

Mothers' Judgement Of The Representativeness Of Their Sons' Language Samples In Relation To Volume Of Language Produced*

Frenette Southwood and Ann F. Russell

Departments of General Linguistics and Speech-Language and Hearing Therapy,
Stellenbosch University, 7600 Stellenbosch, South Africa. E-mail: fs@sun.ac.za

1. Introduction

The analysis of language elicited during naturally occurring situations forms the cornerstone of a child language assessment protocol and is used for the planning and monitoring of intervention (Dunn, Flax, Sliwinski, and Aram 1996; Evans and Miller 1999). A 1997 survey revealed that 85% of speech-language therapists in the United States of America use language samples during language assessment with children (Kemp and Klee 1997). One reason for the frequent use of language samples is the limitations of standardised language tests (cf., amongst others, Hawkins and Spencer 1985). Another reason is the lack of assessment tools, especially culturally fair ones, for clients from nonmainstream groups (Peña, Quinn, and Iglesias 1992; Toronto and Merrill 1983). Consider the situation of Afrikaans, for example: According to the 2001 census results (Statistics South Africa 2003), this language is spoken as mother-tongue by 13% of the South African population (i.e., by 6 million people). Considering that Afrikaans is not widely spoken outside of the country, it could thus be viewed as a nonmainstream language compared to, for instance, English, the language which has the world's third largest number of mother-tongue speakers, viz. 322 million (Grimes 1996). The lack of

assessment tools for clients from nonmainstream groups is exemplified by the fact that there is only one standardised instrument for the assessment of expressive syntax for children who speak Afrikaans, namely the *Toets vir Mondelinge Taalproduksie* (Vorster 1980). However, this test has poor test-retest reliability (according to Vorster himself) and has only been standardised for children in the age group 4 years 6 months to 10 years 5 months. This leaves speech-language therapists who are required to assess the expressive syntactic abilities of Afrikaans-speaking children not represented by the population sample on which the test was standardised with the choice of either devising their own informal assessment tool(s) or making use of a spontaneous language sample.

Several methods of eliciting language samples are discussed in the literature. Freeplay, conversation, and story generation are three prominent methods. According to Wren (1985), both the length and the representativeness of the language sample obtained are important criteria for the selection of an elicitation method. Research with various age groups has been done on the first criterion, namely the length of the language sample obtained by different elicitation methods.

As far back as 1974, Longhurst and Grubb cautioned that different methods of language elicitation could lead to measurable differences in language use. Since this statement was made, several studies have been conducted to compare methods of language elicitation in children. When combining and comparing the results of studies of 3- to 5-year-olds by Atkins and Cartwright (1982); 8- to 9-year olds by Evans and Craig (1992); 4- to 5-year-olds by Stalnaker and Creaghead (1982); and 6-year-olds by Wren (1985), the methods eliciting the most to the least language are: Conversation employing picture interpretation during which open-ended questions were asked, conversation during which questions were asked in an interview format, story-retelling after the story was read to the child, and freeplay consisting of researcher-child interaction with toys of the child's choice. Furthermore, freeplay with puppets was found to elicit more language than a story generation task in which a picture series was used as a prompt. To date, no single study has compared all three methods of language sample elicitation (i.e., freeplay, conversation, and story generation) with preschool children, despite the fact that these methods are popular and are employed by many speech-language therapists with clients from this age group.

For a judgement on Wren's (1985) second criterion, namely that of the representativeness of the obtained language sample, speech-language therapists frequently rely on the caregiver. The question arises as to whether caregivers are able to give a judgement on the representativeness of a language sample, judgements which correlate with quantitative measures, such as the length of the sample. For children in preschool, who are often less reluctant to interact with strangers than are younger children, the speech-language therapist often chooses to elicit the language sample instead of requesting the caregiver to do so. This means that the speech-language therapist needs the caregiver to judge the representativeness of a language sample when the caregiver was merely an observer of the elicitation of that sample. The question is whether the caregiver is able to do so.

According to Thal, O'Hanlon, Clemmons, and Fralin (1999), the status of parent-report has shifted dramatically in recent years. There are at least three advantages of using parent-report, namely that this report is cost-effective, especially when used to substitute formal testing; can be obtained before assessment commences, which enables the speech-language therapist to select certain aspects of the child's language for in-depth assessment; and reflects the child's behaviour in a number of contexts apart from the clinical setting, providing a more representative sample of the child's language than would have been obtained in the clinic alone (Thal et al. 1999). The parent-report referred to by Thal et al. (1999) usually takes the form of a completed questionnaire. If the report takes the form of a judgement, the second and third advantages do not apply, but the first still does. Although parents may be viewed as lacking adequate training for accurate reporting and as being biased in their impressions of their own children, the advantages of making use of parent-report outweigh these disadvantages.

Recently, several studies (e.g., Klee, Pearce, and Carson 2000; Ratner and Silverman 2000; Thal et al. 1999; Thal, Jackson-Maldonado, and Acosta 2000) have been conducted comparing parents' reports of their children's language development with the children's scores on standardised tests and mean length of utterance (MLU) or other behavioural measures. Thal et al. (1999) compared the responses of parents on the MacArthur Communicative Development Inventory: Words and Sentences form (Fenson et al. 1993)

with a language sample, the results of Gardner's (1979) Expressive One-Word Picture Vocabulary Test, and the results of the Memory for Sentences subtest of the Stanford Binet Intelligence Scale (Thorndike, Hagen, and Sattler 1986). Thal et al. (1999) also correlated parents' reports on the MacArthur Communicative Development Inventory: Words and Gestures form with the results of the Preschool Language Scale-Revised (Zimmerman, Steiner, and Pond 1979). More recently, Thal and colleagues (Thal et al. 2000) compared the results of parent-report in the form of the Fundación MacArthur: Inventario del Desarrollo de Habilidades Comunicativas (Jackson-Maldonado, Bates, and Thal 1992) with behavioural measures such as a naming task and a language sample obtained from Spanish-speaking toddlers. Ratner and Silverman (2000) compared the results of four standardised articulation and language tests and a spontaneous language sample with those of a parent-report in the forms of the Speech and Language Assessment Scale (Hadley and Rice 1993) and the CDI-Toddler on children at the onset of stuttering. Lastly, Klee et al. (2000) refined a criterion for language delay used in a previous study (Klee et al. 1998) so as to reduce the number of overreferrals, and found the predictive value of the screening score, based on parent-report, to be good.

While the general finding of these studies was that parent-report is a valid measure of children's language abilities, the studies employed either a questionnaire or an inventory. The parents were not required to indicate whether a specific language sample was typical of their children's language use, nor to compare samples of their children's language. Therefore, these studies did not determine whether parents could give information to speech-language therapists on whether the language samples obtained from their children were representative of their everyday language usage, nor whether parents were able to assist the speech-language therapist in selecting the best sample from a range of samples. Due to the lack of standardised instruments for assessing the expressive language abilities of many language groups, including Afrikaans, and the fact that those instruments that are available often lack cross-cultural validity (Craig and Washington 2000; Toronto and Merrill 1983), speech-language therapists often have to make use of nonstandardised means, such as language sampling and analysis, to obtain information on such language abilities of clients from these groups. In these cases, language sampling does not supplement standardised testing but substitutes it. As the results of language sample

analysis will then be the main, if not only, source of information of the Afrikaans-speaking client's expressive language abilities, it is imperative for accurate diagnosis that the speech-language therapist ascertains whether the language sample analysed is, in fact, representative of the client's expressive language abilities.

In a recent study, Southwood and Russell (2004) determined the number of utterances, variety of syntactic structures, length of utterance, number of syntactic errors, and number of complex syntactic utterances elicited by three methods of language elicitation, namely conversation, freeplay, and story generation, from ten 5-year-old, Afrikaans-speaking children. To limit the number of variables, these children were all male. Males were selected instead of females, as there is considerable evidence that being male nearly doubles the risk of language disorder in preschool years (Lassman, Fisch, Vetter, and La Benz 1980; Tomblin et al. 1997), and preschool-aged males should therefore make up more of the speech-language therapist's case-load than should preschool-aged females (cf. Aram and Nation 1980). It was found that freeplay elicited significantly more utterances than did story generation, but a smaller proportion of complex syntactic structures than did conversation and story generation. Furthermore, story generation elicited longer utterances than did conversation or freeplay. What was not reported in this study, was the judgements the mothers made regarding the representativeness of each sample, i.e., whether the mothers could tell the researcher which sample was most representative of the children's language usage.

In the present study, which was a small-scale study, we examined the judgements of these 10 mothers, who were the primary caregivers of their sons, on the representativeness of the three language samples. This was done by calculating the correlation between the judged representativeness of each language sample with two measures of the volume of language elicited. Note that the term "representative" can be used to describe either language that resembles typical performance or language that resembles maximum performance (Wren 1985). Herein lies a conflict: According to Wren (1985:85), if a language sample represents what a child typically does, the speech-language therapist remains unsure as to the upper limits of the child's abilities, but if maximum performance was elicited, the sample may be of little use when planning remediation, as remediation "needs to begin at

the level of language the child uses every day". Because the results of language sample analysis are often used to plan remediation, especially for Afrikaans-speaking clients for which there exists a lack of standardised assessment instruments, in this study, we examined representativeness in terms of typical language production (cf. par. 2.4 for a discussion on the criteria used to determine representativeness).

We set out to answer the following question: Would mothers' judgements on the representativeness of the language elicited correlate in a significant way with the volume of language elicited, as measured in terms of number of utterances and MLU (measured in morphemes)? It was anticipated that there would be a significant correlation for freeplay and conversation but not for story generation, because mothers are presumably more familiar with the language their children produce during freeplay and conversation than during story generation.

2. Method

2.1 Research design

This study had a qualitative design, in which both qualitative and quantitative research methods were used (De Vaus 2001). The task of ranking the language samples intended to obtain ordinal qualitative data (Woodward and Francis 1988) on the mothers' opinions regarding the representativeness of their sons' language, and therefore represents a qualitative research method (Cresswell 1998; Hakim 1987), whereas establishing the correlation between this rank and results of measures of number of utterances and MLU was quantitative (Leedy and Ormrod 2001).

2.2 Participant selection

All the Afrikaans-speaking boys, between 5 years 0 months and 5 years 11 months, from monolingual Afrikaans-speaking homes, who attended one preschool in a northern suburb of Cape Town, were identified as possible participants. We chose Afrikaans-speaking children, because, according to the 1996 census results (Statistics South Africa 2003), this is the home language of the majority (55%) of the inhabitants in the Western Cape, and because of the high probability that the expressive language abilities of an Afrikaans-speaking preschool-aged child will be assessed by means of language sample analysis,

given the shortage of other options available to the speech-language therapist. Boys in preschool were chosen, as it is our clinical experience that boys are referred for school-readiness testing, and thus language assessment, more often than girls (cf. Lyon 1996; Office of Special Education Programs 1998). The fact that only males participated could be viewed as a limitation of the study. All participants were to have normal intelligence, language abilities, and hearing according to the judgements of their class teachers, and no previous referral to or treatment by a speech-language therapist.

Ten boys were randomly selected from the pool of potential participants and the mother of each of these boys telephonically contacted to obtain consent for the inclusion of herself and her son in the study. When consent was withheld, another participant was randomly selected from the pool.

2.3 Participants

Ten typically developing boys between the ages of 5 years 5 months and 5 years 11 months and their mothers participated in the study. The average age of the boys was 5 years 7 and a half months. Every boy was a member of a family with either two or three children. One boy was the oldest sibling in his family, six were the second oldest, and three were the third- and last-born.

The boys were mother-tongue speakers of what is considered standard Afrikaans and were from monolingual Afrikaans-speaking homes. The community that they lived in was, however, Afrikaans-English bilingual. The boys would have been exposed to English through television and possibly through English-speaking or bilingual family members and neighborhood friends.

The mothers of all the boys had been their primary caregivers before the boys started attending preschool. One mother was her son's primary caregiver in the afternoons, while a residential worker acted as caregiver during the mornings prior to his enrolment in preschool.

The mothers who participated in this study were well-educated home-makers. All had completed at least 12 years of formal schooling. Six mothers had received three or more years of tertiary education, whereas one had had one year's tertiary education. All mothers were from the middle- to upper-middle-class, as determined by level of parental education and choice of occupation. Despite the mothers' educational level, they chose not to work out of their homes on a full-time basis, a common phenomenon in their suburb which is regarded as middle- to upper-middle-class.

2.4 Framework of procedures

After we had obtained verbal consent for participation in the study, we sent a letter to the mothers in which we explained the nature of the study and asked for their written consent. The first author, who is a mother-tongue speaker of the same dialect of Afrikaans as the participants and who was unfamiliar to all the participants, performed the rest of the procedures during a one-hour visit to the Hearing and Speech Clinic of a university training hospital. During the visit, each child's written assent was obtained and each mother completed a short case-history form in order to provide us with information on her and her son's background, particularly regarding his language development and hearing ability.

After an otoscopic examination, the hearing sensitivity and middle-ear functioning of all the children were screened using the American Speech, Language, and Hearing Association's guidelines (ASHA 1975, 1979). Hearing sensitivity and middle-ear functioning were found to be essentially within normal limits bilaterally for all participating children.

Three language samples were elicited from each child while his mother judged the representativeness of each sample by rank ordering the elicitation methods (cf. par. 2.5). The mother assigned a rank of 1 to the sample that sounded to her the most representative of her son's everyday language and a rank of 3 to the sample that sounded the least representative. The mother could assign the same rank to two samples if they sounded equally representative to her.

Mothers were to measure representativeness in terms of volume of language elicited, as we predicted that the criteria mothers would use to judge the representativeness of their sons' language samples would correlate with volume, regardless of the researcher's instructions. Therefore, the correlation between the rank and two measures of the volume of language elicited by each of the elicitation methods was calculated, viz. the number and length of the utterances obtained by each method. These two measures were selected by eliminating others, such as structural complexity, number of grammatical errors, and aspects of pragmatic competence.

Structural complexity was not selected, because it is not clear whether parents can differentiate between structurally complex versus long utterances, not even after receiving training on the difference between length and complexity. Parents could perceive long but structurally simple utterances as complex utterances, whereas short but complex utterances could be perceived as simple. In the present study, we wanted to simulate the clinical setting, where the speech-language therapist usually does not have the time to train parents before commencing language sample elicitation. Because we could not say whether the mothers in the present study would be able to differentiate between long utterances and structurally complex utterances, and because we did not want to train the mothers and, by doing so, decrease the generalisability of the results to the clinical situation, we chose not to include structural complexity as one of the measures which mothers would use to decide on the representativeness of the language elicited from their children.

Number of grammatical errors was not chosen as a measure which mothers would use to judge the representativeness of their children's language samples, because researchers are not in agreement on whether or not parents do correct their children's grammatical errors¹. If parents do not, this could indicate that parents do not identify their children's grammatical errors, in which case number of errors would not be used as a measure to judge their children's utterances.

McTear and Conti-Ramsden (1992) discussed a number of protocols which the researcher or speech-language therapist can use when assessing pragmatic competence. These protocols were devised for the speech-language therapist to complete while observing a

client interact with a familiar person. McTear and Conti-Ramsden mentioned difficulties in the implementation of these protocols as well as difficulties in scoring, in particular in deciding on appropriacy of language behaviour. We predicted that, if speech-language therapists experience difficulty in deciding on appropriate language behaviour of a child interacting with a familiar adult, mothers will experience the same or greater difficulty in evaluating their children's interaction with an unfamiliar adult. For this reason, the number of speech acts, turn-taking behaviour, or some other aspect of their children's pragmatic competence was excluded as a criterion which mothers would use to judge the representativeness of their children's language samples.

2.5 Language elicitation methods

In an effort to increase generalisability to speech-language therapist-preschool child interactions during language assessment, the language elicitation took the form of researcher-child interaction (Evans and Craig 1992). Each child was seen alone in a quiet room, while his mother, with his knowledge, observed through a one-way mirror. Observable recorders were used to make audiovisual and audio-cassette recordings of every language elicitation session. Each session lasted 15 minutes and sessions were separated by 5-minute rest periods. The child spent the rest periods with his mother, while the researcher prepared the room for the next session. The order of elicitation methods was randomly assigned to each mother-child pair.

2.5.1 Freeplay

During freeplay, the objects visible to the participants (excluding the video and cassette recorders) were a table, two chairs, and toys. The toys were chosen for their proven popularity with child clients of both genders who had previously visited the clinic and included a toy stove; toy kitchenware (e.g., a pan and cutlery); a two-storey doll's house, with a removable roof and a facade that opened like a door; doll's house furniture, placed next to the doll's house; building blocks in a cardboard container; and six puppets (representing a family). The layout of the room and the arrangements of the toys were the same for all participants.

The freeplay session was initiated by inviting the child to help place the furniture into the doll's house. The rest of the session was conducted according to the guidelines suggested by Crystal, Fletcher, and Garman (1989), which entailed the researcher playing alone with the child in a manner that she considered to be appropriate and natural. If the child was quiet for extended periods, he was prompted with questions such as "And what are you doing?" or "What should happen next?"

2.5.2 Conversation

During conversation, the researcher and child sat next to one another at a table. Questions from the M.W.M. Program for Developing Language Abilities (Minskoff, Wiseman, and Minskoff 1972) were modified and asked to each child in the same order. Following the recommendation of Minskoff et al., the conversation took place over toy telephones and resembled a phone-in competition to make the activity more interesting for the child.

2.5.3 Story generation

During story generation, the researcher and the child were positioned as they were during the conversation session. The method used for story elicitation was that suggested by Peterson and McCabe (1983), who based theirs on that of Labov (1972). The researcher told stories on topics such as pet adventures, accidents, and holidays, whereafter the child was invited to tell about any similar experiences. While generating stories, the researcher and the child jointly coloured in a picture, according to Peterson and McCabe's (1983) suggestion. This was done to put the child at ease and to minimise self-consciousness.

2.6 Data transcription and analysis

The first 15 minutes of each of the three language samples of every child were transcribed orthographically and all utterances occurring in these 15 minutes were considered as part of the sample². The samples were analyzed into utterances following the specifications of Crystal (1979), Crystal et al. (1989) and Garman (1989): an utterance was considered to be a single word, a single phrase, or a clause with its own prosodic identification. Hereafter, the MLU (in morphemes) was calculated by dividing the number of morphemes used in the sample elicited by the total number of utterances elicited.

2.7 Reliability

The reliability of the transcriptions and analyses of the language samples were checked by a final year student in speech-language pathology. The student independently transcribed and analyzed 10% of every language sample. Any differences between the researcher and this student were resolved through discussion and by referring to Crystal (1979). The interjudge reliability for transcription and for analysis was calculated with the following formula: total number of words agreed on divided by total number of words transcribed by researcher multiplied by 100. The interjudge reliability for transcription was 99% and, for analysis, 98%.

2.8 Statistical analysis

A one-way analysis of variance (ANOVA) was performed to determine which method elicited the most utterances and the most words per utterance. No intra-child comparisons were made. Where the overall F was significant, post-hoc pair-wise comparisons were made by using the Bonferonni multiple comparison procedure.

Spearman's rho and Pearson's r were used to calculate the correlation between the rank that the mother assigned to an elicitation method, and the number of utterances and average length of utterance. A significance level of .05 was used for all statistical tests.

3. Results

The appendix contains excerpts of the verbal interaction between the researcher and participants during the three language elicitation methods, as well as the rank the mother assigned to the language sample from which each excerpt was taken. Table 1 shows the number of utterances and MLU elicited from each child participant during conversation, freeplay, and story generation.

Table 1. Raw scores of number of utterances and MLU elicited from each child participant during conversation, freeplay, and story generation

P ^a	Conversation		Freeplay		Story generation	
	No. ^b	MLU	No.	MLU	No.	MLU
1	138	2.500	159	3.547	84	4.595
2	133	3.609	95	3.642	77	5.247
3	144	4.951	202	3.965	117	5.667
4	136	3.081	12	2.167	38	6.290
5	139	3.453	207	4.193	162	4.561
6	151	4.185	113	3.726	98	3.857
7	133	4.399	106	3.594	98	4.449
8	153	3.503	214	4.051	122	4.360
9	149	4.577	131	4.160	144	4.778
10	153	2.837	203	4.507	77	5.857

^aP = Participant number. ^bNo. = Number of utterances.

3.1 Number of utterances

For number of utterances elicited, the difference among methods was significant, $F(2,18) = 4.69$, $p = .023$, and the coefficient of determination (R^2) was 0.640. Pair-wise comparisons showed that the significant difference was between freeplay and story generation $F(1,9) = 7.18$, $p = .025$, with freeplay eliciting significantly more utterances, and not between the number of utterances elicited by conversation and freeplay, $F(1,9) = 0.005$, $p = .950$. Given that the difference between freeplay and story generation was significant, it is interesting to note that the difference between the number of utterances elicited by conversation and story generation was not significant, $F(1,9) = 14.21$, $p = .128$, even though the mean number of utterances elicited by conversation and freeplay were highly comparable ($M = 142.9$ and 144.2 , respectively).

There were very large differences in the standard deviations of the number of utterances elicited by conversation, freeplay, and story generation ($SD = 8.103$, 65.267 , and 36.040 , respectively). For conversation, the method with the smallest standard deviation, the range in number of utterances per participant was the smallest (133 to 153) (cf. Table 2). This

could be attributed to the fact that the conversation elicitation method was very structured, with all participants being asked the same questions in the same order, and that all participants assumed that they had to answer all the questions posed to them. Story generation showed the second highest standard deviation and elicited a range of 38 to 162 utterances. This method was less structured than that of conversation, with the type of prompt used by the researcher ("Has something like that ever happened to you?") eliciting responses ranging from a single word ("yes" or "no") to detailed accounts of the participants' experiences. Freeplay elicited a range of 12 to 214 utterances. During this elicitation method, there was no pressure on the participant to verbalise. Where participants verbalised actively, no prompts were employed. Quieter participants, who were engrossed in the freeplay activity and for whom prompts such as "What are you doing now?" were used, treated these prompts in the same way as they did the other verbalisations of the researcher, namely as utterances that required either no response or a one-utterance response such as "playing" or "moving her here".

Table 2. Means and standard deviations of number of utterances and MLU elicited by conversation, freeplay, and story generation

	Conversation		Freeplay		Story generation	
	Mean	SD	Mean	SD	Mean	SD
No. of Utterances	142.9	8.103	144.2	65.267	101.7	36.040
MLU	3.709	0.798	3.755	0.637	4.966	0.768

3.2 MLU

The difference among methods in terms of elicited MLU was significant, $F(2,18) = 8.32$, $p = 0.003$, $R^2 = 0.558$. Pair-wise comparisons showed that the differences were not between conversation and freeplay $F(1,9) = 0.02$, $p = .878$, but between conversation and story generation, $F(1,9) = 10.67$, $p = .010$, and between story generation and freeplay, $F(1,9) = 10.92$, $p = .009$, with story generation eliciting significantly longer utterances than did conversation or freeplay (also cf. Table 2). The longer MLU elicited by story generation cannot be attributed to multiple conjoining using and or and then, as sentences conjoined by these connecting devices were counted as separate utterances, unless they were used as definite coordinating conjunctions. Rather, the longer MLU elicited by story generation

could be attributed to the type of responses required from the participants by this task and the lack of opportunity to use acceptable elliptical utterances during story generation.

3.3 Correlation between rank assigned and language elicited

There was no significant correlation between the rank assigned to conversation by the mothers, and the number of utterances and the MLU, respectively. There was a significant correlation between the rank that the mothers assigned to the language elicited during freeplay, and the number of utterances, $r = -0.808$, $p = .004$. There was also a significant correlation between the rank that the mothers assigned to the language elicited during freeplay, and the MLU elicited during freeplay, $r = -0.790$, $p = .006$. For story generation, there was no significant correlation between the assigned rank, and the number of utterances and the MLU, respectively (cf. Table 3).

Table 3. Correlation between rank assigned, and number of utterances and MLU elicited

Correlation between rank assigned and	r	%VAR	F	df	p
No. of utterances for conversation	0.072	0.528	0.042	8	.841
MLU for conversation	-0.369	13.630	1.263	8	.293
No of utterances for freeplay	-0.809	65.415	15.132	8	.004
MLU for freeplay	-0.790	62.457	13.309	8	.006
No. of utterances for story generation	0.333	11.088	0.998	8	.347
MLU for story generation	-0.438	19.263	1.908	8	.204

4. Discussion

The question was posed as to whether mothers' judgement on the representativeness of the language elicited would correlate with the volume of language elicited, as measured in terms of number of utterances and MLU. It was anticipated that there would be a significant correlation for freeplay and conversation but not for story generation, because mothers are presumably more familiar with the language their children produce during freeplay and conversation than during story generation. The results of our small-scale study suggest that the judgements mothers make regarding the representativeness of their sons' language elicited during freeplay, but not during conversation, correlate significantly

with both measures of volume of the language elicited. For story generation, the rank assigned by the mothers did not correlate with any of the two measures of volume of language elicited. The answer to the above-mentioned question was, therefore, "yes and no". The results of the present study support only in part the general finding of the studies by Klee et al. (2000), Ratner and Silverman (2000), and Thal et al. (1999, 2000), and offer an important qualification to the finding that parent-report is a valid measure of children's communication abilities, namely that the particular method of sampling used is a significant factor.

A possible explanation for why the judgement of the mothers in the current study on the language elicited by conversation did not correlate with the volume of the language elicited by conversation, is that the frequent but not exclusive occurrence of one-word answers and elliptical utterances given by the children during the conversation activity made it difficult for the mothers to judge whether the language elicited during conversation was representative of their sons' everyday language usage. An explanation as to why the mothers' judgements on story generation did not correlate with the measures of volume may be that the specific story generation task utilised in this study was too unfamiliar to the mothers (meaning that it did not resemble the type of story generation activity in which the mothers and their sons usually engaged) to enable them to judge the representativeness of the language elicited by it in terms of volume of language elicited or any other criterion. Furthermore, it is possible (but not likely) that the interaction of the mothers with their sons consists mostly of what typically takes place during freeplay, and not of question-asking and story prompts, and that the freeplay method appeared natural, or everyday-like, to the mothers, due to its familiarity to them.

We also found that freeplay elicited significantly more utterances than story generation, but not significantly more than conversation. The reason for the differences between our results and those of Evans and Craig (1992), who found that conversation elicited more language than did freeplay, may be that participants in the Evans and Craig study were language-impaired and older than our participating children. The results of Stalnaker and Creaghead (1982) showed that question asking (as in conversation) or story telling by the researcher offered children more to talk about, as well as avoiding a situation in which the

child becomes involved with toys to the extent that verbalisation does not take place. Our results do not confirm those of Stalnaker and Creaghead, because freeplay encouraged our participants to verbalise more. The different types of questions used in the two studies may explain this difference in results: Stalnaker and Creaghead used mainly single question words (e.g., "Who?" and "Why?"), whereas we used more specific questions (e.g., "What does the front of your house look like?")

There is no consensus on the preferred number of utterances of which a language sample should consist (cf. Crystal et al. 1989 for a summary of various scholars' opinions on adequate sample size), but, according to Crystal et al. (1989), most scholars these days use 100. In their experience, 30 minutes' interaction, consisting of 15 minutes' free play and 15 minutes' conversation, yields between 100 and 200 sentences. In our study, all three elicitation methods produced more than 100 utterances on average in 15 minutes. Therefore, although conversation and freeplay resulted in significantly more utterances than did story generation, all three methods delivered sufficient utterances for clinical use. The reason for story generation eliciting the least utterances could be that not all participating children related to all the story prompts presented by the researcher. Where the child did not relate to the prompt, i.e., where the child indicated that he had not had an experience similar to the one reported by the researcher, the researcher had told a story, which took up more of the 15 minutes sample time than what a single question (as in the case of conversation) would, without eliciting more than a one-word or two-word response from the child. In this case, another story prompt had to be presented, without the guarantee that the child would relate to this prompt, thus reducing the time for verbalisation by the child even further.

In our study, story generation elicited the longest utterances (as in the studies of Dollaghan, Campbell, and Tomlin 1990; Wagner, Nettelbladt, Sahlén, and Hilholm 2000), and freeplay the shortest. These results are contrary to those of Stalnaker and Creaghead (1982), who found that story retelling elicited the longest, toys the second longest, and questions the shortest utterances. Possible reasons for the difference in results include the fact that Stalnaker and Creaghead used story retelling instead of story generation and that

they instructed their participants to describe how they played with the toys, whereas no such overt request for verbalisation was made in our study.

5. Conclusion

In summary, we found that freeplay elicited more utterances than did story generation, but that the utterances elicited by story generation were longer than those elicited by conversation and freeplay. Thal et al. (2000:1088) stated that "parent report may be useful for such clinical tasks as evaluating the representativeness of a laboratory sample". We found that there was a significant correlation between representativeness and volume of the language sample only for samples elicited during freeplay and not during conversation or story generation.

Our results suggest two implications for clinical practice. According to our results, freeplay elicits more utterances than does story generation (although story generation elicits longer utterances), and there is a significant correlation between mothers' judgements on the representativeness of freeplay and the number and length of utterances elicited, whereas this is not the case for the language elicited during conversation. Recall Wren's (1985) suggestion that the length and the representativeness of the obtained language sample are two important criteria for the selection of a language elicitation method. The first implication for clinical practice is then that, if Wren's two criteria are seen as the most important ones (more important than, for example, variety of syntactic structures elicited), freeplay should be the language elicitation method of choice with 5-year-olds. The reason for this is that freeplay elicits as many utterances as does conversation and more than does story generation, but only for freeplay is there a significant correlation between the number of utterances elicited and the mothers' judgements on the representativeness of language elicited.

The second implication for clinical practice concerns story generation as a language elicitation method. Despite the lack of correlation between mothers' judgements of and volume of language elicited by story generation, this method renders valuable information for the speech-language therapist, due to the long MLUs and high proportion of complex utterances elicited (cf. Southwood and Russell 2004 for detailed results on the proportion

of complex utterances elicited by story generation). Recall Wren's (1985) statement that the term "representative" can be used to describe either language that resembles typical performance or language that resembles maximum performance. As stated before, in this study, we examined representativeness in terms of typical language production, for reasons already provided. However, had we chosen to examine representativeness in terms of maximum language production, story generation could have been one of the methods of choice, as it elicits the longest and most complex utterances.

When considering these implications for clinical practice, it should be borne in mind that this study was a small-scale study, with several methodological limitations, including the following: Because the appropriateness of elicitation methods may differ according to the client's age (Lee and Carter 1971), our results cannot be generalised to other age groups. Future research should replicate this study with larger sample sizes of children from different age groups including both genders of children and adults. This would allow for generalisability of the findings. Similar studies with children who do not exhibit normal language development should also be performed, because one cannot necessarily generalise caregivers' judgements regarding normal language to those regarding delayed or disordered language. Furthermore, caregivers' judgements on clinic-elicited language samples should be compared to their judgements on home-elicited samples. Also, in future studies, the use of a rating scale should be followed up by a semi-structured interview with the raters in order to explore the actual criteria used in allocating the ratings.

Laosa (1980, 1982) found that there is a high correlation between the amount of schooling individuals receive and how those individuals as parents interact with their children. Because children from nonmiddle-class homes may be used to a different form of interaction with adults, the study should be replicated with mothers and their children from other socioeconomic groups.

There is the possibility that mothers do not consider the volume of their child's verbalisations (i.e., the number and length of the child's utterances) when deciding on the representativeness of a language sample. That is, contrary to our assumption, mothers might be able to distinguish, without extensive training, between length and complexity or

might be able to use range of utterances or error types as criteria when judging the representativeness of language samples. Therefore, studies similar to ours should be performed, analyzing the language samples according to different criteria, amongst others those previously mentioned, namely structural complexity, number of grammatical errors, and pragmatic criteria (using trained and untrained parents).

Speech-language therapists continually strive to increase the involvement of parents in the assessment and remediation process, amongst others by encouraging them to share the specialist knowledge they have of their children. In order for this shared knowledge to be of use, however, the speech-language therapist needs to know that the information obtained from parents is accurate. This is especially the case when the use of formal assessment instruments is substituted by language sample analysis due to a lack of such instruments (or a lack of appropriate ones) for expressive language abilities of clients of certain language groups, such as Afrikaans. In such cases, where language sample analysis is used for diagnostic purposes, a parent judgement is needed to determine whether a language sample is representative of a child's language abilities, in order to increase the accuracy of the diagnosis. This study, despite its small sample size, provides grounds for optimism that mothers can play a valuable role during assessment by providing accurate information on the representativeness of language samples elicited during freeplay.

Notes

- * We would like to thank two anonymous SPIL-reviewers for their comments on a previous draft of this paper.
1. From as early as 1970, research results (cf., amongst others, Braine 1971; Brown and Hanlon 1970) have shown that, although parents correct errors of fact in their children's utterances, they do not correct their children's grammatical errors. A study reported on in Cho and O'Grady (1987:489) showed that mothers do revise their 2-year-olds' utterances (14% of grammatical utterances were revised and 26% of ungrammatical ones), but not the utterances of older children. In a more recent study, Levy (1999) presented evidence that parents (Hebrew-speaking, in this case) do correct grammatical errors in their children's utterances. It appears then that some researchers found no parental correction of ungrammatical utterances, some found correction of ungrammatical utterances, but only when produced by 2-year-olds, and others found evidence of wider spread parental correction of ungrammatical utterances. As mentioned before, there seems to be no consensus on whether or not parents can systematically identify the ungrammaticalities in their children's utterances.
 2. The researcher's and child's speaker time during each 15 minute sample was not timed, but from visual examination of the transcripts, it is clear that these times differed across samples. It is acknowledged that the researcher's vs the child's speaking time during the three language samples obtained from the child may have influenced the mothers' judgement of the samples. However, this is one variable that could not be controlled for, given that, where a child was not interacting in verbal communication, the researcher had to prompt the child to verbalise and still had to engage in what would be regarded as natural communicative behaviour, i.e., the researcher had to carry on talking and creating opportunities for the child to participate in the verbal interaction.

References

- American Speech and Hearing Association. 1975. Guidelines for identification audiometry. *ASHA* 17: 94-99.
- American Speech and Hearing Association. 1979. Committee on audiometric evaluation: Guidelines for acoustic immittance screening of middle ear function. *ASHA* 21: 283-288.
- Aram, D.M., and Nation, J.E. 1980. Preschool language disorders and subsequent language and academic difficulties. *Journal of Communication Disorders* 13: 159-170.
- Atkins, C.P., and Cartwright, L.R. 1982. An investigation of the effectiveness of three language elicitation procedures on Head Start children. *Language, Speech, and Hearing Services in Schools* 13: 33-36.
- Braine, M.D.S. 1971. On two types of models of the internalization of grammars. In D. I. Slobin (ed.) *The ontogenesis of grammar: A theoretical symposium*. New York: Academic Press. pp. 153-186.
- Brown, R. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Brown, R., and Hanlon, C. 1970. Derivational complexity and order of acquisition in child speech. In J.R. Hayes (ed.) *Cognition and the development of language*. New York: John Wiley and Sons, Inc. pp. 11-53.
- Cho, S.W., and O'Grady, W. 1987. Language acquisition: The emergence of grammar. In W. O'Grady, M. Dobrovolsky, and F. Katamba (eds.) *Contemporary linguistics. An introduction*. London and New York: Longman. pp. 464-502.
- Craig, H.K., and Washington, J.A. 2000. An assessment battery for identifying language impairments in African American children. *Journal of Speech, Language, and Hearing Research* 43: 366-379.
- Cresswell, J.W. 1998. *Qualitative inquiry and research design*. London: SAGE Publications, Inc.
- Crystal, D. 1979. *Working with LARSP*. London: Edward Arnold.
- Crystal, D., Fletcher, P., and Garman, M. 1989. *The grammatical analysis of language disability* (2nd ed.). London: Edward Arnold.
- De Vaus, D. 2001. *Research design in social research*. London: Sage Publications.

- Dollaghan, C.A., Campbell, T.F., and Tomlin, R. 1990. Video narration as a language sampling context. *Journal of Speech and Hearing Disorders* 55: 582-590.
- Dunn, M., Flax, J., Sliwinski, M., and Aram, D. 1996. The use of spontaneous language measures as criteria for identifying children with specific language impairment: An attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research* 39: 643-654.
- Evans, J.L., and Craig, H.K. 1992. Language sampling collection and analysis: Interview compared to freeplay assessment contexts. *Journal of Speech and Hearing Research* 35: 343-353.
- Evans, L.E., and Miller, J. 1999. Language sample analysis in the 21st century. *Seminars in Speech and Language* 20: 101-116.
- Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D., Hartung, J., and Reilly, J. 1993. *Technical manual for the MacArthur communicative developmental inventory*. San Diego, CA: Developmental Psychological Laboratory, San Diego State University.
- Garman, M. 1989. The role of linguistics in speech therapy: Assessment and interpretation. In P. Grumwell, and A. James (eds.) *The functional evaluation of language disorder*. London: Croom Helm. pp. 29-57.
- Gardner, R. 1979. *Expressive one-word picture vocabulary test*. Novato, CA: Academic Therapy Services.
- Grimes, B.F. 1996. *Ethnologue* (13th ed.). Dallas, TX: Summer Institute of Linguistics, Inc.
- Hadley, P., and Rice, M. 1993. Parental judgements of preschoolers' speech and language development: A resource for assessment and IEP planning. *Seminars in Speech and Language* 14: 278-288.
- Hakim, C. 1987. *Research design: Strategies and choices in the design of social research*. London: Allen and Unwin.
- Hawkins, P., and Spencer, H. 1985. Imitative versus spontaneous language assessment: A comparison of CELI and LARSP. *British Journal of Disorders of Communication* 20: 191-200.
- Jackson-Maldonado, D., Bates, E., and Thal, D. 1992. *Fundación MacArthur: Inventario del desarrollo de habilidades comunicativas*. San Diego, CA: San Diego State University.

- Kemp, K., and Klee, T. 1997. Clinical language sampling practices: Results of a survey of speech-language pathologists in the United States. *Child Language Teaching and Therapy* 13: 161-176.
- Klee, T., Carson, D.K., Gavin, W.J., Hall, L., Kent, A., and Reece, S. 1998. Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research* 41: 627-641.
- Klee, T., Pearce, K., and Carson, D.K. 2000. Improving the positive predictive value of screening for developmental language disorder. *Journal of Speech, Language, and Hearing Research* 43: 821-833.
- Labov, W. 1972. *Language in the inner city*. Philadelphia, PA: University of Pennsylvania Press.
- Laosa, L.M. 1980. Maternal teaching strategies in Chicano and Anglo-American families: The influence of culture and education on maternal behaviour. *Child Development* 51: 759-765.
- Laosa, L.M. 1982. School, occupation, culture, and family: The impact of parental schooling on the parent-child relationship. *Journal of Educational Psychology* 76(6): 791-827.
- Lassman, F.M., Fisch, R.O., Vetter, D.K., and La Benz, E.S. 1980. *Early correlates of speech, language, and hearing*. Littleton, MA: PSG Publishing.
- Lee, L.L., and Carter, S.M. 1971. Developmental sentence scoring: A clinical procedure for estimating syntactic development in children's spontaneous speech. *Journal of Speech and Hearing Disorders* 36: 315-340.
- Leedy, P.D., and Ormrod, J.E. 2001. *Practical research: Planning and design*. (7th ed.). New Jersey: Prentice-Hall, Inc.
- Levy, Y. 1999. Early metalinguistic competence: Speech monitoring and repair behavior. *Developmental Psychology* 35(3): 822-834.
- Longhurst, T.M., and Grubb, S. 1974. A comparison of language samples collected in four situations. *Language, Speech, and Hearing Services in Schools* 5: 71-78.
- Lyon, G.R. 1996. Learning disabilities. *The Future of Children* 61(1): 54-76.
- McTear, M., and Conti-Ramsden, G. 1992. *Pragmatic disability in children*. London: Whurr Publishers.

- Miller, F., and Chapman, R.S. 1981. The relationship between age and mean length of utterance. *Journal of Speech and Hearing Research* 24: 154-161.
- Minskoff, E.H., Wiseman, D.E., and Minskoff, J.G. 1972. *The M.W.M. program for developing language abilities: Expression manual*. New Jersey: Educational Performance Associates.
- Office of Special Education Programs. 1998. Gender as a factor in special education eligibility, services, and results. *Twentieth Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act*. Washington, DC: Office of Special Education Programs and Rehabilitative Services, U.S. Department of Education. pp. II.25-II.43.
- Peña, E., Quinn, R., and Iglesias, A. 1992. The application of dynamic methods to language assessment: A nonbiased procedure. *The Journal of Special Education* 26: 269-280.
- Peterson, C., and McCabe, A. 1983. *Developmental psycholinguistics: Three ways of looking at child's narrative*. New York: Plenum Press.
- Ratner, N.B., and Silverman, S. 2000. Parental perceptions of children's communicative development at stuttering onset. *Journal of Speech, Language, and Hearing Research* 43: 1252-1263.
- Southwood, F., and 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: 366-376.
- Stalnaker, L.D., and Creaghead, N.A. 1982. An examination of language samples obtained under three experimental conditions. *Language, Speech, and Hearing Services in Schools* 13: 121-128.
- Statistics South Africa. 2003. *Census in brief*. (No. 03-02-03-2001). Pretoria, South Africa: Author.
- Templin, M.C. 1957. Certain language skills in children: Their development and interrelationships. *Institute of Child Welfare Monograph Series* (No. 26). Minneapolis, MN: University of Minneapolis Press.
- Thal, D., Jackson-Maldonado, D., and Acosta, D. 2000. Validity of parent-report measure of vocabulary and grammar for Spanish-speaking toddlers. *Journal of Speech, Language, and Hearing Research* 43: 1087-1100.

- Thal, D., O'Hanlan, L., Clemmons, M., and Fralin, L. 1999. Validity of parent-report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech, Language, and Hearing Research* 42: 482-496.
- Thorndike, R., Hagen, E., and Sattler, J. 1986. *The Stanford Binet intelligence scale* (4th ed.) Chicago: Riverside Publishing Company.
- Tomblin, J., Records, N., Buckwalter, P., Zhang, X., Smith, E., and O'Brien, M. 1997. Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research* 40: 1245-1260.
- Toronto, A.S., and Merrill, S. 1983. Developing local normed assessment instruments. In D.R. Omark, and J.G. Erickson (eds.) *The bilingual exceptional child*. San Diego, CA: College-Hill Press. pp. 105-121.
- Vorster, J. 1980. *Toets vir mondeling taalproduksie*. Pretoria, South Africa: South African Institute for Psychological and Psychometric Research.
- Wagner, C.R., Nettelbladt, U., Sahlén, B., and Hilholm, C. 2000. Conversation versus narration in pre-school children with language impairment. *International Journal of Language and Communication Disorders* 35: 83-93.
- Woodward, M., and Francis, L.M.A. 1988. *Statistics for health management and research*. London: Edward Arnold.
- Wren, C.T. 1985. Collecting language samples from children with syntax problems. *Language, Speech, and Hearing Services in Schools* 16: 83-102.
- Zimmerman, I.L., Steiner, V.G., and Pond, R. 1979. *Preschool language scale*. Columbus, OH: Merrill.

APPENDIX

Excerpts of the Verbal Interaction Between the Researcher and Participants during the Three Language Elicitation Methods

Key

- P Participant
R First researcher
/ End of utterance
-- Silence longer than 2 s
(XXX) Unintelligible

Conversation

Participant 4. Rank assigned by mother: 1.

- R: van watter speelgoed hou jy die meeste/
P: van karretjies/
R: hoe speel jy met dit/
P: mens laat die goed ry/
R: mens laat die goed ry/ met wie speel jy graag by die skool/
P: um/ met 'n maatjie van my/
R: en by die huis/ met wie speel jy dan graag/
P: al die maatjies wat in my straat woon/
R: o/ vertel my/ hoe speel 'n mens wegkruipertjie/
P: mens kruip weg/ dan moet dan moet die een tel/ dan dan moet die ander een vir
mens soek/

Participant 1. Rank assigned by mother: 3.

- R: van watter speelgoed hou jy die meeste/
P: van van die blokkies/ nee/ van Lego/
R: hoe speel mens met dit/
P: mens bou/
R: o/ met wie speel jy graag by die skool/
P: met Giovanni/
R: en by die huis/ met wie speel jy dan graag/

- P: met Nina/
R: vertel my/ hoe speel 'n mens wegkruipertjie/
P: mens kruip weg/ en 'n ou tel/ en dis al/

Freeplay

Participant 10. Rank assigned by mother: 1.

- P: hoe lyk die meubels/
R: so/ hierso/ ek dink ons moet die deur hier vaskry/ -- ai toggie/
P: ons moet baie ander goed agter hom sit/ sit baie ander goed agter hom/
R: daar's hy/
P: sit hom hierso/ (XXX)/ hey/ hey/ is hier ooit mannetjies/
R: uh/ hierso's so enetjie/ so ou kleintjie/ en dan's hier sulkes wat 'n mens aan jou hand sit/
P: 'n pop/ pop/
R: ja/ handpoppe/ [researcher puts puppet on her hand]
P: hey/ [addressing puppet]

Participant 4. Rank assigned by mother: 3.

- R: OK/ daai deur gaan seker nou heelyd toeval/ kom ons sit [researcher puts building block in front of door]/ oeps/ dit werk nie/ kom ons probeer hierdie een [researcher adds another building block]/ ook nie/ -- dan pak ons hulle só/ né/ -- OK/ dan soek ons uit watter meubels moet in watter kamer kom/ ek dink/ jy kan nou nie eintlik lekker daar bykom nie [referring to furniture]/ hierso's vir jou -- [researcher wants to hand participant a piece of furniture]/ wat dink jy/ wat moet waar kom/ -- ek moet/ ek dink hierdie kamer moet die sitkamer wees/ né/ want dit het 'n lekker mat in/ dan sit ons hierdie daar [referring to couch]/ so ja/ -- goed/ dit lyk mooi/ dink jy ons moet hierdie die koffietafeltjie maak/ OK/
P: 'n stoof/

Story Generation

Participant 1. Rank assigned by mother: 1.

- R: [tells story about her cat catching a fish out of a fish pond] het jy enige troeteldiere/
P: huh-uh/

- R: nie/
P: huh-uh/
R: ken jy iemand met 'n troeteldier/
P: ja/
R: o/ het die troeteldier al ooit so iets stouts aangevang/
P: huh-uh/
R: nie/ [tells story about taking her dog for a walk and a big dog biting her dog's leg]
en weet jy wat sê die veearts toe/
P: huh-uh/
R: hy het gesê die nare hond het Sandy se been afgebyt/ toe moes ons haar been in
gips sit/ sodat dit weer kan aangroei/ het so iets al met SP se troeteldier gebeur/
P: ek weet nie/

Participant 10. Rank assigned by mother: 3.

- R: [tells story about her cat catching a fish out of a fish pond] het jy troeteldiere/ huh/
P: kat en 'n hond/
R: o/ het hulle al so iets stouts aangevang/
P: [shakes head]
R: o/ -- [tells story about taking her dog for a walk and a big dog biting her dog's leg]
weet jy wat sê die veearts toe/
P: wat/
R: hy sê daai nare groot hond het Sandy se been gebreek/ toe moes Sandy se been in
gips kom/ sodat dit weer kan aangroei/ het so iets al met jou gebeur/
P: huh-uh/
R: en met julle troeteldiere/
P: ja/ niks het daar het al 'n kat met ons ander kat baklei met ons een kat baklei/
N: met julle een kat baklei/ en toe/
P: niks het gebeur met my kat nie/