JSLHR

Research Article

Conversational Language in 3-Year-Old Children Born Very Preterm and at Term

Katherine Sanchez,[a,b] Alicia J. Spittle,[a,c,d] Jessica O. Boyce,[a,b] Linda Leembruggen,[b] Anastasia Mantelos,[b] Stephanie Mills,[b] Naomi Mitchell,[b] Emily Neil,[b] Miya St John,[a] Jasmin Treloar,[b] and Angela T. Morgan[a,b,e]

Purpose: Language difficulties are prevalent among children born preterm. Existing studies have largely used standardized language tests, providing limited scope for detailed descriptive examination of preterm language. This study aimed to examine differences in conversational language between children born < 30 weeks and at term as well as correlations between language sample analysis (LSA) and a standardized language tool.
Method: Two hundred four 3-year-olds (103 born < 30 weeks, 101 born at term) recruited at birth provided a 10-min language sample and completed the Preschool Language Scales–Fifth Edition (I. Zimmerman, Steiner, & Pond, 2011). LSA was conducted using the Systematic Analysis of Language Transcripts and Index of Productive Syntax. Group differences were analyzed using linear regression, and Pearson correlation coefficient (coef) was used to determine correlations between measures.
Results: Children born < 30 weeks scored lower than term-born peers on multiple metrics when controlled for confounding factors (sex, high social risk, multilingualism, and diagnosed neurodevelopmental disorders), including mean length of utterance in morphemes (coef = –0.28, 95% confidence interval [CI] [–0.56, 0.01]) and words (coef = –0.29, 95% CI [–0.53, –0.05]), number of different word roots (coef = –10.04, 95% CI [–17.93, –2.14]), and Index of Productive Syntax sentence structures (coef = –1.81, 95% CI [–3.10, –0.52]). Other variables (e.g., number of utterances, number of nouns and adjectives) were not significantly different between groups. LSA and the Preschool Language Scales–Fifth Edition were at most moderately correlated (≤ .45).
Conclusions: Three-year-old children born preterm demonstrated poorer conversational language than children born at term, with some specific areas of deficit emerging. Furthermore, formal assessment and LSA appear to provide relatively distinct and yet complementary data to guide diagnostic and intervention decisions.
Supplemental Material: https://doi.org/10.23641/asha.11368073

Children born preterm (prior to 37 weeks of gestational age) are vulnerable to a range of developmental problems, including adverse language outcomes (Barre, Morgan, Doyle, & Anderson, 2011; E. Zimmerman, 2018). Language difficulties can have deleterious long-term implications, being associated with poor literacy and educational attainment, restricted employment opportunities, social–emotional challenges, and reduced quality of life (Botting, Toseeb, Pickles, Durkin, & Conti-Ramsden, 2016; Conti-Ramsden, Durkin, Toseeb, Botting, & Pickles, 2018). Poor self-esteem has also been noted among children born preterm with language difficulties (Islam et al., 2018; E. Zimmerman, 2018). Given these significant potential impacts, there is a growing need to better understand and remediate language in children born preterm, especially given the increasing number of younger and more medically complex infants surviving in this group (Saigal & Doyle, 2008).

Previous research exploring language outcomes in preterm children has relied heavily on parent report and standardized tests, particularly in preschool-age children (Barre et al., 2011). Standardized scores provided by these tools are beneficial for understanding a child's level of ability

[a]Clinical Sciences, Murdoch Children's Research Institute, Parkville, Victoria, Australia
[b]Department of Speech Pathology and Audiology, The University of Melbourne, Parkville, Victoria, Australia
[c]Department of Physiotherapy, The University of Melbourne, Parkville, Victoria, Australia
[d]The Royal Women's Hospital, Parkville, Victoria, Australia
[e]Department of Speech Pathology, the Royal Children's Hospital, Parkville, Victoria, Australia

Correspondence to Katherine Sanchez: katherine.sanchez@mcri.edu.au

Editor-in-Chief: Sean M. Redmond
Editor: Stephen M. Camarata

Received January 24, 2019
Revision received March 6, 2019
Accepted September 19, 2019
https://doi.org/10.1044/2019_JSLHR-19-00153

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

relative to their chronological peer group and the broader population and are typically required in most institutions or jurisdictions to provide evidence for provision of support. Yet, despite these important contributions made by standardized tools, they are limited in terms of ecological validity and are arguably more impacted by cultural bias (Hewitt, Hammer, Yont, & Tomblin, 2005). The summary scores generated by standardized language tests have also been criticized for their potential to mask subtle language differences (Barre et al., 2011), an effect that has been noted in studies of the preterm population (Guarini et al., 2010). As such, formal assessments are generally agreed to be inadequate as the sole measure of language ability and as sensitive tools for setting intervention goals.

An alternative approach to standardized language assessment is language sample analysis (LSA). This tool has been even been referred to by some as the "gold standard" for describing children's language (Heilmann, Nockerts, & Miller, 2010). LSA is argued to provide a more natural, representative, and less biased sample of linguistic output, generating a rich data source for identifying strengths and weaknesses that can then be translated into therapeutic goals (Heilmann et al., 2010). Despite these reported benefits, the number of studies that have applied LSA to a preterm population is minute, relative to those that use formal assessment (see Table 1). Furthermore, studies that have used LSA have been limited in only reporting a restricted range of outcomes (Craig, Evans, Meisels, & Plunkett, 1991; Félix, Santos, & Benítez-Burraco, 2017; Grunau, Kearney, & Whitfield, 1990; Le Normand, Vaivre-Douret, & Delfosse, 1995; Rice, Spitz, & O'Brien, 1999), such as mean length of utterance (MLU) only. Hence, to date, these data have helped to complement what has been measured by accompanying standardized language test outcomes but have not provided a comprehensive depth of language assessment to support differential diagnosis and guide decisions about specific and targeted intervention. In addition, the few studies that provide a more detailed range of LSA outcomes in those born preterm are focused on school-age children (Crosbie, Holm, Wandschneider, & Hemsley, 2011; Mahurin-Smith, DeThorne, Logan, Channell, & Petrill, 2014) or lack a term-born comparison group (Feldman, Janosky, Scher, & Wareham, 1994), limiting their generalizability to toddlers and preschool-age children.

A more detailed descriptive analysis of preterm language would assist clinicians and researchers to build a preterm language profile, which could guide surveillance, assessment, and intervention decisions. Furthermore, as an association between language outcomes and neuroimaging metrics in the preterm population emerges (Mürner-Lavanchy et al., 2018; Northam et al., 2012), a detailed picture of preterm language could support our understanding of the neurobiological underpinnings of language impairment in this group.

Naturally, given the lack of studies investigating language using LSA approaches in the preterm group, there are even fewer studies exploring commonalities or differences between language assessment methods and reported outcomes in the preterm population. In nonpreterm populations, studies have observed congruence between LSA metrics such as MLU and the outcomes of standardized tools (Owens & Pavelko, 2017). However, those that have examined this in children born preterm have found a less consistent picture, with children born preterm typically performing worse on standardized assessment than LSA (Crosbie et al., 2011; Imgrund, Loeb, & Barlow, 2019; Mahurin-Smith et al., 2014). It is important to establish the congruence between measures to better understand the potential impact of using different assessment methods and to gain a true, uninflated picture of language deficits in this population. This knowledge can in turn provide clinicians with guidance as to the value of different assessment methods in preschoolers born preterm in a critical period of language development. In this study, we aimed to (a) examine differences in conversational language between 3-year-old children born at < 30 weeks and at term and (b) determine correlations between language outcomes obtained from conversational language analysis and a standardized language assessment tool in 3-year-old children born at < 30 weeks.

Method

Participants

Infants born at term (≥ 37 weeks) and at < 30 weeks of gestational age were recruited within 2 weeks of birth and prior to hospital discharge from the Royal Women's Hospital and Frances Perry House as part of a larger prospective longitudinal cohort study of preterm neurobehavior and neurodevelopment (Spittle et al., 2014). The cut-point of 30 weeks was selected as these infants are thought to be at highest risk of adverse neurodevelopmental outcomes (Spittle et al., 2014). Infants were excluded if they had congenital abnormalities that affect neurodevelopment or if their parents did not speak English (Spittle et al., 2014). The larger study was approved by the Royal Women's Hospital Human Research Ethics Committee (HREC 10/07), and the 3-year follow-up was approved by the Royal Children's Hospital Research Ethics Committee (34147A). Caregivers provided written consent to participate. Demographic and medical information were collected at recruitment, including sex, gestational age at birth, and social risk (Spittle et al., 2014). The Social Risk Index was used to describe social risk, scoring six aspects of social status: family structure, education of primary caregiver, occupation of primary income earner, employment status of primary income earner, language spoken at home, and maternal age at birth (Spittle et al., 2014).

Three-Year Data Collection

At 3 years of age (corrected age for participants born preterm), families were invited to attend an assessment appointment at the Murdoch Children's Research Institute. As part of a battery of speech and language assessments, participants completed the Preschool Language Scales–Fifth Edition (PLS-5) Australia and New Zealand adapted

Table 1. Studies of connected conversational or narrative language in children born preterm.

		Population				
Study	**Year**	***n***	**Age**	**Location**	**Assessment**	**Outcome**
Grunau et al.	1990	23 children with birth weight < 1,000 g, 23 children born at term	3 years	Vancouver, Canada	30 min of spontaneous language during unstructured play	Sentence complexity lower in preterm group, but MLUm equivalent between groups
Craig et al.	1991	30 children born between 27 and 37 weeks	3 years	Michigan, USA	20 min of spontaneous language during parent–child unstructured and structured play	13% of children demonstrated language delays
Feldman et al.	1994	18 boys born ≤ 36 weeks	3 years	Pennsylvania, USA	25 min of spontaneous language during parent–child play	MLU and IPSyn within normal range
Le Normand et al.	1995	37 children born ≤ 36 weeks, 74 children born at term	2 years and 3;6 years;months	Paris, France	20 min of spontaneous language during play	MLUm lower in preterm group, but vocabulary equivalent between groups
Rice et al.	1999	69 children who had been neonatal intensive care inpatients	4 years	Kansas, USA	100 spontaneous utterances	14% of children scored below MLUm reference values
Crosbie et al.	2011	15 children born < 33 weeks, 15 children born at term	9;8–10;11 years;months	Queensland, Australia	Expression, Reception, and Recall of Narrative Instrument	No significant differences between groups on any standard score
Mahurin-Smith et al.	2014	57 children born ≤ 32 weeks, 57 children born at term	7, 8, and 10 years	Ohio, USA	15 min of spontaneous language during play with modelling clay	Preterm group produced most target structures less frequently than term group, but no significant differences
Félix et al.	2017	19 children born between 28 and 37 weeks	4–5 years	Lisbon, Portugal	30 min of spontaneous language during play	Preterm MLUw lower than reference values; no clear increase in MLUw with age
Imgrund et al.	2019	29 children born < 34 weeks, 29 children born at term	4–5 years	Kansas, USA	100 complete and intelligible utterances obtained over 90 min of unstructured play	Scores on all conversational semantic and grammatical metrics lower in preterm group

Note. MLUm = mean length of utterance in morphemes; MLUw = mean length of utterance in words; IPSyn = Index of Productive Syntax.

(I. Zimmerman, Steiner, & Pond, 2011) and approximately 20 min of unstructured play with one of five trained pediatric speech-language pathologists. The primary outcomes from the PLS-5 assessment for the entire cohort are reported in a separate paper (Sanchez et al., 2019). Parents were also asked whether their child had received a diagnosis of a neurodevelopmental disorder, such as autism spectrum disorder or cerebral palsy, and whether their child understood or spoke more than one language.

The play protocol, drawn from previous studies of conversational language ability, involved presenting the child with a standard set of toys (foam blocks and a farm set) and two popular children's books (DeThorne, Deater-Deckard, Mahurin-Smith, Coletto, & Petrill, 2011; Guo & Eisenberg, 2015; Rice, Redmond, & Hoffman, 2006). Clinicians were instructed to follow the child's lead in play and conversation; engage in parallel talk and "glossing" (repeating back the child's utterances to facilitate transcription); and minimize the use of corrections, yes/no, and *wh*-questions (DeThorne et al., 2011; Guo & Eisenberg, 2015; Rice et al., 2006). Approximately 20 min of play were recorded to allow the participants 10 min to "warm up." Conversational language was analyzed in the second 10-min segment—a sample length that has been found to be effective in assessing the language of 3-year-old children (Guo & Eisenberg, 2015). For samples of 20 min or more, the second 10 min of the sample was transcribed. For samples of less than 20 min, the final 10 min was transcribed. Samples with fewer than 50 utterances within a 10-min sample were omitted from analysis to reduce the risk of measurement error (Guo & Eisenberg, 2015).

Language sample transcription and analysis was conducted according to the protocols and conventions of the Systematic Analysis of Language Transcripts 16 (SALT 16) —Australia and New Zealand (Miller, Gillon, & Westerveld, 2015). All transcribers (L. L., A. M., S. M., N. M., E. N., J. T.) were final-year graduate speech-language pathology students who completed 15 hr of online training in SALT transcription and analysis. One transcriber (J. T.) completed 8 additional hours of personalized training with an expert-level transcriber from the SALT company to achieve > 95% transcription and coding accuracy for the purpose of establishing strong interrater reliability. Index of Productive Syntax (IPSyn) scores—which examine the emergence of 56 different morphosyntactic structures from noun phrase, verb phrase, questions, negations, and sentence structure categories—were calculated using an automated online tool (Hassanali, Liu, Iglesias, Solorio, & Dollaghan, 2014; Scarborough, 1990).

Variables of interest for this study included the following:

- measures of volubility and linguistic fluency: number of utterances, total number of completed words, words per minute, percent of intelligible and complete utterances, and maze words as a percentage of total words;
- measures of morphosyntactic complexity: mean length of utterance in words and morphemes (MLUw and MLUm) and IPSyn noun phrase score, IPSyn verb phrase score, IPSyn question/negation score, IPSyn sentence structure score, and IPSyn total score (a composite of the other IPSyn scores); and
- measures of lexical variety: total number of words (TNW); number of different word roots; and number of nouns, verbs, adjectives, and prepositions.

All variables were defined according to the instructions provided by the SALT and IPSyn developers (Miller et al., 2015; Scarborough, 1990). An online randomizer was used to select 10 preterm and 10 term samples for retranscription to establish interrater reliability for LSA. The intraclass correlation coefficient was calculated, with values of > .75 considered excellent, between .40 and .75 considered fair to good, and < 0.40 considered poor (Portney & Watkins, 2009). Results were tabulated in Supplemental Material S1.

Statistical Analysis

Data were entered into REDCap and analyzed using Stata 14 (Harris et al., 2009; StataCorp, 2015). Differences between children born at term and preterm were analyzed twice: once with no controlling for potential confounders using univariable analysis and once controlling for known predictors of language difference (sex, high social risk, multilingualism, and diagnosed neurodevelopmental disorders; Reilly et al., 2010) using multivariable analysis. Both univariable and multivariable analyses were conducted as there is a known relationship between preterm birth and other predictors of language impairment (e.g., sex, high social risk), and it was considered that including both types of analyses might assist in understanding preterm birth both as an independent contributor to outcomes and as part of a cluster of related risk factors.

Only two children in the cohort had any diagnosed hearing impairment; therefore, hearing impairment was not included as a predictor in analyses. Linear regressions were fitted using generalized estimating equations, accounting for multiple births, to report group differences. Pearson correlation coefficient was used to examine the relationship between PLS-5 measures and variables reported from LSA. The relationship between classifications was determined by identifying participants who were 1.5 *SD*s below the mean PLS-5 Total Language Score, MLUm, and TNW of the term-born children included in our study. MLUm and TNW were selected to represent morphosyntactic and lexical ability because they have been used in previous studies as primary outcomes of LSA (Owens & Pavelko, 2017).

Results

Two hundred four children (101 born at term, 103 born < 30 weeks) attended a clinical assessment appointment for the 3-year follow-up. In six cases, language samples were unusable or not collected due to technical failures or running out of time. Of the 198 language samples collected (97 from children born preterm, 101 from children born at term), 31 were excluded as they contained < 50 utterances. This left 167 samples in the final analysis: 87 from children born at < 30 weeks and 80 from children born at term. The characteristics of the analyzed sample can be seen in Table 2. For a comparison of participants who produced fewer than 50 utterances versus those who produced 50 or more utterances, see Supplemental Material S1. The excluded group had a relatively higher proportion of children from a multilingual background, or with a diagnosed neurodevelopmental disorder, and a relatively lower proportion of multiple births and children who completed a full PLS-5. Interrater reliability values (shown in Supplemental Material S2) ranged from fair (.53) to excellent (.98), with a mean of .79 (excellent).

Table 3 shows the mean values for a range of linguistic variables and compares the term and preterm groups. The results of both uncontrolled (univariable) and controlled (multivariable; for sex, high social risk, multilingualism, and diagnosed neurodevelopmental disorders) analyses are shown. Based on these analyses, it is clear that the term and preterm groups differed on MLU in both words (uncontrolled coefficient [coef.] = –0.41, 95% confidence interval [CI] [–0.65, –0.16], p = .001; controlled coef. = –0.29, 95% CI [–0.53, –0.05], p = .02) and morphemes (uncontrolled coef. = –0.39, 95% CI [–0.66, –0.12], p = .01; controlled coef. = –0.28, 95% CI [–0.56, 0.01], p = .04), the IPSyn Sentence Structures score (uncontrolled coef. = –1.97, 95% CI [–3.26, –0.69], p = .003; controlled coef. = –1.81, 95% CI [–3.10, –0.52], p = .01), and the number of different word roots (uncontrolled coef. = –11.29, 95% CI [–18.94, –3.63], p = .004; controlled coef. = –10.04, 95% CI [–17.93, –2.14], p = .01). When analyses did not control for possible confounders, IPSyn total score (uncontrolled coef. = –4.52, 95% CI [–8.55, –0.51], p = .03), IPSyn Noun Phrase subscale (uncontrolled coef. = –0.99, 95% CI [–1.80, –0.17], p = .02), number of verbs (uncontrolled coef. = –6.25, 95% CI [–12.41, –0.10], p = .046) and number of prepositions (uncontrolled coef. = –2.39, 95% CI [–4.69, –0.09], p = .04) were also significantly different between groups. Other variables—number of utterances (uncontrolled coef. = 3.12, 95% CI [–5.16, 11.30], p = .46; controlled coef. = 3.65, 95% CI [–4.85, 12.15], p = .40), total number of completed words (uncontrolled coef. = –33.52, 95% CI [–68.46, 1.42], p = .06; controlled coef. = –18.58, 95% CI [–55.02, 17.86], p = .32), words per minute (uncontrolled coef. = –3.35, 95% CI [–6.86, 0.17], p = .06; controlled coef. = –1.90, 95% CI [–5.57, 1.76], p = .31), maze words as a percentage of total words (uncontrolled coef. = –0.19, 95% CI [–1.33, 0.96], p = .75; controlled coef. = –0.04, 95% CI [–1.16, 1.09], p =.95), IPSyn verb phrase (uncontrolled coef. = –1.07, 95% CI [–2.73, 0.58], p = .20; controlled coef. = –0.39, 95% CI [–2.03, 1.25], p = .64) and question/negation scores (uncontrolled coef. = 0.01, 95% CI [–1.62, 1.64], p = .99; controlled coef. = 0.59, 95% CI [–1.31, 2.49], p = .54), number of nouns (uncontrolled coef. = –3.53, 95% CI [–12.08, 5.02], p = .42; controlled coef. = –3.09, 95% CI [–12.14, 5.96], p = .50), and number of adjectives (uncontrolled coef. = –1.26, 95% CI [–3.12, 0.60], p = .18; controlled coef. = –1.02, 95% CI [–2.95, 0.92], p = .30)—did not differ significantly between groups.

Figure 1 shows how many of the 155 children (78 born < 30 weeks, 77 born at term) who had both PLS-5 Total Language and LSA scores were classified as having a language problem (i.e., performed at least 1.5 *SD*s below the term mean) using the PLS-5 Total Language Score, MLUm, and TNW, respectively. One hundred thirty-five of these 155 participants (87%) scored within or above 1.5 *SD*s of the term mean on all three measures, whereas 20 (12 born at < 30 weeks, eight born at term) scored below 1.5 *SD*s of the term mean on at least one measure. The PLS-5 Total Language Score identified 14 children(11 preterm, three term), MLUm identified 13 children (seven preterm, six term), and TNW identified five children (three preterm, two preterm). Of the 12 children who did not complete the PLS-5, six (five born < 30 weeks, one born at term) scored

Table 2. Participant characteristics.

Variable	Born < 30 weeks *n* = 87	Born ≥ 37 weeks *n* = 80
Gestational age at birth (weeks), *M* (*SD*)	27.86 (1.41)	40.06 (1.14)
Male sex, *n* (%)	45 (52)	42 (53)
Multiple birth, *n* (%)	38 (44)	2 (3)
High social risk, *n* (%)	32 (40)	18 (25)
Multilingual, *n* (%)	9 (10)	10 (13)
Diagnosis of neurodevelopmental disorder, *n* (%)	4 (5)	1 (1)
Receiving speech/language therapy at the time of assessment, *n* (%)	18 (21)	1 (1)
Corrected age (years) at assessment, *M* (*SD*)	3.20 (0.19)	3.20 (0.21)
Completed full PLS-5, *n* (%)	78 (90)	77 (96)
PLS-5 Auditory Comprehension scale, *M* (*SD*)	97.26 (14.48)	101.38 (11.03)
PLS-5 Expressive Communication scale, *M* (*SD*)	98.76 (10.07)	104.71 (11.24)
PLS-5 total scale, *M* (*SD*)	98.14 (11.85)	103.40 (11.26)

Note. PLS-5 = Preschool Language Scales–Fifth Edition.

Table 3. Language sample analysis outcomes.

Variable	Preterm *M* (*SD*)	Term born *M* (*SD*)	Group differences (uncontrolled) Magnitude coef. (CI)	*p*	Group differences (controlled) Magnitude coef. (CI)	*p*
Volubility and linguistic fluency						
Number of utterances	101.26	98.05	3.12	.46	3.65	.40
	(28.15)	(26.23)	[–5.16, 11.40]		[–4.85,12.15]	
Total number of completed words	277.67	311.26	–33.52	.06	–18.58	.32
	(106.18)	(123.29)	[–68.46,1.42]		[–55.02, 17.86]	
Words per minute	27.83	31.17	–3.35	.06	–1.90	.31
	(10.77)	(12.33)	[–6.86,0.17]		[–5.57, 1.76]	
Percent of intelligible and complete utterances	82.57	85.24	–2.55	.13	–3.94	.02
	(10.91)	(10.52)	[–5.83, 0.72]		[–7.28, –0.61]	
Maze words as a percentage of total words	4.34	4.56	–0.19	.75	–0.04	.95
	(3.35)	(4.27)	[–1.33, 0.96]		[–1.16, 1.09]	
Morphosyntactic complexity						
Mean length of utterance in words	2.81	3.20	–0.41	.001	–0.29	.02
	(0.78)	(0.80)	[–0.65, –0.16]		[–0.53, –0.05]	
Mean length of utterance in morphemes	3.11	3.48	–0.39	.01	–0.28	.04
	(0.85)	(0.91)	[–0.66, –0.12]		[–0.56, 0.01]	
IPSyn total score	60.90	65.39	–4.52	.03	–2.86	.15
	(13.53)	(12.05)	[–8.55, –0.51]		[–6.78, 1.05]	
IPSyn Noun Phrase subscale	17.44	18.40	–0.99	.02	–0.58	.16
	(2.81)	(2.48)	[–1.80, –0.17]		[–1.40, 0.24]	
IPSyn Verb Phrase subscale	20.17	21.14	–1.07	.20	–0.39	.64
	(5.16)	(5.24)	[–2.73, 0.58]		[–2.03, 1.25]	
IPSyn Question/Negation subscale	7.54	7.49	0.01	.99	0.59	.54
	(7.49)	(3.71)	[–1.62, 1.64]		[–1.31, 2.49]	
IPSyn Sentence Structures subscale	16.26	18.24	–1.97	.003	–1.81	.01
	(4.62)	(3.89)	[–3.26, –0.69]		[–3.10, –0.52]	
Lexical variety						
Total number of words	217.78	251.88	–33.96	.02	–27.57	.08
	(88.16)	(105.16)	[–63.35, –4.56]		[–58.69, 3.56]	
Number of different word roots	80.63	91.85	–11.29	.004	–10.04	.01
	(23.43)	(26.37)	[–18.94, –3.63]		[–17.93, –2.14]	
Number of nouns	58.70	61.82	–3.53	.42	–3.09	.50
	(26.96)	(27.15)	[–12.08, 5.02]		[–12.14, 5.96]	
Number of verbs	35.41	41.65	–6.25	.046	–5.33	.09
	(17.89)	(22.18)	[–12.41, –0.10]		[–11.46, 0.81]	
Number of adjectives	7.38	8.60	–1.26	.18	–1.02	.30
	(6.30)	(6.00)	[–3.12, 0.60]		[–2.95, 0.92]	
Number of prepositions	10.84	13.38	–2.39	.04	–2.25	.08
	(7.88)	(7.57)	[–4.69, –0.09]		[–4.74, 0.23]	

Note. CI = 95% confidence interval; coef. = regression coefficient; IPSyn = Index of Productive Syntax.

within or above 1.5 *SD*s of the term mean on MLUm and TNW, and six (four born < 30 weeks, two born at term) scored ≥ 1.5 *SD*s below the term mean on MLUm or TNW. In total, this meant that 16/87 (18%) preterm children and 10/80 (13%) term-born children scored ≥ 1.5 *SD*s below the term mean on at least one language measure.

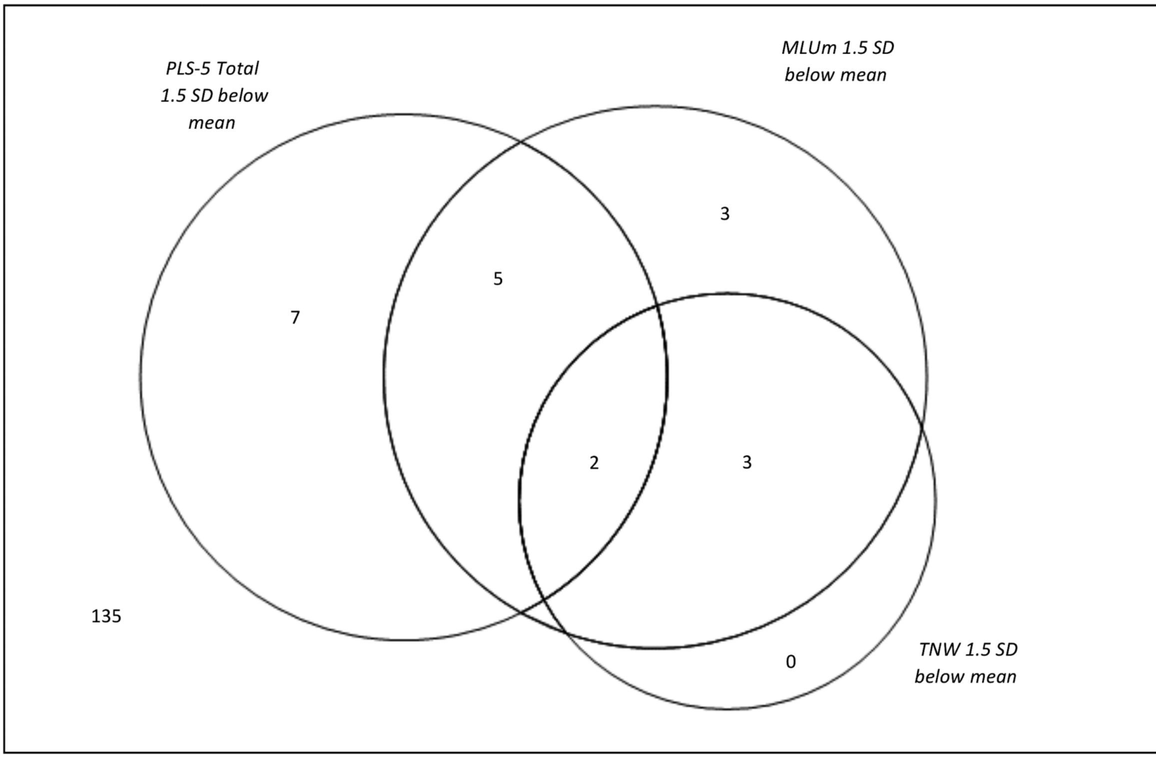
Correlations between the PLS-5 and LSA are shown in Table 4. PLS-5 results include nine children who completed the Auditory Comprehension scale, but not the Expressive Communication scale. There were statistically significant correlations between PLS-5 scores and all but three of the LSA metrics reported (IPSyn Question/Negation subscale, number of nouns, and number of adjectives) across the entire group, yet all correlations were weak or moderate in magnitude (Cohen, West, & Aiken, 1984).

For the purposes of comparing the clinical utility of LSA and the PLS-5, it is also noted that, of the 31 children who produced fewer than 50 utterances for LSA, 24 (77%) were able to complete a full PLS-5, and of the 167 children who produced more than 50 utterances, 12 (7%) did not complete a full PLS-5. Furthermore, 20/188 children (15/92 born < 30 weeks, 5/96 born at term) scored at least 1.5 *SD*s below the term mean on the Auditory Comprehension scale of the PLS-5, and 7/179 (5/85 born < 30 weeks, 2/94 born at term) scored at least 1.5 *SD*s below the term mean on the Expressive Communication scale.

Discussion

In this study of language outcomes in 167 children (87 born at < 30 weeks of gestational age, 80 born at term),

Figure 1. Overlap in language classification between measures.



PLS-5 = Preschool Language Scales–Fifth Edition; MLUm = mean length of utterance in morphemes; TNW = total number of words.

Table 4. Correlation between Preschool Language Scales scores and language sample analysis.

	Total sample		Children born < 30 weeks		Children born at term	
Variable	**PLS-AC**	**PLS-EC**	**PLS-AC**	**PLS-EC**	**PLS-AC**	**PLS-EC**
Morphosyntactic complexity						
Mean length of utterance in words	.41*	.45*	.52*	.54*	.21	.30*
Mean length of utterance in morphemes	.39*	.43*	.52*	.55*	.19	.28*
IPSyn total score	.24*	.30*	.39*	.38*	–.05	.18
IPSyn Noun Phrase subscale	.36*	.31*	.55*	.53*	–.00	.08
IPSyn Verb Phrase subscale	.24*	.24*	.42*	.40*	–.04	.10
IPSyn Question/Negation subscale	.02	.05	.10	.06	–.21	.05
IPSyn Sentence Structure subscale	.25*	.33*	.29*	.30*	.10	.30*
Lexical variety						
Total number of words	.27*	.29*	.36*	.29*	.13	.24*
Number of different word roots	.34*	.40*	.39*	.40*	.24*	.35*
Number of nouns	–.05	–.07	–.03	–.07	–.12	–.10
Number of verbs	.20*	.21*	.31*	.23*	.04	.15
Number of adjectives	.14	.08	.14	.05	.12	.06
Number of prepositions	.23*	.24*	.30*	.27*	.06	.16

Note. PLS-AC = Preschool Language Scales–Fifth Edition Auditory Comprehension score; PLS-EC = Preschool Language Scales–Fifth Edition Expressive Communication score; IPSyn = Index of Productive Syntax.

*Indicates a statistically significant correlation.

we found that the conversational language of children born at < 30 weeks differed considerably from that of their term-born peers across a range of morphosyntactic and lexical metrics, notably MLU in words and morphemes, sentence structures, and number of different word roots, but that the groups were similar in number of utterances, total number of completed words, words per minute, maze words as a percentage of total words, verb phrases, questions/negations, and number of nouns and adjectives. Furthermore, we identified weak to moderate correlations between a widely used standardized language measure (the PLS-5) and conversational LSA outcomes (such as MLUm and TNW) in this population. The incidence of low language performance (scoring ≥ 1.5 *SD*s below the term mean PLS-5 total score, MLUm, or TNW) in children who produced > 50 utterances over 10 min in the term group (13%) was comparable to previous incidence studies using direct assessment methods (Raghavan et al., 2018; Reilly et al., 2010), whereas the incidence of low language in the preterm group (18%) was somewhat lower than expected (Foster-Cohen, Friesen, Champion, & Woodward, 2010), perhaps reflecting relatively low medical acuity in this study group.

Our findings are consistent with some of the literature in this area suggesting that children born preterm demonstrate reductions in conversational MLU (Félix et al., 2017; Imgrund et al., 2019; Le Normand et al., 1995; Rice et al., 1999), sentence complexity (Grunau et al., 1990; Imgrund et al., 2019), and lexical diversity, particularly when it comes to verbs (Imgrund et al., 2019; Le Normand & Cohen, 1999), although these findings are by no means universal (Crosbie et al., 2011; Feldman et al., 1994; Grunau et al., 1990; Mahurin-Smith et al., 2014). Demographic differences between study participants may account for some of this variation—for example, two of the studies with significantly different results to ours examined an older group of children, which may reflect differences in the clinical validity of LSA in different age groups (Crosbie et al., 2011; Mahurin-Smith et al., 2014).

In this study, we reported on a range of LSA metrics in the largest cohort of preterm children to date and compared outcomes with a similar term-born population. These results are therefore likely to be representative of language abilities in 3-year-old children born < 30 weeks with similar medical histories to our cohort. Importantly, we have identified specific areas of language deficit in the preterm population that may inform language assessment and therapy goals. For example, depending on individual assessment results, speech-language pathologists working with preschoolers born preterm may choose to prioritize sentence structures, grammatical complexity, and verb vocabulary.

Several values (IPSyn Total Score, IPSyn Noun Phrase subscale, TNW, number of verbs, number of prepositions) differed significantly between groups on uncontrolled analyses; however, the difference was reduced when analyses controlled for sex, high social risk, multilingualism, and diagnosed neurodevelopmental disorders. This may suggest that variables associated with preterm birth, rather than preterm birth itself, place preterm-born children at higher risk of poor language outcomes. An a priori study design would be best placed to rigorously evaluate this possibility and may inform targeting of resources to higher risk children.

Another interesting finding of this study was that correlations between a standardized assessment tool (the PLS-5) and LSA were moderate at best. While some studies have found similarly weak correlations—including insignificant correlations—between LSA and standardized assessment tools (Ebert & Scott, 2014), others have found stronger relationships (Owens & Pavelko, 2017), at least for some measures. However, our results were more congruent than those of previous studies that contrasted standardized assessment with LSA in the preterm population (Crosbie et al., 2011; Imgrund et al., 2019; Mahurin-Smith et al., 2014). This may be a result of the relatively young age of our participants or may be related to our use of the PLS-5, as opposed to different standardized assessments (e.g., the Clinical Evaluation of Language Fundamentals–Fifth Edition) used in other studies. Nevertheless, these results and, in particular, our findings reported in Figure 1, could support previous literature stating that standardized language assessment and LSA give different information about a child's language abilities and should be used complementarily rather than relying on use of one or the other. Furthermore, these weak correlations suggest that the existing literature on preterm language development—based largely on formal assessment measures—is incomplete. Measurement of LSA at a range of ages would facilitate a better understanding of language in this population and may lead to the application of more targeted interventions.

It is possible that this study overestimates the language abilities in both the term and preterm groups in an LSA context, given that 10 children born preterm and 21 children born at term failed to produce more than 50 utterances and were thus omitted. As shown in Supplemental Material S1, this decision led to the omission of six of the 25 children in the study group from multilingual backgrounds (24%), two of the seven children with a neurodevelopmental disorder (29%), eight of 58 children at high social risk (14%), and 14 of 101 boys (14%). While this did not result in any unexpected imbalance between groups (see Table 2), it may have resulted in the omission of some children with clinically significant language problems. However, only three (10%) of the children with language samples of < 50 utterances performed ≥ 1.5 *SD*s below the term-born mean on the PLS-5. Future studies may choose to use a more structured elicitation protocol, allow a longer time for eliciting a language sample, or use parent–caregiver interaction in a home environment to increase the number of participants achieving the minimum utterances required for inclusion. These inconsistencies in assessment completion also illustrate the importance of considering both structured and unstructured assessment contexts, as 77% (24/31) of the participants who produced fewer than 50 utterances in the language sample were able to complete the PLS-5 and, conversely, 7% (12/167) of the children who produced a language sample of ≥ 50 utterances did not complete the full PLS-5. Other recommendations for future studies would be including a measure of elicitation fidelity, as we cannot rule out differences in how clinicians interpreted the provided protocol, and recording the reasons for noncompletion of direct assessment.

In this article, we provide a comprehensive report of conversational language skills in preschool-age children born preterm. Results demonstrated that children born < 30 weeks differ from their term-born peers across a number of conversational language variables and that LSA provides important data to complement formal language assessment in the preterm population. Future research exploring the trajectories and predictors of conversational language ability in children born preterm may assist researchers and clinicians to better understand the language profile that characterizes this population. For clinicians, these findings suggest that neither formal language assessment nor LSA in isolation provides sufficient data to describe language outcomes in all children born preterm who require early language intervention. Rather, these methods are best used complementarily to provide a richer linguistic profile for the individual child. With this holistic language profile, clinicians can generate informed and targeted intervention goals and strategies, informed by the profile we have reported in this unique population.

Acknowledgments

Funding for this study was received from the National Health and Medical Research Council (Career Development Award 607315 and Practitioner Fellowship 1105008 [awarded to A. M.]; Career Development Fellowship 11098714, Centre for Research Excellence in Newborn Medicine Grant 1060733 [awarded to A. S.]; Centre for Research Excellence in Newborn Medicine Grant 1060733 [awarded to K. S.]) and an Australian Government Research Training Program Scholarship (awarded to K. S.). Murdoch Children's Research Institute research was supported by the Victorian government's Operational Infrastructure Support Program.

We would like to acknowledge the Victorian Infant Brain Studies team, who contributed to recruitment, data collection, and analysis, and the families and children who participated in this study. Thanks also to Cristina Mei and Jodie Smith, who contributed to data collection.

References

Barre, N., Morgan, A., Doyle, L. W., & Anderson, P. J. (2011). Language abilities in children who were very preterm and/or very low birth weight: A meta-analysis. *The Journal of Pediatrics, 158*(5), 766–774.e1. https://doi.org/10.1016/j.jpeds.2010.10.032

Botting, N., Toseeb, U., Pickles, A., Durkin, K., & Conti-Ramsden, G. (2016). Depression and anxiety change from adolescence to adulthood in individuals with and without language impairment. *PLOS ONE, 11*(7), e0156678. https://doi.org/10.1371/journal.pone.0156678

Cohen, P., West, S. G., & Aiken, L. S. (1984). *Applied multiple regression/correlation analysis for the behavioral sciences.* New York, NY: Routledge.

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., & Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *International Journal of Language & Communication Disorders, 53*(2), 237–255. https://doi.org/10.1111/1460-6984.12338

Craig, H. K., Evans, J. L., Meisels, S. J., & Plunkett, J. W. (1991). Linguistic production abilities of 3-year-old children born premature with low birth weight. *Journal of Early Intervention, 15*(4), 326–337. https://doi.org/10.1177/105381519101500402

Crosbie, S., Holm, A., Wandschneider, S., & Hemsley, G. (2011). Narrative skills of children born preterm. *International Journal of Language & Communication Disorders, 46*(1), 83–94. https://doi.org/10.3109/13682821003624998

DeThorne, L. S., Deater-Deckard, K., Mahurin-Smith, J., Coletto, M.-K., & Petrill, S. A. (2011). Volubility as a mediator in the associations between conversational language measures and child temperament. *International Journal of Language & Communication Disorders, 46*(6), 700–713. https://doi.org/10.1111/j.1460-6984.2011.00034.x

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

Feldman, H. M., Janosky, J. E., Scher, M. S., & Wareham, N. L. (1994). Language abilities following prematurity, periventricular brain injury, and cerebral palsy. *Journal of Communication Disorders, 27*(2), 71–90. https://doi.org/10.1016/0021-9924(94)90035-3

Félix, J., Santos, M. E., & Benítez-Burraco, A. (2017). Spontaneous language of preterm children aged 4 and 5 years. *Revista CEFAC, 19*(6), 742–748. https://doi.org/10.1590/1982-021620171968017

Foster-Cohen, S. H., Friesen, M. D., Champion, P. R., & Woodward, L. J. (2010). High prevalence/low severity language delay in preschool children born very preterm. *Journal of Developmental & Behavioral Pediatrics, 31*(8), 658–667. https://doi.org/10.1097/DBP.0b013e3181e5ab7e

Grunau, R. V. E., Kearney, S. M., & Whitfield, M. F. (1990). Language development at 3 years in pre-term children of birth weight below 1000 g. *International Journal of Language & Communication Disorders, 25*(2), 173–182. https://doi.org/10.3109/13682829009011972

Guarini, A., Sansavini, A., Fabbri, C., Savini, S., Alessandroni, R., Faldella, G., & Karmiloff-Smith, A. (2010). Long-term effects of preterm birth on language and literacy at eight years. *Journal of Child Language, 37*(4), 865–885. https://doi.org/10.1017/S0305000909990109

Guo, L.-Y., & Eisenberg, S. (2015). Sample length affects the reliability of language sample measures in 3-year-olds: Evidence from parent-elicited conversational samples. *Language, Speech, and Hearing Services in Schools, 46*(2), 141–153. https://doi.org/10.1044/2015_LSHSS-14-0052

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Hassanali, K., Liu, Y., Iglesias, A., Solorio, T., & Dollaghan, C. (2014). Automatic generation of the Index of Productive Syntax for child language transcripts. *Behavior Research Methods, 46*(1), 254–262. https://doi.org/10.3758/s13428-013-0354-x

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.

Hewitt, L. E., Hammer, C. S., Yont, K. M., & Tomblin, J. B. (2005). Language sampling for kindergarten children with and without SLI: Mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders, 38*(3), 197–213. https://doi.org/10.1016/j.jcomdis.2004.10.002

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

Islam, U. A., Poole, K. L., Schmidt, L. A., Ford, J., Saigal, S., & Van Lieshout, R. J. (2018). Childhood language skills and adolescent self-esteem in preterm survivors. *Journal of Child Health Care, 22*(1), 34–45. https://doi.org/10.1177/1367493517739158

Le Normand, M.-T., & Cohen, H. (1999). The delayed emergence of lexical morphology in preterm children: The case of verbs. *Journal of Neurolinguistics, 12*(3–4), 235–246. https://doi.org/10.1016/S0911-6044(99)00016-0

Le Normand, M.-T., Vaivre-Douret, L., & Delfosse, M. J. (1995). Language and motor development in pre-term children: Some questions. *Child: Care, Health and Development, 21*(2), 119–133. https://doi.org/10.1111/j.1365-2214.1995.tb00414.x

Mahurin-Smith, J., DeThorne, L. S., Logan, J. A. R., Channell, R. W., & Petrill, S. A. (2014). Impact of prematurity on language skills at school age. *Journal of Speech, Language, and Hearing Research, 57*(3), 901–916. https://doi.org/10.1044/1092-4388(2013/12-0347)

Miller, J. F., Gillon, G. T., & Westerveld, M. F. (2015). Systematic Analysis of Language Transcripts 16 (SALT 16)—New Zealand/Australia. Madison, WI: SALT Software.

Mürner-Lavanchy, I. M., Kelly, C. E., Reidy, N., Doyle, L. W., Lee, K. J., Inder, T., . . . Anderson, P. J. (2018). White matter microstructure is associated with language in children born very preterm. *NeuroImage: Clinical, 20,* 808–822. https://doi.org/10.1016/j.nicl.2018.09.020

Northam, G. B., Liégeois, F., Tournier, J.-D., Croft, L. J., Johns, P. N., Chong, W. K., . . . Baldeweg, T. (2012). Interhemispheric temporal lobe connectivity predicts language impairment in adolescents born preterm. *Brain, 135*(12), 3781–3798. https://doi.org/10.1093/brain/aws276

Owens, R. E., & Pavelko, S. L. (2017). Relationships among conversational language samples and norm-referenced test scores. *Clinical Archives of Communication Disorders, 2*(1), 43–50. https://doi.org/10.21849/cacd.2017.00052

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Raghavan, R., Camarata, S., White, K., Barbaresi, W., Parish, S., & Krahn, G. (2018). Population health in pediatric speech and language disorders: Available data sources and a research agenda for the field. *Journal of Speech, Language, and Hearing Research, 61*(5), 1279–1291. https://doi.org/10.1044/2018_JSLHR-L-16-0459

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E. L., Prior, M., Cini, E., . . . Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from Early Language in Victoria study. *Pediatrics, 126*(6), e1530–e1537. https://doi.org/10.1542/peds.2010-0254

Rice, M. L., Redmond, S. M., & Hoffman, L. (2006). Mean length of utterance in children with specific language impairment and in younger control children shows concurrent validity and stable and parallel growth trajectories. *Journal of Speech, Language, and Hearing Research, 49*(4), 793–808.

Rice, M. L., Spitz, R. V., & O'Brien, M. (1999). Semantic and morphosyntactic language outcomes in biologically at-risk children. *Journal of Neurolinguistics, 12*(3–4), 213–234. https://doi.org/10.1016/S0911-6044(99)00015-9

Saigal, S., & Doyle, L. W. (2008). An overview of mortality and sequelae of preterm birth from infancy to adulthood. *The Lancet, 371*(9608), 261–269. https://doi.org/10.1016/S0140-6736(08)60136-1

Sanchez, K., Boyce, J. O., Mei, C., St. John, M., Smith, J., Spittle, A. J., & Morgan, A. T. (2019). Communication in children born very preterm: A prospective cohort study. *Developmental Medicine and Child Neurology*. https://doi.org/10.1111/dmcn.14360

Scarborough, H. S. (1990). Index of Productive Syntax. *Applied Psycholinguistics, 11*(1), 1–21. https://doi.org/10.1017/S0142716400008262

Spittle, A. J., Thompson, D. K., Brown, N. C., Treyvaud, K., Cheong, J. L. Y., Lee, K. J., . . . Anderson, P. J. (2014). Neurobehaviour between birth and 40 weeks gestation in infants born < 30 weeks gestation and parental psychological wellbeing: Predictors of brain development and child outcomes. *BMC Pediatrics, 14*(1), 111. https://doi.org/10.1186/1471-2431-14-111

StataCorp. (2015). Stata statistical software: Release 14. College Station, TX: StataCorp.

Zimmerman, E. (2018). Do infants born very premature and who have very low birth weight catch up with their full term peers in their language abilities by early school age? *Journal of Speech, Language, and Hearing Research, 61*(1), 53–65. https://doi.org/10.1044/2017_JSLHR-L-16-0150

Wiig, E. H., Semel, E.& Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals Australian and New Zealand–Fifth Edition (CELF-5)*. San Antonio, TX: Pearson.

Zimmerman, I., Steiner, V., & Pond, R. (2011). *Preschool Language Scales (Australian and New Zealand language adapted)–Fifth Edition*. San Antonio, TX: The Psychological Corporation.