144	—	—
20 min vs. 30 min	MLU-w	0.537	0.440	0.758	0.217	0.280
	NDW/m	—	—	0.355	0.217	0.355
	TNW/m	0.440	—	1.000	—	—
20 min vs. 40 min	MLU-w/m	0.217	0.758	0.440	0.643	0.643
	NDW/m	—	—	0.273	0.144	0.100
	TNW/m	0.273	—	0.715	—	—
20 min vs. 60 min	MLU-w	0.217	0.355	0.537	0.643	1.000
	NDW/m	—	—	*0.028	*0.011	*0.018
	TNW/m	0.273	—	0.144	—	—
30 min vs. 40 min	MLU-w	0.537	0.643	0.643	0.090	0.537
	NDW/m	—	—	0.355	0.355	0.165
	TNW/m	0.537	—	0.438	—	—
30 min vs. 60 min	MLU-w	0.537	0.877	0.355	0.440	0.280
	NDW/m	—	—	0.064	0.064	*0.045
	TNW/m	0.440	—	0.063	—	—
40 min vs. 60 min	MLU-w	1.000	0.537	0.165	0.355	0.643
	NDW/m	—	—	0.273	0.273	0.465
	TNW/m	1.000	—	0.273	—	—

Note. — indicates where no significant differences were found upon calculation of the Friedman test, therefore pairwise comparison for these values was not performed. MLU-w = mean length of utterance in words; NDW/m = number-of-words measures; TNW/m = total-number-of-words measures. \**p* value significant at a 5% level of significance.

participant also presented with inconsistent initial consonant deletion errors.

In the 5;0–5;11 and 6;0–6;11 age categories, no phonological errors were noted. In the oldest group (7;0–9;6), one participant distorted the /r/ sound, whereas none of the other participants had any speech sound errors.

### Semantic Skills

The NDW scores indicated the expected mastery of semantic skills with age, as these scores increased steadily with age. There were no significant differences in the

NDW scores between the 5;0–5;11 and the 6;0–6;11 cohorts ( $p = 1.000$ ) or between the 6;0–6;11 and the 7;0–9;6 cohorts ( $p = .169$ ). The NDW scores in Figure 1 depict the developmental age trajectory for semantic abilities, which increases from a mean score of 231.50 different words for the youngest group to a mean of 421.00 different words for the oldest group.

### Pragmatics

A similar increase in the mastery of pragmatic skills was observed across the age categories. In Figure 2, the



**Table 3.** TNU and TNW used at 30 min of the interaction.

Measure	Variable	3;6–3;11	4;0–4;11	5;0–5;11	6;0–6;11	7;0–9;6
TNU	<i>M</i> ( <i>SD</i> )	163.5 (58.2)	155.8 (88.6)	199.8 (53.0)	214.5 (47.4)	249.1 (94.2)
	<i>Mdn</i> ( <i>IQR</i> )	157.0 (82.0)	152.5 (186.1)	182.5 (105.0)	198.0 (95.8)	233.8 (141.3)
TNW	<i>M</i> ( <i>SD</i> )	874.0 (477.4)	719.9 (472.7)	1542.2 (403.7)	1641.0 (664.7)	1909.1 (660.2)
	<i>Mdn</i> ( <i>IQR</i> )	793.0 (817.5)	846.5 (909.4)	1491.5 (785.5)	1618.0 (1300.0)	1705.0 (1221.0)

Note. TNU = total number of utterances; TNW = total number of words; IQR = interquartile range.

stabilization of pragmatic skills with age can be seen in the smaller variation of the pragmatic protocol scores as children mature.

The maturation of specific pragmatic skills is clear from Figure 3, especially pertaining to verbal aspects. The younger groups, ages 3;6–4;11, displayed low levels of lexical accuracy and specificity ( $n = 7$ ). This finding may be attributed to the still-developing vocabularies of these groups, concurring with the significantly lower NDW scores of these two categories compared to the older groups. Eye gaze improved for some of the participants ( $n = 4$ ) as the interaction progressed. The paralinguistic aspects remain relatively stable (93.33%–100%) throughout the chronological progression. In the 5;0–5;11 age group, verbal aspects increase while paralinguistic aspects decrease.

The slightly lower nonverbal aspects score for the two older categories (6;0–6;11 and 7;0–9;6) is attributed to many participants ( $n = 7$ ) opting for the interaction to take place at a tabletop. This preference diminished the potential observability of natural physical proximity as the children were sitting at appropriate physical proximity due to the table layout rather than due to appropriate pragmatic skill. The initial overall pragmatic scores of typically developing Afrikaans-speaking children, varying between 86.11% and 100%, indicate the sophisticated language use of this population.

## Discussion

LSA has been described as the gold standard for natural language assessment (Channell et al., 2018; Heilmann et al., 2010); however, discrepancies in guidelines regarding suggested length of sample hamper the efficient use of the method in clinical practice (Pavelko & Owens, 2017; Pavelko et al., 2016). In the case of the Afrikaans clinical population, the use of LSA is further complicated by the limited availability of developmental language data (Southwood & Van Dulm, 2015). This pilot study aimed to provide clinical guidelines for a representative, yet effective

sample length in addition to preliminary developmental spoken language data regarding Afrikaans-speaking children.

## Length of Sample

The study found that an Afrikaans language sample is representative of morphosyntactic language skill at 20 min. It further confirmed that the sample can be taken from 0 to 30 min of interaction with no significant effect ( $p = .094$ – $.156$ ) when the first part of the interaction was disregarded. Previous studies suggest that when using unstructured conversational and narrative tasks to elicit language samples, a language section longer than the frequently advised 50 utterances may be necessary (Gavin & Giles, 1996; Heilmann et al., 2010). The results of this study confirm that longer samples are needed to reliably obtain a representative language sample. This pilot study does not agree, however, with the recommendations of Heilmann et al. (2010) who compared the reliability of measures in much shorter intervals of 1-min versus 3-min versus 7-min samples. The current results indicate that longer intervals may provide more representative guidelines for the length of samples. In the SUGAR study (Pavelko & Owens, 2017), a 30-min sample was also elicited, but only the first 50 utterances were analyzed for the reason of clinical relevance and generalizability of the guidelines across age categories. The authors did not suggest a guideline regarding the length in minutes that would yield an interaction containing 50 child utterances. Considering the NDW/m results, this study may concur with the suggestion of 30-min samples, as only one cohort's NDW/m score indicated a significant difference when compared to subsequent intervals.

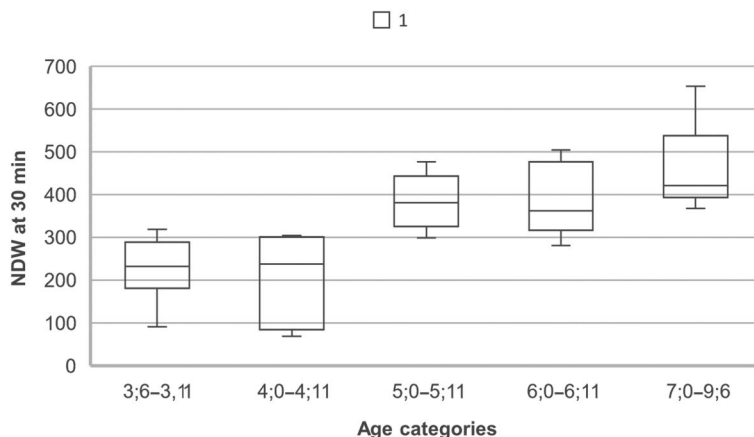
Further support of this study's findings that samples longer than 50 utterances are necessary to represent spoken language skills can be found in the report by Oosthuizen and Southwood (2009). These authors calculated MLU-w using the alternate method employed by this study and recommended that SLPs use samples of at least 100 utterances as opposed to the traditionally recommended 50 utterances (Oosthuizen & Southwood, 2009).

**Table 4.** Descriptive statistics per group at 30 min.

Measure	3;6–3;11		4;0–4;11		5;0–5;11		6;0–6;11		7;0–9;6	
	<i>Mdn</i> (IQR)	<i>M</i> (SD)	<i>Mdn</i> (IQR)	<i>M</i> (SD)	<i>Mdn</i> (IQR)	<i>M</i> (SD)	<i>Mdn</i> (IQR)	<i>M</i> (SD)	<i>Mdn</i> (IQR)	<i>M</i> (SD)
MLU-w	5.32 (2.67)	4.99 (1.84)	4.21 (1.94)	4.26 (1.63)	8.23 (6.30)	8.50 (3.98)	8.66 (5.59)	7.89 (3.06)	7.74 (1.63)	7.79 (0.95)
NDW	231.50 (108.25)	226.89 (77.51)	237.50 (217.00)	206.17 (104.10)	381.50 (118.00)	384.17 (70.73)	362.17 (160.00)	384.39 (84.91)	421.00 (144.75)	460.50 (104.46)
PP	26.50 (3.875)	26.92 (2.01)	27.00 (2.00)	27.00 (1.67)	29.75 (2.25)	29.08 (1.28)	29.50 (1.63)	29.25 (0.88)	29.00 (0.63)	28.92 (0.80)

*Note.* IQR denotes interquartile range, and median, as a measure of central tendency for non-normal distributed data, denotes the middle number in the data set. PP indicates the scores of the Prutting and Kirchner Pragmatic Protocol. MLU-w = mean length of utterance in words; NDW = number of different words.

**Figure 1.** Boxplots representing the minimum, first quartile, median, third quartile, and maximum of number of different words (NDW) scores at 30 min for each age group.



The method used in the current pilot study yielded more than 100 utterances in a 30-min sample and found that a sample duration of 20 min is sufficient for MLU-w calculation. However, for assessing the content of language, the NDW/m measure seems to suggest that a 30-min sample would be more representative. A guideline in terms of the length of interaction in minutes is more clinically relevant than one specifying the number of utterances, while also providing the most representative sample in terms of typical interaction behavior and naturalistic language skills.

### The Spoken Language Skills of Typically Developing Afrikaans-Speaking Children

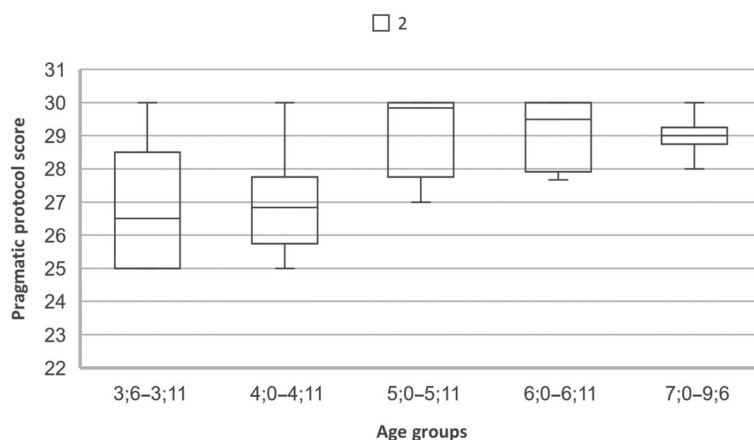
The results obtained from this study shows promise for the use of LSA to describe developmental trajectories for spoken language skills (Bowles et al., 2020; Heilmann et al., 2010; Manning et al., 2020). Preliminary developmental data

for some encompassing linguistic parameters, one for each language domain (i.e., language form, content, and use), were obtained. Age-related changes were noted in all the LSA measures (MLU-w, NDW, and PP) included in this study.

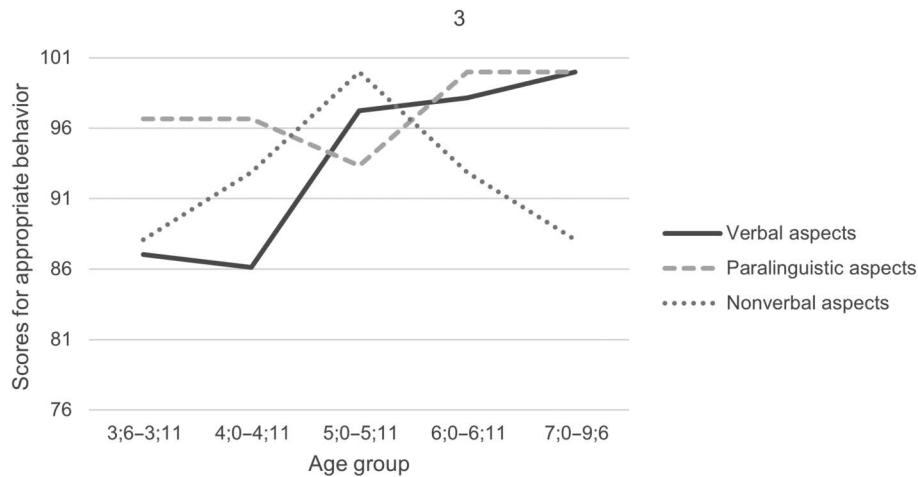
### Morphosyntactic Skills

Using MLU-w to describe morphosyntactic skills has long been recognized as a reliable measure of overall language ability (Heilmann et al., 2010; Manning et al., 2020; Pavelko & Owens, 2017). This study used the alternate method of MLU calculation in words (MLU-w) to quantify and describe the children's morphosyntactic skills. The results, although not statistically significant, show that there is a tendency for the number of utterances to increase with age (see Table 3). The current results may not support MLU-w as a robust indicator for Afrikaans morphosyntactic skills; however, due to the small sample

**Figure 2.** Boxplots representing the minimum, first quartile, median, third quartile, and maximum of pragmatics protocol scores for each age category.



**Figure 3.** Percentage scores for appropriate behaviors in each aspect of the Pragmatic Protocol (Puttling & Kirchner, 1987).



included in this study, further investigation is necessary to determine the robustness of MLU-w. For all age groups except one, the age-related developmental trajectory of MLU-w in the current pilot study was consistent with the existing preliminary data for Afrikaans (Oosthuizen & Southwood, 2009) as well as with normative data from English (Pavelko & Owens, 2017). The 5-year-old cohort performed better than expected from previously published data.

### Phonology

The literature has provided some evidence regarding the development of speech sound production in Afrikaans-speaking children (Geertsema, 2016; Lotter, 1974). Informal descriptions of phonological and articulation errors have been compared with the age-of-acquisition guidelines (Geertsema, 2016). Accurate production of the /s/ and voiced alveolar trill /r/ sound is only expected at the ages of 6–7 years (Geertsema, 2016). The results obtained in this study are aligned with the expected mastery of these sounds, except in the case of one participant in the oldest age category (i.e., GES 7;0–9;6).

The error patterns noted for Afrikaans-speaking children decreased only slightly later than indicated by the norms for English age-matched peers (Bowen, 1998). For example, according to English norms, final consonant deletion should diminish by age 3;3, whereas according to the current results, it is still present up to age 4;1. This finding provides preliminary evidence that phonological patterns in Afrikaans may differ slightly from those of other languages, which highlights the need for future research to report on developmental phonological data for this population.

### Semantic Skills

Semantic skills could also be analyzed using LSA by calculating the NDW, a commonly used measure for

lexical diversity (Ebert, 2020; Ebert & Scott, 2014; Imgrund et al., 2019; Pavelko & Owens, 2017). Familiarity with the context yields greater lexical diversity and semantic complexity, as seen when children talk about their own experiences (Channell et al., 2018; Squires et al., 2020). The initial results of this study support this observation as the NDW/m measure decreased, although not significantly, as time progressed and unfamiliar activities were introduced. When considering the clinical utility of the LSA procedures described here, it is important to take into account the influence of sample length on the opportunity for children to produce language representative of their spoken language skills. Bearing this in mind, these findings support the use of a sample of at least 30 min.

To the best of the researchers' knowledge, no NDW data for Afrikaans are currently available. When comparing the results of the different age categories in the current pilot study, the steady development of semantic skills and refined narrative skills is evident (Owens, 2016). The NDW scores increase for each age group, which indicates the growth of vocabulary throughout these chronological ages and therefore confirms the usefulness of this measure as a metric of spoken language skill (Charest et al., 2020).

### Pragmatics

LSA also enabled the researchers to explore the participants' use of language. The current preliminary results indicated age-related development and mastery of pragmatic skills before the age of 10 years, as reported before (Gentileau-Lambin et al., 2019). The results show steady increases in pragmatic skills with improved verbal aspects, such as lexical specificity and accuracy, with age. This is most likely due to maturation and development of more sophisticated pragmatic skills and vocabulary development. The paralinguistic aspects of unintelligibility and

disfluency lessened with age, as fewer phonological processes and developmental disfluencies were observed. Nonverbal aspects such as foot/leg and hand/arm movement became more appropriate in the older age groups, while eye gaze also became more consistent and appropriate with age. However, the impeded observability of physical proximity due to the setting at a tabletop influenced scores regarding nonverbal aspects in the older groups (ages 6;0–9;6).

The current findings concur with previous results that the conversational skills of preschool children improve with age, and LSA also allows the investigation of narrative pragmatics skills that develop mainly at school age (Gentileau-Lambin et al., 2019). The increased and appropriate intentional use of facial expressions confirmed the increasingly sophisticated narrative pragmatics skills.

The raters of this study observed the behaviors at random intervals throughout the recording to ensure comprehensive observation of pragmatic skills, and time-related improvements in eye gaze and turn-taking were observed. It is advised that SLPs score pragmatic behaviors at random intervals of a recorded interaction, ensuring that not only the initial or the last part of interaction is included (Owens, 2016). The comprehensive assessment potential of LSA is highlighted in its capacity to provide a naturalistic context for reliable assessment of language use.

## Limitations and Future Research

Although this study met its dual aim of providing a preliminary description of the spoken language skills of typically developing Afrikaans-speaking children using LSA and providing length-of-sample guidelines for LSA, some limitations should be mentioned. The most notable limitation was the relatively small sample size, typical of pilot studies, which means that the results obtained are not representative of the entire population that was investigated. Only six participants per age category could be included, which negatively affected the power of the results. To compute the achieved power of the statistical tests, the level of significance ( $\alpha = 0.05$ ), the sample size ( $n = 6$ ), and the effect size are required. Using Cohen's (Cohen, 1969) recommendation for detecting moderate to large effect sizes, the achieved power equals 0.559 and 0.358 for the WSR and MW tests, respectively. However, as a pilot study, the aim of the study is not to control for the Type I error.

Furthermore, the sample only represents children with middle-to-high SES, whose language skills may not be comparable to those of peers with low SES. Afrikaans is further known to have multiple dialects, and this study only focused on children from one geographical area (i.e., Tshwane), which may limit the generalizability of the data to other Afrikaans dialects.

The utterance separation guidelines were found to be limited and subjective. Although SUGAR procedures address utterance boundaries, variability in TNU was noted between the raters and may have affected MLU-w scores.

The length of sample has been proven to influence NDW scores (Charest et al., 2020). Using moving-average type–token ratio (MATTR) may be a more reliable measure of semantic skills. Although Charest et al. (2020) found that different measures of lexical diversity are appropriate for different clinical purposes, future research may provide developmental data for MATTR and determine the reliability of both NDW and MATTR for the Afrikaans population.

The pragmatics analysis for this study was conducted using video recordings of the interactions (postevent analysis) to enable the researchers to calculate interrater reliability. However, the feasibility and reliability of a real-time analysis may be explored in future research in an attempt to reduce the time needed to analyze the samples, while capitalizing on the comprehensive capabilities of LSA to analyze language use.

Further research with larger samples should be conducted to provide reliable and representative developmental data. The inclusion of formal phonology measures in LSA procedures may further increase the potential usefulness of the method as a clinical tool.

## Conclusions

The study concluded that interactions with a 30-min duration provide a representative language sample considering language form, content, and use for children between the ages of 3;6 and 9;6. The guidelines for the collection, transcription, and analysis of language samples, using SUGAR procedures adjusted for Afrikaans, should be carefully followed to ensure the reliability of the samples when the procedure is replicated in practice. The study revealed urgent gaps in the literature regarding the Afrikaans language assessment and development, which may draw the attention of other experts in the field to these challenges often encountered in nonmainstream languages. The current pilot study obtained promising preliminary developmental data and clinical guidelines that hold potential for the future of fair language assessment and the reliable use of LSA as a method to describe spoken language skills in nonmainstream languages.

## Data Availability Statement

Data and results are available via the University of Pretoria Data Repository. Access may be requested from the authors via the repository.

## Acknowledgments

Compiling a child speech database for the South African context: Speech samples of typically developing Afrikaans and Sesotho sa Leboa-speaking children were made possible with the support from the South African Centre for Digital Language Resources (SADiLaR). SADiLaR is a research infrastructure established by the Department of Science and Technology of the South African government as part of the South African Research Infrastructure Roadmap. It was further supported by the “Suid-Afrikaanse Akademie vir Wetenskap en Kuns” (South African Academy for Science and Art). The authors wish to thank the following individuals for their invaluable contributions to this publication: Klarissa Nel, Annemarie Horn, Alex Winter, Maria du Toit, and Elsie Naudé. This study formed part of the first author’s dissertation for the Masters in Speech-Language Pathology degree at the University of Pretoria.

## References

- Adams, C. (2002). Practitioner review: The assessment of language pragmatics. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43(8), 973–987. <https://doi.org/10.1111/1469-7610.00226>
- Akoglu, H. (2018). User’s guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3), 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>
- Alvesson, M., & Kärreman, D. (2011). Decolonizing discourse: Critical reflections on organizational discourse analysis. *Human Relations*, 64(9), 1121–1146. <https://doi.org/10.1177/0018726711408629>
- Bakker, M., & Wicherts, J. M. (2014). Outlier removal, sum scores, and the inflation of the type I error rate in independent samples *t* tests: The power of alternatives and recommendations. *Psychological Methods*, 19(3), 409–427. <https://doi.org/10.1037/met0000014>
- Barratt, J., Khoza-Shangase, K., & Msimang, K. (2012). Speech-language assessment in a linguistically diverse setting: Preliminary exploration of the possible impact of informal “solutions” within the south African context. *South African Journal of Communication Disorders*, 59(1), 34–44. <https://doi.org/10.4102/sajcd.v59i1.20>
- Bedford, H., Walton, S., & Ahn, J. (2013). *Measures of child development: A review*. UCL Institute of Child Health. <https://discovery.ucl.ac.uk/1521166/>
- Bliss, L. S., McCabe, A., & Miranda, A. E. (1998). Narrative assessment profile. *Journal of Communication Disorders*, 31(4), 347–363. [https://doi.org/10.1016/S0021-9924\(98\)00009-4](https://doi.org/10.1016/S0021-9924(98)00009-4)
- Bornman, J., Ronski, M., Tönsing, K., Sevcik, R., White, R., Barton-Hulsey, A., & Morwane, R. (2018). Adapting and translating the Mullen scales of early learning for the south African context. *South African Journal of Communication Disorders*, 65(1), e1–e9. <https://doi.org/10.4102/sajcd.v65i1.571>
- Bowen, C. (1998). *Developmental phonological disorders: A practical guide for families and teachers*. ACER Press.
- Bowles, C. (2018). *Predicting expressive vocabulary change in young children growing up in Ireland*. University College Cork.
- Bowles, R. P., Justice, L. M., Khan, K. S., Piasta, S. B., Skibbe, L. E., & Foster, T. D. (2020). Development of the narrative assessment protocol-2: A tool for examining young children’s narrative skill. *Language, Speech, and Hearing Services in Schools*, 51(2), 390–404. [https://doi.org/10.1044/2019\\_LSHSS-19-00038](https://doi.org/10.1044/2019_LSHSS-19-00038)
- Brothers, K. B., Page Glascoe, F., & Robertshaw, N. S. (2008). PEDS: Developmental milestones—An accurate brief tool for surveillance and screening. *Clinical Pediatrics*, 47(3), 271–279. <https://doi.org/10.1177/0009922807309419>
- Casby, M. W. (2011). An examination of the relationship of sample size and mean length of utterance for children with developmental language impairment. *Child Language Teaching and Therapy*, 27(3), 286–293. <https://doi.org/10.1177/0265659010394387>
- Chambers, M., Bliss, K., & Rambur, B. (2020). Recruiting research participants via traditional snowball vs facebook advertisements and a website. *Western Journal of Nursing Research*, 42(10), 846–851. <https://doi.org/10.1177/0193945920904445>
- Channell, M. M., Loveall, S. J., Conners, F. A., Harvey, D. J., & Abbeduto, L. (2018). Narrative language sampling in typical development: Implications for clinical trials. *American Journal of Speech-Language Pathology*, 27(1), 123–135. [https://doi.org/10.1044/2017\\_AJSLP-17-0046](https://doi.org/10.1044/2017_AJSLP-17-0046)
- Charest, M., Skoczylas, M. J., & Schneider, P. (2020). Properties of lexical diversity in the narratives of children with typical language development and developmental language disorder. *American Journal of Speech-Language Pathology*, 29(4), 1866–1882. [https://doi.org/10.1044/2020\\_AJSLP-19-00176](https://doi.org/10.1044/2020_AJSLP-19-00176)
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. Academic Press.
- De Lamo White, C., & Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language & Communication Disorders*, 46(6), 613–627. <https://doi.org/10.1111/j.1460-6984.2011.00049.x>
- Ebert, K. D. (2020). Language sample analysis with bilingual children. *Topics in Language Disorders*, 40(2), 182–201. <https://doi.org/10.1097/tld.0000000000000209>
- Ebert, K. D., & Scott, C. M. (2014). Relationships between narrative language samples and norm-referenced test scores in language assessments of school-age children. *Language, Speech, and Hearing Services in Schools*, 45(4), 337–350. [https://doi.org/10.1044/2014\\_LSHSS-14-0034](https://doi.org/10.1044/2014_LSHSS-14-0034)
- ELAN (Version 6.1) [Computer software]. (2021). *Nijmegen: Max Planck Institute for Psycholinguistics*. <https://archive.mpi.nl/tla/elan>
- Farahani, M. V. (2013). The role of discourse analysis in translation. *International Journal of Applied Linguistics and English Literature*, 2(1), 112–116. <https://doi.org/10.7575/ijalel.v.2n.1p.112>
- Gavin, W. J., & Giles, L. (1996). Sample size effects on temporal reliability of language sample measures of preschool children. *Journal of Speech and Hearing Research*, 39(6), 1258–1262. <https://doi.org/10.1044/jshr.3906.1258>
- Geertsema, S. (2016). *Application of principles of motor learning in the treatment of speech sound disorders*. [Unpublished doctoral dissertation]. University of Pretoria.
- Gentileau-Lambin, P., Nicli, J., Richard, A. F., Macchi, L., Barbeau, C., Nguyen, S., Medjkane, F., & Lemaître, M. P. (2019). Assessment of conversational pragmatics: A screening tool for pragmatic language impairment in a control population of children aged 6–12 years. *Archives de Pédiatrie*, 26(4), 214–219. <https://doi.org/10.1016/j.arcped.2019.03.004>

- Glascocoe, F. P.** (2013). *Parents' Evaluation of Developmental Status: Developmental Milestones (PEDS:DM)*. <http://www.pedstest.com>
- Govindarajan, K., & Paradis, J.** (2019). Narrative abilities of bilingual children with and without developmental language disorder (SLI): Differentiation and the role of age and input factors. *Journal of Communication Disorders*, 77, 1–16. <https://doi.org/10.1016/j.jcomdis.2018.10.001>
- Heilmann, J., Nockerts, A., & Miller, J. F.** (2010). Language sampling: Does the length of the transcript matter? *Language, Speech, and Hearing Services in Schools*, 41(4), 393–404. [https://doi.org/10.1044/0161-1461\(2009/09-0023\)](https://doi.org/10.1044/0161-1461(2009/09-0023))
- Hux, K., Sanger, D., Reid, R., & Maschka, A.** (1997). Discourse analysis procedures: Reliability issues. *Journal of Communication Disorders*, 30(2), 133–150. [https://doi.org/10.1016/S0021-9924\(96\)00060-3](https://doi.org/10.1016/S0021-9924(96)00060-3)
- Imgrund, C. M., Loeb, D. F., & Barlow, S. M.** (2019). Expressive language in preschoolers born preterm: Results of language sample analysis and standardized assessment. *Journal of Speech, Language, and Hearing Research*, 62(4), 884–895. [https://doi.org/10.1044/2018\\_JSLHR-L-18-0224](https://doi.org/10.1044/2018_JSLHR-L-18-0224)
- Kathard, H., & Pillay, M.** (2013). Promoting change through political consciousness: A south African speech-language pathology response to the world report on disability. *International Journal of Speech-Language Pathology*, 15(1), 84–89. <https://doi.org/10.3109/17549507.2012.757803>
- Kramer, C. A., James, S. L., & Saxman, J. H.** (1979). A comparison of language samples elicited at home and in the clinic. *The Journal of Speech and Hearing Disorders*, 44(3), 321–330. <https://doi.org/10.1044/jshd.4403.321>
- Leedy, P. D., & Ormrod, J. E.** (2020). *Practical research: Planning and design*. Pearson.
- Liebenberg, P., Rabé, M., Bornman, J., De Wet, F., Graham, M. A., & Van der Linde, J.** (in press). Code switching patterns in natural spoken language in a group of young Afrikaans-speaking children: An exploratory study. In *Suid Afrikaanse Akademie vir Wetenskap en Kuns: Tydskrif vir Geesteswetenskappe* [South African Academy of Science and Art: Journal of Humanities].
- Lotter, E. C.** (1974). 'n Ondersoek na die ontwikkeling van artikulasie by die Afrikaanssprekende kind met die oog op die opstel van 'n geskikte artikulasietoets [An investigation of the development of articulation skills in Afrikaans-speaking children with the goal of developing an appropriate articulation test]. [MA dissertation, University of Pretoria]. <https://repository.up.ac.za/handle/2263/29411?show=full>
- Mahomed-Asmail, F., Swanepoel, D., Eikelboom, R. H., Myburgh, H. C., & Hall, J., III.** (2016). Clinical validity of hearScreen smartphone hearing screening for school children. *Ear and Hearing*, 37(1), e11–e17. <https://doi.org/10.1097/AUD.0000000000000223>
- Manning, B. L., Harpole, A., Harriott, E. M., Postolowicz, K., & Norton, E. S.** (2020). Taking language samples home: Feasibility, reliability, and validity of child language samples conducted remotely with video chat versus in-person. *Journal of Speech, Language, and Hearing Research*, 63(12), 3982–3990. [https://doi.org/10.1044/2020\\_JSLHR-20-00202](https://doi.org/10.1044/2020_JSLHR-20-00202)
- Moonsamy, S., & Kathard, H.** (2015). *Speech-language therapy in a school context: Principles and practices* (1st ed.). Van Schaik Publishers.
- Mowbray, F. I., Fox-Wasylyshyn, S. M., & El-Masri, M. M.** (2019). Univariate outliers: A conceptual overview for the nurse researcher. *Canadian Journal of Nursing Research*, 51(1), 31–37. <https://doi.org/10.1177/0844562118786647>
- Nicklin, C., & Plonsky, L.** (2020). Outliers in L2 research in applied linguistics: A synthesis and data re-analysis. *Annual Review of Applied Linguistics*, 40, 26–55. <https://doi.org/10.1017/S0267190520000057>
- Oosthuizen, H., & Southwood, F.** (2009). Methodological issues in the calculation of mean length of utterance. *The South African Journal of Communication Disorders*, 56(1), 76–87. <https://doi.org/10.4102/sajcd.v56i1.194>
- Owens, R. E.** (2016). *Language development: An introduction* (9th ed.). Pearson.
- Pascoe, M., Rogers, C., & Norman, V.** (2013). Are we there yet? On a journey towards more contextually relevant resources in speech-language therapy and audiology. *The South African Journal of Communication Disorders*, 60(1), 2–9. <https://doi.org/10.7196/sajcd.256>
- Pavelko, S. L., & Owens, R. E.** (2017). Sampling Utterances and Grammatical Analysis Revised (SUGAR): New normative values for language sample analysis measures. *Language, Speech, and Hearing Services in Schools*, 48(3), 197–215. [https://doi.org/10.1044/2017\\_LSHSS-17-0022](https://doi.org/10.1044/2017_LSHSS-17-0022)
- Pavelko, S. L., Owens, R. E., Ireland, M., & Hans-Vaughn, D. L.** (2016). Use of language sample analysis by school-based SLPs: Results of a nationwide survey. *Language, Speech, and Hearing Services in Schools*, 47(3), 246–258. [https://doi.org/10.1044/2016\\_LSHSS-15-0044](https://doi.org/10.1044/2016_LSHSS-15-0044)
- Pezold, M. J., Imgrund, C. M., & Storkel, H. L.** (2020). Using computer programs for language sample analysis. *Language, Speech, and Hearing Services in Schools*, 51(1), 103–114. [https://doi.org/10.1044/2019\\_LSHSS-18-0148](https://doi.org/10.1044/2019_LSHSS-18-0148)
- Prutting, C. A., & Kirchner, D. M.** (1987). A clinical appraisal of the pragmatic aspects of language. *The Journal of Speech and Hearing Disorders*, 52(2), 105–119. <https://doi.org/10.1044/jshd.5202.105>
- Saenz, T. I., & Huer, M. B.** (2003). Testing strategies involving least biased language assessment of bilingual children. *Communication Disorders Quarterly*, 24(4), 184–193. <https://doi.org/10.1177/15257401030240040401>
- Shipley, K. G., & McAfee, J. G.** (2016). *Assessment in speech-language pathology* (5th ed.). Cengage Learning.
- South African Department of Basic Education.** (2010). The status of the language of learning and teaching (LOLT) in South African public schools: A quantitative overview. <https://www.education.gov.za/Portals/0/Documents/Reports/StatusofLOLT.pdf?ver=2011-03-30-231358-000>
- Southwood, F.** (2005). A comparison of the responses to three comprehension and three production tasks assessing the morpho-syntactic abilities of Afrikaans-speaking preschoolers. *Per Linguam*, 21(1), 36–59. <https://doi.org/10.5785/21-1-71>
- Southwood, F.** (2013). Towards a dialect-neutral assessment instrument for the language skills of Afrikaans-speaking children: The role of socioeconomic status. *Journal of Child Language*, 40(2), 415–437. <https://doi.org/10.1017/S0305000912000037>
- Southwood, F., & Russell, A. F.** (2004). Comparison of conversation, freeplay, and story generation as methods of language sample elicitation. *Journal of Speech, Language, and Hearing Research*, 47(2), 366–376. [https://doi.org/10.1044/1092-4388\(2004/030\)](https://doi.org/10.1044/1092-4388(2004/030))
- Southwood, F., & Van Dulm, O.** (2015). The challenge of linguistic and cultural diversity: Does length of experience affect south African speech-language therapists' management of children with language impairment? *South African Journal of Communication Disorders*, 62(1), E1–E14. <https://doi.org/10.4102/sajcd.v62i1.71>
- Spencer, E., Bryant, L., & Colyvas, K.** (2020). Minimizing variability in language sampling analysis. *Topics in Language Disorders*, 40(2), 166–181. <https://doi.org/10.1097/TLD.0000000000000212>

- Squires, B., Bird, E. K. R., & Cahill, P.** (2020). "Undating" and "quarter ponies": Comparing the use of derived words and compounds across discourse types and age groups. *Language, Speech, and Hearing Services in Schools*, 51(3), 761–777. [https://doi.org/10.1044/2020\\_LSHSS-19-00076](https://doi.org/10.1044/2020_LSHSS-19-00076)
- Statistics South Africa.** (2012). Census 2011: Census in brief. [http://www.statssa.gov.za/census/census\\_2011/census\\_products/Census\\_2011\\_Census\\_in\\_brief.pdf](http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf)
- Streiner, D. L.** (2018). Statistics commentary series. *Journal of Clinical Psychopharmacology*, 38(3), 170–171. <https://doi.org/10.1097/JCP.0000000000000865>
- Tilstra, J., & McMaster, K.** (2007). Productivity, fluency, and grammaticality measures from narratives. *Communication Disorders Quarterly*, 29(1), 43–53. <https://doi.org/10.1177/1525740108314866>
- Tomblin, J. B., Oleson, J. J., Ambrose, S. E., Walker, E., & Moeller, M. P.** (2014). The influence of hearing aids on the speech and language development of children with hearing loss. *JAMA Otolaryngology–Head & Neck Surgery*, 140(5), 403–409. <https://doi.org/10.1001/jamaoto.2014.267>
- Tommerdahl, J., & Kilpatrick, C. D.** (2014). The reliability of morphological analyses in language samples. *Language Testing*, 31(1), 3–18. <https://doi.org/10.1177/0265532213485570>
- van der Linde, J.** (2015). *Developmental screening and communication delays in infants: A South African primary health care perspective*. [Doctoral dissertation]. University of Pretoria. <https://repository.up.ac.za/handle/2263/53474?show=full>
- van der Merwe, M. N., Mosca, R., Swanepoel, D. W., Glascoe, F. P., & van der Linde, J.** (2017). *Early detection of developmental delays in vulnerable children by community care workers using an mHealth tool*. [MA dissertation]. University of Pretoria. <https://repository.up.ac.za/handle/2263/71437>
- van Dulm, O.** (2007). *The grammar of English-Afrikaans code switching: A feature checking account*. [Doctoral dissertation]. Radboud Universiteit Nijmegen. <http://www.lotpublications.nl/publish/articles/002423/bookpart.pdf>
- Van Dulm, O., & Southwood, F.** (2014). Child language assessment and intervention in multilingual and multicultural South Africa: Findings of a national survey. *Stellenbosch Papers in Linguistics*, 42(1), 55–76. <https://doi.org/10.5774/42-0-147>
- Verdon, S., McLeod, S., & Wong, S.** (2015). Supporting culturally and linguistically diverse children with speech, language and communication needs: Overarching principles, individual approaches. *Journal of Communication Disorders*, 58, 74–90. <https://doi.org/10.1016/j.jcomdis.2015.10.002>
- Villarroel, J. D., Miñón, M., & Nuño, T.** (2011). The origin of counting: A study of the early meaning of "one", "two" and "three" among Basque- and Spanish-speaking children. *Educational Studies in Mathematics*, 76(3), 345–361. <https://doi.org/10.1007/s10649-010-9291-0>
- Wale, V. D., Sthool, V. A., Balagond, S., Jadhav, J. D., & Upadhye, S. K.** (2020). Analysis of rainfall and rainy days trend in Shirala Tahsil of Sangli district of Maharashtra (India). *Journal of Pharmacognosy and Phytochemistry*, 9(5), 314–317.
- Wildsmith-Cromarty, R., & Balfour, R. J.** (2019). Language learning and teaching in South African primary schools. *Language Teaching*, 52(3), 296–317. <https://doi.org/10.1017/S0261444819000181>



---

## Appendix

### Conventions and Adaptations to Sampling Utterances and Grammatical Analysis Revised (SUGAR; Pavelko & Owens, 2017)

---

Adaptation	Theoretical justification
The SUGAR procedures state that transcription should stop at 50 utterances; however, to address the secondary aim of the study, the whole interaction was transcribed, as different lengths of transcriptions were analyzed and compared.	The transcribed sample was analyzed at 5, 10, 20, 30, 40, and 60 min and the total number of words (TNW) and total number of utterances (TNU) were counted for each time segment. The measures at each different segment were compared to suggest guidelines for length of sample.
Contractions were transcribed as the child used them.	As per SUGAR analysis procedures, all contractions were transcribed the way they were used. These were counted as one word.
All personal information, such as names, was removed from the transcribed sample after analysis, as required by the conditions of ethics approval.	The removal of personal information was necessary to maintain the confidentiality of the participants. All personal information was indicated with a preceding hash (#xxx) and removed from samples after analyses were conducted.
All single-morpheme utterances were removed from the transcription.	To ensure greater sensitivity of mean length of utterance (MLU) measures (Oosthuizen & Southwood, 2009) and to ensure an accurate reflection of the child's linguistic abilities (Oosthuizen & Southwood, 2009) the single-morpheme utterances "yes" and "no" were not included in the transcription. Although yes/no questions were avoided throughout the data collection process (Channell et al., 2018; Pavelko & Owens, 2017), it remains part of the typical discourse and could therefore not be completely excluded from the raw samples.
Code-switching refers to the practice of switching between two or more languages during discourse, and is a common phenomenon in South African conversations (van Dulm, 2007). Also, in the current data, frequent code-switching was noted. It is beyond the scope of this study to investigate this phenomenon and further research is being done (Liebenberg et al., in press).	Code-switching was indicated in the transcriptions using braces ({xx}).
Unintelligible or overlapping speech Onomatopoeia or nonverbal behaviors	This was indicated using block brackets ([xyxy]). Where onomatopoeia or nonverbal behaviors, such as laughing, were used, forward slashes (/xyxy/) indicated this and it did not form part of the word count for number of different words (NDW) or TNW.
Number words	Limited evidence was found regarding the inclusion or exclusion of rote-counting or number words (Villarroel et al., 2011). Counting prompts such as give-N were not used in the elicitation of interactions, although some children rote-counted some objects during play. This counting was transcribed as it was not likely to influence lexical diversity in the NDW calculation significantly.
Orthographic transcription and phonological analyses	The samples were only orthographically transcribed for this study. Although phonetic transcription may provide in-depth insight into the phonological development of Afrikaans-speaking children, it was beyond the scope of this study.

---