Research in Developmental Disabilities 148 (2024) 104711

Contents lists available at ScienceDirect

Research in Developmental Disabilities

journal homepage: www.elsevier.com/locate/redevdis

Lexical skills and gesture use: A comparison between expressive and receptive/expressive late talkers

Caterina Verganti [a,*], Chiara Suttora [a], Mariagrazia Zuccarini [b], Arianna Aceti [c,d], Luigi Corvaglia [c,d], Arianna Bello [e], M. Cristina Caselli [f], Annalisa Guarini [a,1], Alessandra Sansavini [a,*,1]

[a] Department of Psychology "Renzo Canestrari", University of Bologna, Italy
[b] Department of Education Studies "Giovanni Maria Bertin", University of Bologna, Italy
[c] Neonatal Intensive Care Unit, IRCCS Azienda Ospedaliero Universitaria Bologna, Italy
[d] Department of Medical and Surgical Sciences, University of Bologna, Italy
[e] Department of Education, Roma Tre University, Italy
[f] Institute of Cognitive Sciences and Technologies, CNR, Rome, Italy

ARTICLE INFO

Keywords:
late talkers
gestures
lexical skills
picture naming game
MacArthur–Bates Communicative Development Inventory

ABSTRACT

Background: Studies on late talkers (LTs) highlighted their heterogeneity and the relevance of describing different communicative profiles.
Aims: To examine lexical skills and gesture use in expressive (E-LTs) vs. receptive-expressive (R/E-LTs) LTs through a structured task.
Methods and procedures: Forty-six 30-month-old screened LTs were distinguished into E-LTs (n= 35) and R/E-LTs (n= 11) according to their receptive skills. Lexical skills and gesture use were assessed with a Picture Naming Game by coding answer accuracy (correct, incorrect, no response), modality of expression (spoken, spoken-gestural, gestural), type of gestures (deictic, representational), and spoken-gestural answers' semantic relationship (complementary, equivalent, supplementary).
Outcomes and results: R/E-LTs showed lower scores than E-LTs for noun and predicate comprehension with fewer correct answers, and production with fewer correct and incorrect answers, and more no responses. R/E-LTs also exhibited lower scores in spoken answers, representational gestures, and equivalent spoken-gestural answers for noun production and in all spoken and gestural answers for predicate production.
Conclusions and implications: Findings highlighted more impaired receptive and expressive lexical skills and lower gesture use in R/E-LTs compared to E-LTs, underlying the relevance of assessing both lexical and gestural skills through a structured task, besides parental questionnaires and developmental scales, to describe LTs' communicative profiles.

* Corresponding authors.
E-mail addresses: caterina.verganti2@unibo.it (C. Verganti), alessandra.sansavini@unibo.it (A. Sansavini).
[1] Annalisa Guarini and Alessandra Sansavini share last co-authorship

https://doi.org/10.1016/j.ridd.2024.104711
Received 30 November 2023; Received in revised form 5 February 2024; Accepted 22 February 2024
Available online 22 March 2024

What this paper adds?

The present study addresses a notable research gap on late talkers, by investigating the contribution of lexical skills and gesture use in characterizing different communicative profiles. Specifically, the study shows that receptive/expressive late talkers were characterized by more weaknesses than expressive late talkers not only in receptive skills, but also in expressive skills, gesture use, and cognitive development. Furthermore, they showed lower lexical and gestural abilities in retrieving the meaning of the target word both for nouns and predicates and more difficulties in completing the naming task, particularly for predicate production. Overall, these findings emphasize the need for a more comprehensive approach to late talkers' assessment of both lexical and gestural skills through different tools, such as structured tasks alongside parental questionnaires and developmental scales. This has relevant implications for the early identification of late talkers at higher risk for subsequent language and/or learning impairment. Additionally, it emphasizes the need for tailored strategies and interventions for this at-risk population.

1. Introduction

Children with language delay, defined as late talkers (LTs), are children aged 18-35 months with a limited expressive vocabulary in the absence of any known primary cause, such as neurological, sensorial, socio-emotional or cognitive deficits (Hawa & Spanoudis, 2014). The criteria for identifying LTs are an expressive vocabulary size of fewer than 50 words or $\leq 10^{th}$ percentile, as measured with parental questionnaires (e.g. MacArthur-Bates Communication Developmental Inventories, MB-CDI, Fenson et al., 2007), at 24 months, and/or the absence of word combination at 30 months of age (Desmarais et al., 2008).

LTs' prevalence varies between 9-21% in 2-3-year-old children (Korpilahti et al., 2016). Among them, more than half catch up by age 3 (Dale et al., 2003), while the others may receive a diagnosis of language and/or learning impairments at later ages (Sansavini, Favilla et al., 2021). LTs represent a heterogeneous group (Desmarais et al., 2010); their language profiles and outcomes may depend on biological and environmental risk factors (Collisson et al., 2016; Zambrana et al., 2014), and early communicative-linguistic abilities.

Family history of language or learning impairments, male gender, preterm birth, and low birth weight have been documented as biological risk factors (Collisson et al., 2016; Reilly et al., 2007; Sansavini, Guarini, Savini et al., 2011). Specifically, preterm birth has been widely recognized as a condition that can negatively affect language development with cascading effects on later development in multiple domains, depending on neonatal immaturity, pre-, peri- and postnatal complications, neurological alterations, and the association with other biological and environmental risk factors (Sansavini, Guarini & Caselli, 2011). Several studies conducted on preterm infants have shown that the risk for language delay increases as the level of neonatal immaturity increases, so it is greater among extremely preterm infants (gestational age $<$ 28 weeks) and very preterm infants (gestational age $<$ 32 weeks) compared to full-term children (Barre et al., 2011; Guarini et al., 2016; Nguyen et al., 2018; Sansavini et al., 2010, 2014, 2015; Woodward et al., 2009). By contrast, fewer and inconsistent findings are available for low-risk preterm children (i.e., those having a gestational age $<$ 37 weeks and lack of major cerebral damage or severe perinatal complications; Pérez-Pereira, 2021; Sansavini, Zuccarini et al., 2021; Suttora et al., 2020, 2022; Zambrana et al., 2021; Zuccarini et al., 2023).

Low parental education level and socioeconomic status and poor parental linguistic input have been proven as environmental risk factors (Chilosi et al., 2019; Suttora et al., 2021). However, biological and environmental factors showed limited power in predicting a persistent language delay up to preschool age (Fisher, 2017; Sansavini, Favilla et al., 2021).

By contrast, early communicative-linguistic abilities seem to better predict and differentiate LTs' outcomes in the short- and medium-term (Sansavini, Favilla et al., 2021). It is thus relevant to understand early communicative-linguistic differences among LTs to more accurately detect those at higher risk of persistent language difficulties.

1.1. Late talkers' linguistic skills

Although LTs show a primary delay in expressive language, a significant proportion of them display difficulties also in comprehension (Carson et al., 1998). Since receptive skills constitute the basis of expressive language development, recent studies have examined both these abilities to better understand LTs' interindividual variability and whether the association of poor expressive and receptive skills is a more robust and reliable predictor of persistent language delays (Bello et al., 2018; Chilosi et al., 2019; Thal et al., 2013).

Concerning LTs' interindividual variability, Desmarais et al. (2010) identified three distinct linguistic profiles in a group of 18-35-month-old LTs: the first with impaired language comprehension, expression, and communicative engagement; the second with less impaired comprehension skills but impaired language expression and communicative engagement; the third with higher scores in language comprehension and expression, but still characterized by an expressive language delay. Similarly, a recent study (Zuccarini et al., 2023), examining a sample of 30-month-old LTs (born either full-term or low-risk preterm), identified three distinct communicative profiles: *severe*, exhibiting a limited receptive and expressive vocabulary size; *moderate,* with a less impaired receptive vocabulary size but a limited expressive vocabulary; and *mild*, with a receptive vocabulary size similar to that of the *moderate* one and a partially limited expressive vocabulary size. The *severe* and *moderate* profiles also had significantly lower noun and predicate production, grammar, and cognitive skills than the *mild* one. In addition, the *severe* profile had lower verbal imitation and pragmatic skills, proving that children with both expressive and receptive language delays had the most widespread delays across linguistic and non-linguistic domains.

Concerning the predictive power of expressive and receptive language delays on later language outcomes, O'Neill et al. (2019),

examining longitudinally a group of LTs, found that receptive skills at 24 months showed the most pervasive correlations with language measures at 4-5 years of age and predicted language outcomes at the same age. Additionally, a study by Chilosi et al. (2019), following longitudinally a group of 28-month-old LTs outlined three distinct groups based on their outcomes: *late bloomers*, exhibiting a higher lexical quotient and syntactic comprehension than the other two groups and catching up their expressive delay at three years of age; *slow learners,* lagging behind late bloomers, but showing a higher lexical quotient and expressive grammar than children with persistent language delay at three years of age and recovering language delay by four years of age; and *children with persistent language delay* who maintained poor receptive and expressive skills up to four years of age.

The studies reported above confirmed the wide heterogeneity among LTs, suggesting a crucial role of receptive language skills in describing heterogeneity and predicting outcomes. However, more extensive evidence is needed on LTs' language profiles to early detect those exhibiting persistent language delay.

1.2. Late talkers' gestural skills

Some studies have investigated gestures as early markers of language delay, highlighting the importance of an early assessment of gesture production to detect children at risk for persistent language delays. For instance, Lüke et al. (2017) showed that a scarce use of pointing with the index finger at 12 months predicted language delay at 24 months. Furthermore, Sansavini et al. (2019) found very limited or absent pointing gestures in 18-month-old children belonging to two at-risk populations, i.e., children born preterm and siblings of children with autism spectrum disorder, identified as LTs at 30-36 months. Converging evidence was found by Manwaring et al. (2019) in 18-month-old children, identified as LTs (12 of them with autism spectrum disorder) at 36 months, which showed lower receptive and expressive language scores and fewer deictic (i.e., pointing, showing, giving, reaching) and representational gestures, specifically conventional ones (i.e., gestures which have a culturally defined form and meaning, such as waving ''hi'' or ''bye,'' shaking head ''no'', etc.) than typically developing children.

Few studies have investigated gestural communication in children already identified as LTs. As regards LTs' gestural interindividual variability, O'Neill & Chiat (2015) found poorer deictic and representational gestures (including conventional ones and iconic ones, the latter used to depict actions or attributes of something, such as waving an arm in the air to indicate airplane flying) and symbolic comprehension of gestures in receptive/expressive 2–3-year-old LTs than expressive only LTs, highlighting that poorer receptive linguistic skills were associated with a lower frequency of communicative gestures. Less frequent use of pointing among LTs was also found by Zuccarini et al. (2023) in children with a *severe* communicative profile compared to those with a *moderate* communicative profile at 30 months. Furthermore, Suttora et al. (2022), examining 30-month-old LTs during a parent-child book-sharing session, found that lower spoken-gestural combinations, particularly complementary (i.e., combinations in which the gesture identifies the referent and the word labels it) and supplementary (i.e., combinations in which the gesture and the word convey two different meanings so that each of the combined elements adds information to the other one) forms, were associated with lower receptive and expressive skills.

Considering LTs' developmental outcomes, a seminal study (Thal & Tobias, 1992) showed that LTs with a transient language delay, caught up by three years of age, compensated for their language delay between 18 and 28 months by producing more deictic and representational gestures (mainly conventional ones) during parent-infant play than typically developing children and children with persistent language delay. Additional findings were recently achieved by Rinaldi et al. (2022) who followed LTs longitudinally from 2 to 4 years of age. Three groups, i.e., typically developing children, children with transient language delay, and children with developmental language disorder, were identified. Groups did not differ in the number of spontaneous gestures and spoken-gestural semantic relationships at 32 months. However, children with transient language delay produced more often unimodal gestural utterances than typically developing children and used representational gestures to replace spoken words for which they could access the meaning but not produce the word yet. Conversely, children who developed a developmental language disorder did not show a unimodal gestural advantage and used representational gestures more frequently than the other children before the spoken answer to access the word meaning. Another study (O'Neill et al., 2019) showed that receptive language, frequency of gesture use, and symbolic comprehension, assessed at 2-3 years, were strong predictors of language outcomes at 4-5 years, with poorer outcomes for LTs having receptive/expressive language delay.

Apart from the works reported above, literature about LTs' gesture use and language profile is still scarce.

1.3. Aim of the present study

As LTs are characterized by significant heterogeneity and considering the relevance of an early assessment of gesture production, the current study aimed to examine noun and predicate comprehension, production, and gesture use according to the type of language delay (i.e., Expressive Late Talkers, E-LTs vs Receptive/Expressive Late Talkers, R/E-LTs). To this aim, answer accuracy in noun and predicate comprehension and production, modality of expression, types of gestures, and spoken-gestural semantic relationship were analyzed through a structured task. The present study hypothesized more impaired noun and predicate comprehension and production skills, and a lower gesture use in R/E-LTs than in E-LTs.

2. Method

2.1. Participants

Forty-six children (29 males, 17 females), consisting of 20 low-risk preterm and 26 full-term children, were identified as LTs through a screening project, conducted on children born at the Policlinico di Sant'Orsola Hospital of the IRCCS Azienda Ospedaliero-Universitaria di Bologna, Italy (for details about the screening, see Sansavini, Zuccarini et al., 2021; Zuccarini et al., 2023), at about 30 months of age (*M*= 30.41 months; *SD*= 1.15; range= 26.76 - 33.06), and assessed at about 31 months of age (*M*= 31.12 months; *SD*= 1.41; range= 27.14 - 34.26). Concerning low-risk preterm children's screening and assessment, age was corrected for weeks of prematurity to consider their level of neurobiological maturation, as done in previous studies (Sansavini, Zuccarini et al., 2021; Zuccarini et al., 2023).

Inclusion criteria consisted of (a) being monolingual or mainly exposed (> 65% of daily exposure) to the Italian language; (b) being born either full-term (i.e., with a gestational age ≥ 37 weeks) or low-risk preterm (i.e., with a gestational age < 37 weeks and lack of major cerebral damage or severe perinatal complications); (c) not having major cerebral damage (i.e., intraventricular haemorrhage, hydrocephalus, periventricular leukomalacia) or congenital malformations, visual (i.e., blindness, retinopathy of prematurity), hearing (i.e., deafness) or motor impairments, or severe cognitive deficits (identified by a Bayley Scales of Infant and Toddler Development, Third Edition- BSID-III - cognitive composite score < 70); (d) not having a diagnosis or clinical sign of neurodevelopmental disorders (e.g., autism spectrum disorder); and (e) having an expressive vocabulary ≤ 10^{th} percentile at the Italian version of the MacArthur Bates Communicative Development Inventories (MB-CDI), Words and Sentences (W&S) - Short Form (Caselli et al., 2015).

The 46 assessed children were distinguished into two groups according to the type of language delay, E-LTs or R/E-LTs (for identification criteria, see the Procedure paragraph). The biological and sociodemographic characteristics of the children and parents are presented in Table 1. Children of the two groups were similar in all biological and sociodemographic variables, except for their attendance at a childcare centre, which was significantly higher for E-LTs. Concerning the twin condition, for each pair of twins either one or both were included in the sample if they satisfied the inclusion criteria. All 11 twins included in the study (24% of the whole sample) were born preterm. This could be expected as the twin condition is much more represented in the preterm than in the full-term population (Chauhan et al., 2010; Guarini et al., 2016). Specifically, 6 out of 11 twins belonged to 3 pairs of twins including both children, whereas the other 5 twins were single children from 5 pairs of twins, in which the other twin did not fit the inclusion criteria.

2.2. Tools

2.2.1. MacArthur–Bates communicative development inventories (MB-CDI) words and sentences (W&S) – short form

The Italian version of the MB-CDI W&S - Short Form was used to assess children's expressive vocabulary size. This reliable and widely used tool has been validated on children aged 18 to 36 months and has already been adopted in screening programs on 24- to

Table 1
Comparison Between Expressive and Receptive/Expressive Late Talkers' Biological and Sociodemographic Characteristics.

	E-LTs	R/E-LTs		
	(*n* = 35)	(*n* = 11)	Test	
Participants' Characteristics	*M/n (SD/%)*	*M/n (SD/%)*	χ^2/t	*p*
Neonatal Condition (low-risk preterm), *n (%)*	17 (48)	3 (27)	1.545	.214
Gestational Age (weeks), *M (SD)*	36.58 (3.44)	38.01 (3.14)	-1.228	.226
Birthweight (grams), *M (SD)*	2576.34 (852.23)	3035.09 (819.89)	-1.571	.123
Length of Hospital Stay (days)[a], *M (SD)*	14.71 (32.03)	7.36 (13.29)	0.737	.465
Gender (male), *n (%)*	24 (69)	5 (46)	1.920	.166
Firstborn, *n (%)*	17 (49)	4 (36)	0.503	.478
Twins, *n (%)*	8 (23)	3 (27)	0.090	.765
Otitis Media, *n (%)*	1 (3)	0 (0)	0.321	.571
Family History Language/Learning Disorders, *n (%)*	5 (14)	2 (18)	0.098	.754
Child-care Centre Attendance, *n (%)*	30 (86)	5 (45)	7.456	**.006**
Other Parent Input besides Italian, *n (%)*	3 (9)	3 (27)	0.118	.146
Mother's Age (years), *M (SD)*	38.94 (4.70)	38.73 (6.64)	0.100	.922
Father's Age (years), *M (SD)*	41.30 (6.33)	39.70 (6.08)	0.699	.489
Mothers with Educational Level >13 Years, *n (%)*	24 (69)	7 (63)	0.093	.761
Fathers with Educational Level >13 Years, *n (%)*	14 (40)	2 (18)	1.756	.185
Mother's Nationality (Italian), *n (%)*	33 (94)	9 (81)	1.639	.201
Father's Nationality (Italian), *n (%)*	31 (89)	9 (81)	0.337	.562

Note. [a] According to the Emilia-Romagna Region (BURERT - Bulletin of Emilia Romagna Region, 2019), and the international guidelines (Benitz et al., 2015; Phillips et al., 2013), in the absence of neonatal complications, the hospital discharge of a full-term newborn occurs between 48 and 72 hours of life, respectively for vaginal birth and caesarean section, with at least two nights of hospital stay. For late preterm newborns this period is extended based on criteria varying from centre to centre, with the decision generally based on reaching a specific gestational age, and/or achieving appropriate weight gain along with successful feeding. *M* = mean; *SD* = standard deviation. Significant results are in bold.

36-month-old children for identifying LTs (Bello et al., 2018; Sansavini, Zuccarini et al., 2021). The questionnaire consists of three sections respectively including a list of 100 words, a list of 12 pairs of sentences with different levels of morphosyntactic complexity, and a list of communicative behaviours. In the present study, the first section, was used. Parents were asked to check off the words their child spontaneously produced. The total number of words spontaneously produced indicated the child's expressive vocabulary size. Children were identified as LTs if their expressive vocabulary size was $\leq 10^{th}$ percentile at the MB-CDI W&S Short Form (Caselli et al., 2015), according to the Italian normative data (Rinaldi et al., 2019).

2.2.2. Bayley scales of infant and toddler development (BSID-III)

The BSID-III (Bayley et al., 2006; Ferri et al., 2015) was used to assess cognitive and language development. The BSID-III is a valid, reliable, and widely used tool for research and clinical practice that measures cognitive, linguistic, and motor skills in typically and atypically developing infants aged one to 42 months and 15 days. The BSID-III provides standardized cognitive and language composite scores, each with a mean of 100 and *SD* of 15, and standardized scaled scores, each with a mean of 10 and *SD* of 3, respectively for the receptive and expressive language subtests. Delay on the cognitive and language scales was defined as a standardized score $\leq$ –1.5 *SD* below the mean of the normative data, as done in previous studies (Lobo & Galloway, 2013).

2.2.3. Picture naming game (PiNG)

The picture naming game (PiNG, Bello et al., 2010, 2012) was used to assess noun and predicate comprehension and production. The PiNG has been validated on Italian toddlers aged 19 to 37 months old (Bello et al., 2012) and is suitable for assessing lexical skills in typical and atypical populations, such as LTs (Rinaldi et al., 2022; Sansavini et al., 2015). The PiNG consists of four subtests: Noun Comprehension, Noun Production, Predicate Comprehension, and Predicate Production, each containing 20 lexical targets and two training items. The noun subtests depict objects and tools (e.g., comb), transport (e.g., truck), animals (e.g., lion), food (e.g., banana), and clothing (e.g., gloves). The predicate subtests depict actions with objects (e.g., eating), actions without objects (e.g., swimming), adjectives (e.g., small), and adverbs of place (e.g., inside).

Following Bello et al.'s procedure (2010, 2012), the child sat next to the experimenter at a table where the pictures were placed. In the comprehension subtests, triplets were presented to the child, each containing a picture of the lexical target and two pictures of distracters, one semantically related and the other unrelated to the target. The experimenter asked the child to point to or touch the picture corresponding to the labelled word. Another chance was given when a child did not answer because of not paying attention. Concerning the production subtests, the experimenter asked the child to name the presented picture. For noun production, the experimenter asked, 'What is this?' (e.g., dog); for predicate production, 'What is he/she doing?' for an action word (e.g., eating), 'What is this like?' for a descriptive word (e.g., small), or 'Where is it?' for a locative word (e.g., inside). The experimenter repeated the question twice if the child did not answer or provided an incorrect answer.

According to the validation study (Bello et al., 2012), for children who did not answer or provided incorrect answers to five or more consecutive items on the production subtests, the administration of these subtests was interrupted.

Each subtest took about 10 minutes to be completed. All sessions were videotaped for later transcription and coding.

2.3. Procedure

Children were identified as LTs through a screening project at about 30 months of age (for details about the screening, see (Sansavini, Zuccarini et al., 2021) if their expressive vocabulary size was $\leq 10^{th}$ percentile at the MB-CDI W&S Short Form (Caselli et al., 2015), according to the Italian normative data (Rinaldi et al., 2019). Within about two weeks after the screening, the identified LTs were invited with their parents to the Developmental Psychology Lab, at the Department of Psychology "Renzo Canestrari", University of Bologna, for a direct assessment of their cognitive and linguistic skills with the BSID-III Cognitive and Language scales and the PiNG test (Bello et al., 2010, 2012).

Based on the results at the BSID-III Receptive Language subscale, the 46 LTs were distinguished into two groups: 35 E-LTs (i.e., expressive vocabulary size $\leq 10^{th}$ percentile at the MB-CDI W&S Short Form and a receptive communication scaled score $\geq$ –1.5 *SD* at

Table 2
Comparison Between Expressive and Receptive/Expressive Late Talkers' MB-CDI Expressive Vocabulary Size and BSID-III Language and Cognitive Scores.

	E-LTs	R/E-LTs		
	(n = 35)	(n = 11)	T-test	
Measures	*M (SD)*	*M (SD)*	*t*	*p*
MB-CDI W&S Short Form				
Expressive Vocabulary Size (words)	23.29 (12.67)	19.82 (11.94)	0.802	.427
BSID-III				
Language Composite Score	85.34 (7.61)	68.55 (4.20)	6.957	**<.001**
Receptive Scaled Score	9.06 (1.70)	4.82 (0.98)	7.845	**<.001**
Expressive Scaled Score	5.83 (1.60)	3.90 (1.38)	3.878	**<.001**
Cognitive Composite Score	91.29 (8.69)	84.55 (10.11)	2.159	**.036**

Note. The MB-CDI W&S Short Form consists of 100 words. *M* = mean; *SD* = standard deviation. Significant results are in bold.

the BSID-III); 11 R/E-LTs (i.e., expressive vocabulary size $\leq 10^{th}$ percentile at the MB-CDI W&S Short Form, and receptive communication scaled score $\leq$ –1.5 *SD* at the BSID-III).

Descriptive data for each group and comparisons between groups on the MB-CDI and BSID-III scores are reported in Table 2. The mean MB-CDI expressive vocabulary score of both groups fell below the 5^{th} percentile (Caselli et al., 2015). As regards the BSID-III, E-LTs' mean language composite score, receptive communication scaled score, and cognitive composite score fell within the normal range, whereas their mean expressive communication scaled score was –1.5 *SD* (Ferri et al., 2015). In contrast, R/E-LTs' mean language composite score and expressive communication scaled score were –2 *SD*, their receptive communication scaled score was –1.5 *SD,* and the cognitive composite score –1.5 *SD* (Ferri et al., 2015). All R/E-LTs' language and cognitive scores were significantly lower than those of the E-LTs (Table 2).

The study met ethical guidelines for human subject protections, including adherence to the legal requirements of the study country, and received formal approval by the local Research Ethical Committee Comitato Etico di Area Vasta Emilia Centro (approval numbers: 76/2013/U/Sper/AOUBo; EM 194–2017; EM 193–2018; EM 1229–2020). Parents of all children gave informed written consent for participation in the study, data analysis, and data publication.

2.4. Coding

We transcribed and coded all communicative exchanges between the experimenter and the child by starting when each picture was presented to the child and ending when it was removed.

2.4.1. Answer accuracy

The answers provided by children during the PiNG test were coded according to the validation study (Bello et al., 2012). Regarding the comprehension subtests, children's first answer was coded as: "*correct*" when they touched, showed, or pointed to the picture corresponding to the target word; "*incorrect*" when they touched, showed, or pointed to the picture corresponding to a different word; "*no response*" when they did not touch, show, or point to any picture. Regarding the production subtests, both verbal and gestural answers were coded. Children's spoken answers were coded in terms of accuracy as: *"correct",* when they provided the target word corresponding to the picture (phonologically altered forms of correct words, e.g., 'spinse' instead of 'spinge' [PUSH], and onomatopoeic forms, e.g., 'coccodé' instead of 'gallina' [chicken] were accepted); "*incorrect*" when they provided a word different from the target word. Incorrect answers were divided into incorrect words *semantically related* to the target word (e.g., 'sea' instead of 'beach'); and incorrect words *semantically unrelated* to the target, including visual–perceptual errors (e.g., 'milk' instead of 'flag') and answers referring to other elements of the photograph (e.g., 'hands' instead of 'washing hands'). Spoken answers were coded as "*unintelligible*" when the meaning could not be understood from the phonological form (e.g., 'gabola' instead of 'bicchiere' [glass]). "*No response*" was defined when the child provided no spoken answer. When the child did not provide the target answer on the first attempt, a second chance was given (without correcting the child), adopting a "best response" criterion. Due to fatigue or difficulties in answering the test, some children could not complete one or more items in the subtests and those items were coded as "*not administered"*.

2.4.2. Modality of expression

All answers provided during the production subtests were coded as unimodal spoken, bimodal spoken/gestural, or unimodal gestural, based on their modality of expression (Stefanini et al., 2009). *Unimodal spoken* answers included all (correct, incorrect, or unintelligible) spoken productions not associated with gestural answers (e.g., the child said 'table' without producing any gesture). *Bimodal spoken/gestural* answers included correct and incorrect spoken answers associated with gestural production (e.g., the child produced the word 'lion' and held his/her hands as claws in a menacing way near the head). *Unimodal gestural* answers included gestural answers not associated with a spoken response or associated with an unintelligible spoken answer (e.g., the child ran his/her hand through the hair with spread fingers for 'comb' without producing any word or producing an unintelligible word).

2.4.3. Types of gestures

All deictic and representational gestures produced by children during the production subtests were coded, including those occurring before and after the child's spoken answer, based on the following criteria (Stefanini et al., 2009): (a) the gesture had to be produced after the experimenter requested to name the picture; (b) the gesture could be performed with an empty hand or while holding the picture, and (c) the gesture must not be an imitation of any adult's preceding gesture. *Deictic gestures* included pointing, showing, and giving. Pointing was defined as an extension of the index finger directed to (or touching or patting) the target picture or an object in the surrounding environment. Instances of pointing with other fingers or with the whole hand were also included in this category. Showing was defined as an arm extension while holding an object (often the picture) in hand. In the case of giving, the object (i.e., the picture) was transferred to another person. *Representational gestures* included manual gestures and representational actions involving postures, body movements, and facial expressions spontaneously produced by children (Kendon, 2004). Four types of representational gestures were coded according to the following criteria (Marentette et al., 2016): *own-body*, i.e., a gesture that depicted the movement of an animate (e.g., the child swung his arms for the item 'swimming') or nonhuman animate agent (e.g., the child held his hands as claws in a menacing way for the item 'lion'); *hand-as-object*, i.e., a gesture in which the hand acts as the object itself, (e.g., the child held a hand over his head for the item 'umbrella'); *hand-as-hand*, i.e., a gesture in which the hand acts as a hand, representing how an object is held or manipulated and the motions associated (e.g., the child moved his fist to his mouth for the item 'fork'); *size-and-shape*, i.e., a gesture depicting an object's size or shape (e.g., the child brings the index and thumb close together, visually representing the concept of smallness for the item 'small'). Only one gesture per item was considered; when a child produced

both a deictic and representational gesture, the representational gesture was considered.

2.4.4. Spoken-gestural semantic relationship

The semantic relationship between gestures and co-occurring speech productions was coded for bimodal spoken/gestural answers as (Iverson et al., 2003; Stefanini et al., 2007): *complementary,* when the gesture identified the referent and the word labeled it (e.g., the child pointed to the target picture while saying 'umbrella'); *equivalent,* when the spoken and gestural elements were referred to the same referent and conveyed the same meaning (e.g., the child rubbed her/his fingers through his/her hair, as for combing his/her hair, while saying 'comb'); and *supplementary,* when the gesture and the word conveyed two different meanings so that each of the combined elements added information to the other one (e.g., the child turned her/his finger to her/his cheek while saying 'banana').

2.5. Reliability

The intercoder reliability of the PiNG lexical and gestural measures was independently assessed by two coders. Concerning noun and predicate comprehension accuracy measures, two independent coders (i.e., the first and third authors of this paper) rated children's answers at 38 over 46 sessions (83%) with intraclass correlation coefficients (ICC) being > 0.93. As for the measures regarding the PiNG production subtests, two independent coders (i.e., the first and second authors of this paper) coded 10 over 46 sessions (22%). The coders' agreement (data from noun and predicate subtests were aggregated) on the measure of answer accuracy corresponded to a Cohen's Kappa of 0.98. Concerning the modality of expression, type of gestures, and spoken-gestural semantic relationship, Cohen's Kappas corresponded respectively to 0.76, 0.89, 0.71. Overall, good intercoder reliability was achieved. In cases of disagreement, the item was coded by coders together leading to a coding consensus between them.

2.6. Statistical analyses

IBM Statistics 25 was used, and statistical significance was set at 0.05. Data distribution was checked for normality with Kolmogorov-Smirnov and Shapiro-Wilk tests. As all measures were not normally distributed ($p < 0.05$), non-parametric tests were used, and median (*Md*), interquartile range (*IQR*), and mean rank were reported for distributions.

A first set of Mann-Whitney tests compared groups' answer accuracy. A second set investigated groups' differences in the modality of expression, types of gestures, and spoken-gestural semantic relationship. The effect size of each comparison was calculated to investigate the strength of the differences between groups. Effect sizes (r) were computed based on the standardized test statistic z ($|z|$) of the Mann-Whitney U test and the number of pairs (n). An effect size $r < 0.30$ indicates a small effect, a r between 0.30 and 0.50 suggests a medium effect, and a $r > 0.50$ implies a large effect. This nuanced evaluation enhances the robustness and interpretability of our Mann-Whitney U Test results.

3. Results

3.1. Accuracy

Concerning the noun comprehension subtest (see Table 3), R/E-LTs provided significantly fewer correct answers and more incorrect answers and no responses than E-LTs, with a medium effect size (Table 3). Compared to the Italian normative data (Bello et al., 2010), 82% (9) R/E-LTs and 37% (13) E-LTs had a score $\leq 10^{th}$ percentile.

Concerning the predicate comprehension subtest (see Table 3), R/E-LTs provided significantly fewer correct answers, with a large effect size, and more no responses, with a medium effect size, than E-LTs (see Table 3). Compared to the Italian normative data (Bello

Table 3
Comparison Between Expressive and Receptive/Expressive Late Talkers' Accuracy at the PiNG Noun and Predicate Comprehension Subtests.

	E-LTs ($n = 35$)			R/E-LTs ($n = 11$)			Mann-Whitney U Test		
PiNG Measures	*Md*	*IQR*	*Mean rank*	*Md*	*IQR*	*Mean rank*	*U*	*p*	*r*
Noun Comprehension									
Correct	17	15-19	27.10	13	11-15	12.05	66.500	**.001**	.48
Incorrect	3	1-5	20.36	6	4-7	33.50	82.500	**.004**	.42
No response	0	0-0	21.00	1	0-2	31.45	105.000	**.003**	.45
Not administered	0	0-0	22.81	0	0-0	25.68	168.500	.205	.19
Predicate Comprehension									
Correct	14	11-16	27.23	8	5-12	11.64	62.000	**<.001**	.50
Incorrect	5	4-8	22.93	7	4-9	25.32	172.500	.604	.08
No response	0	0-0	21.21	1	0-2	30.77	112.500	**.011**	.37
Not administered	0	0-0	22.39	0	0-1	27.05	153.500	.086	.25

Note. The PiNG noun comprehension subtest includes 20 lexical targets depicting objects and tools, transport, animals, food and clothing. The PiNG predicate comprehension subtest includes 20 lexical targets depicting actions with objects, actions without objects, adjectives, and adverbs of place. *Md* = median; *IQR* = interquartile range. Significant results are in bold.

Table 4
Comparison Between Expressive and Receptive/Expressive Late Talkers' Accuracy at the PiNG Noun and Predicate Production Subtests.

	E-LTs (n = 35)			R/E-LTs (n = 11)			Mann-Whitney U Test		
PiNG Measures	*Md*	*IQR*	*Mean rank*	*Md*	*IQR*	*Mean rank*	*U*	*p*	*r*
Noun Production									
Correct	4	2-9	25.83	2	0-4	16.09	111.000	**.035**	.31
Incorrect Semantically Related	2	1-3	27.07	0	0-1	12.14	67.500	**<.001**	.49
Incorrect Semantically Unrelated	1	0-3	25.21	0	0-2	18.05	132.500	.105	.24
Unintelligible	2	1-6	24.07	1	0-5	21.68	172.500	.603	.08
No Response	4	2-6	24.34	2	1-6	20.82	163.000	.444	.11
Not Administered	0	0-2	20.97	12	0-15	31.55	104.000	**.014**	.36
Predicate Production									
Correct	1	0-5	26.70	0	0-0	13.32	80.500	**.002**	.46
Incorrect Semantically-Related	2	1-5	26.89	0	0-0	12.73	74.000	**.001**	.47
Incorrect Semantically-Unrelated	3	0-5	26.34	0	0-1	14.45	93.000	**.009**	.39
Unintelligible	1	0-3	25.99	0	0-1	15.59	105.500	**.018**	.35
No Response	2	1-4	26.91	0	0-0	12.64	73.000	**.002**	.47
Not Administered	1	0-15	19.73	20	15-20	35.50	60.500	**<.001**	.51

Note. The PiNG noun production subtest includes 20 lexical targets depicting objects and tools, transport, animals, food and clothing. The PiNG predicate production subtest includes 20 lexical targets depicting actions with objects, actions without objects, adjectives, and adverbs of place. *Md* = median; *IQR* =interquartile range. Significant results are in bold.

et al., 2010), 82% (9) R/E-LTs and 49% (17) E-LTs had a score $\leq 10^{th}$ percentile.

Concerning the noun production subtest (see Table 4), R/E-LTs provided significantly fewer correct and incorrect semantically related answers and a significantly higher number of not administered items, with a medium effect size, than E-LTs (see Table 4). Compared to the Italian normative data (Bello et al., 2010), 91% (10) R/E-LTs and 86% (30) E-LTs had a score $\leq 10^{th}$ percentile.

Concerning the predicate production subtest (see Table 4), R/E-LTs provided significantly fewer correct, incorrect semantically related, incorrect semantically unrelated, unintelligible, and no response answers, with a medium effect size, and a significantly higher number of not administered items, with a large effect size, than E-LTs (see Table 4). Compared to the Italian normative data (Bello

Table 5
Comparisons Between Expressive and Receptive/Expressive Late Talkers' Modality of Expression, Type of Gestures, and Spoken-Gestural Semantic Relationship at the PiNG Noun and Predicate Production Subtests.

	E-LTs (n = 35)			R/E-LTs (n = 11)			Mann-Whitney U Test		
PiNG Measures	*Md*	*IQR*	*Mean rank*	*Md*	*IQR*	*Mean rank*	*U*	*p*	*r*
Noun Production									
Modality of Expression									
Unimodal Spoken	4	1-7	26.49	0	0-2	14.00	88.000	**.006**	.40
Unimodal Gestural	4	1-6	24.43	3	1-5	20.55	160.000	.397	.12
Bimodal Spoken/Gestural	3	1-7	25.09	1	0-4	18.45	137.000	.148	.21
Type of Gestures									
Deictic Gestures	7	2-13	24.69	5	1-8	19.73	151.000	.284	.16
Representational Gestures	1	0-2	26.16	0	0-0	15.05	99.500	**.008**	.40
Spoken-Gestural									
Semantic Relationship									
Complementary	2	0-5	24.09	1	0-4	21.64	172.000	.587	.08
Equivalent	0	0-1	25.23	0	0-0	18.00	132.000	**.037**	.30
Supplementary	0	0-1	24.37	0	0-0	20.73	162.000	.321	.15
Predicate Production									
Modality of Expression									
Unimodal Spoken	4	0-9	26.30	0	0-0	14.59	94.500	**.008**	.39
Unimodal Gestural	3	1-4	26.57	0	0-1	13.73	85.000	**.005**	.42
Bimodal Spoken/Gestural	2	0-5	26.77	0	0-1	13.09	78.000	**.002**	.45
Type of Gestures									
Deictic Gestures	4	1-6	26.87	0	0-2	12.77	74.500	**.002**	.45
Representational Gestures	1	0-4	26.80	0	0-0	13.00	77.000	**.001**	.48
Spoken-Gestural									
Semantic Relationship									
Complementary	2	0-3	26.07	0	0-1	15.32	102.500	**.015**	.36
Equivalent	0	0-1	25.86	0	0-0	16.00	110.000	**.010**	.38
Supplementary	0	0-0	24.76	0	0-0	19.50	148.500	.086	.25

Note. Significant results in bold. *Md* = median; *IQR* = interquartile range.

et al., 2010), 91% (10) R/E-LTs and 94% (33) E-LTs and had a score $\leq 10^{th}$ percentile.

3.2. Modality of expression, type of gestures, and spoken-gestural semantic relationship

Concerning the noun production subtest, R/E-LTs produced significantly fewer unimodal spoken answers, representational gestures, and equivalent bimodal spoken/gestural answers, with a medium effect size, than E-LTs (Table 5). Representational gesture production was limited and primarily observed in E-LTs, thus only descriptive data were reported. The E-LTs children produced mainly "hand-as-object" gestures (*M*= 0.89, *SD*= 1.28; *Md*= 0, *IQR* = 0-1), followed by "body" (*M*= 0.43, *SD*= 0.78; *Md*= 0, *IQR* = 0-1), "hand-as-action" (*M*= 0.06, *SD*= 0.34; *Md*= 0, *IQR* = 0-0) and "size-and-shape" (*M*= 0.03, *SD*= 0.17; *Md*= 0, *IQR*= 0-0) gestures. R/E-LTs produced very few "hand-as-object" (*M*= 0.09, *SD*= 0.30; *Md*= 0, *IQR*= 0-0) gestures.

Concerning the predicate production subtest, R/E-LTs produced significantly fewer unimodal spoken, unimodal gestural, and bimodal spoken/gestural answers, deictic and representational gestures, and complementary and equivalent bimodal answers, with a medium effect size, than E-LTs (Table 5). Representational gesture production was limited and only observed in E-LTs. Specifically, these children produced predominantly "body" gestures (*M*= 1.20, *SD*= 1.59; *Md*= 1, *IQR*= 0-2), followed by "size-and-shape" (*M*= 0.46, *SD*= 0.70; *Md*= 0, *IQR*= 0-1), "hand-as-object" (*M*= 0.17, *SD*= 0.38; *Md*= 0, *IQR*= 0-1) and "hand-as-action" (*M*= 0.31, *SD*= 0.58; *Md*= 0, *IQR*= 0-0) gestures.

4. Discussion

Our findings brought new evidence about receptive and expressive lexical skills characterizing E-LTs and R/E-LTs with a particular focus on noun and predicate acquisition assessed with a structured task. Moreover, the analysis of gestures spontaneously produced during the naming task broadened our understanding of distinct communication profiles, providing valuable information on the nature of E-LTs and R/E-LTs' difficulties and linguistic processes.

4.1. Expressive and receptive/expressive late talkers' lexical skills

With respect to noun and predicate comprehension, most R/E-LTs and less than half E-LTs scored below the 10^{th} percentile. Discrepancies between the two groups were revealed by answer accuracy analysis showing that R/E-LTs provided fewer correct answers and more incorrect answers and no responses than E-LTs. These findings suggest that R/E-LTs faced widespread receptive lexical difficulties for nouns and predicates, impacting their overall language comprehension abilities to a greater extent. On the other side, even if, as expected, E-LTs seemed to be characterized by stronger receptive lexical skills than R/E-LTs, a notable part of them still struggled with acquiring lexical comprehension, with greater difficulties in predicates.

As regards noun and predicate production, our findings highlighted that the majority of both R/E-LTs and E-LTs scored below the 10^{th} percentile. However, issues in expressive skills were more severe in R/E-LTs than in E-LTs, as the former provided significantly fewer correct answers in noun and predicate production and had a higher number of not administered items, which might be due to difficulties in getting through the task. Differences between groups also emerged in the analysis of incorrect responses as E-LTs provided more attempts of responses, producing more incorrect semantically related answers than R/E-LTs. In addition, in predicate production, E-LTs also produced more incorrect semantically unrelated answers, unintelligible answers, and no responses. These differences shed light on distinct profiles of LTs. The higher number of correct and semantically related incorrect answers in E-LTs suggests that these children could represent word meanings, but they struggled with lexical retrieval of the target word and often activated a semantic strategy for trying to retrieve the meaning. By contrast, the higher number of not administered items on both noun and predicate production in R/E-LTs suggests potential difficulties in maintaining attention to the task and in representing and retrieving meanings.

The combined findings in comprehension and production subtests revealed the presence of a greater expressive delay in LTs also displaying difficulties in receptive language skills. These challenges might be related to broader language and cognitive processing issues involved in early language comprehension and production tasks, which could influence their performance. Indeed, using behavioural measures, such as the picture naming game, involves attention allocation, audio-visual processing, and lexical retrieval, all being affected by speed of information processing. Previous research (Fernald & Marchman, 2012; Macroy-Higgins & Montemarano, 2016) showed that LTs displayed difficulties with attention allocation and highlighted that differences among children in any of the processes mentioned above could lead to different language proficiency profiles. Thus, weaknesses in these areas could explain the difficulties highlighted in R/E-LTs in initiating and completing the tasks. The lower cognitive scores presented at the BSID-III cognitive scale in our study by R/E-LTs compared to E-LTs may support this hypothesis that was also supported by Desmarais et al. (2010), who showed that LTs characterized by higher scores in language comprehension and expression (i.e., those belonging to the third profile) had higher cognitive scores than more severely impaired LTs (i.e., those belonging to the first or second profile). In the same direction, Zuccarini et al. (2023) found that LTs with the *severe* and *moderate* profiles had significantly lower cognitive skills than those with the *mild* profile. Relationships between early receptive and cognitive processes in LTs and their cascading-effects on expressive skills deserve deepened investigation in future studies focusing on specific cognitive skills, such as attention allocation, speed of information processing, coordination of audio and visual information, and working memory, beside using global cognitive scores.

Our findings also underlined the importance of using specific tools, beyond more general developmental scales, to assess the lexicon and its components. At a general level, our findings showed that the results at the PiNG test were in line with those obtained at the

BSID-III language scales with R/E-LTs reporting significantly lower scores that E-LTs in both the BSID-III receptive and expressive scales and the PiNG comprehension and production subtests. However, the use of the PiNG test allowed a more deepened understanding of LTs' receptive and expressive skills in nouns and predicates. Indeed, it highlighted that predicate acquisition is generally more affected than noun acquisition in LTs at 30 months. These findings aligned with previous studies on typically developing children (Bello et al., 2010, 2012; D'Odorico & Fasolo, 2007), suggesting that LTs exhibit a lexical acquisition trend similar, even if delayed, to that observed in typically developing children with an earlier occurrence of noun than predicate acquisition. Indeed, concrete nouns are earlier and more easily acquired as they are associated with separate objects by primarily employing extra-linguistic (e.g., perceptual, motor, and socio-communicative) cues, whereas predicates refer to relational concepts and require the use of both linguistic (e.g., morphosyntactic) and extra-linguistic cues to be acquired (Gentner, 2006). Because the progressive changes in vocabulary composition are closely related to vocabulary size expansion, comprehension and production subtests on nouns appear more feasible than those on predicates for assessing lexical skills in LTs and, particularly, in R/E-LTs whose delay is more widespread.

4.2. Expressive and receptive/expressive late talkers' modality of expression and gesture use

Overall, our findings shed light on distinct patterns of modality of expression and gestural communication in LTs. R/E-LTs, compared to E-LTS, produced fewer unimodal spoken responses, representational gestures, and equivalent gesture-word combinations in noun production and lower frequencies of all modalities of expression, deictic and representational gestures, and complementary and equivalent bimodal responses in predicate production. These findings are coherent with R/E-LTs' more widely compromised lexical skills, discussed in the paragraph above, but also highlight that E-LTs use gestures and, particularly, representational gestures, as an efficient means of communicating meanings, whereas R/E-LTs scarcely used representational gestures for noun production and not at all for predicate production. Noteworthy, "hand-as-object" gestures were the most frequent representational strategy used by both groups for noun production, even though with a very limited frequency for R/E-LTs, whereas "body" gestures were the most frequent representational strategy used by E-LTs for predicate production. These findings add new evidence about LTs' language acquisition processes and use of representational gestures that, conveying meaning through a visuospatial representation of objects and actions, help express the intended words' meanings more accurately or replace words not yet produced (Iverson & Braddock, 2011; Iverson & Goldin-Meadow, 2005). It is worth noting that representational gestures, mirroring the spatial-dynamic characteristics of words, constitute a symbolic representation form that facilitates the effective communication of meanings. This is also confirmed by the finding that the "hand-as-object" representational strategy was more frequently activated for noun production, whereas the "body" representational strategy was more frequently activated for predicate production, a finding that is in line with the embodied cognition theory (Evans et al., 2001). In addition, the greater use of equivalent bimodal spoken-gestural answers by E-LTs suggests that they rely more on representational gestures to reinforce the information conveyed through the spoken modality. By doing so, they appear to be able to support the meaning expressed by spoken words with a physical motor representation. Furthermore, based on the idea that the semantic knowledge pertaining to objects and actions relies on the integration of multimodal abilities, i.e., lexical labeling, perceptual attributes (such as shape and function), and proprioceptive information derived from direct experiential encounters (Capone, 2007), in the context of our study, the greater frequency of representational gestures and higher occurrence of semantically related responses among E-LTs suggest that the challenges in this group were primarily related to the linkage of meaning with speech. By contrast, in the case of R/E-LTs, the difficulty might be more associated with the acquisition of semantic knowledge related to the target objects or actions that need to be labelled. These findings align with previous research demonstrating a positive association between the use of representational gestures and the accuracy of vocabulary comprehension (Cattani et al., 2019) and are coherent with the lower receptive skills characterizing R/E-LTs. Indeed, Caselli et al. (2012) argued that word comprehension constitutes a bridge between action/gesture production and word production, suggesting that the transition from action/gesture to word production is ontogenetically mediated by word comprehension.

4.3. Limitations and strengths of the study

The PiNG test allowed to obtain quantitatively and qualitatively detailed analysis concerning the linguistic and gestural skills of E-LTs and R/E-LTs. However, some limitations emerged. First, several cognitive skills are required, constituting potential confounding factors, especially for the predicate subtests. It is crucial for future research to integrate the lexical assessment with specific cognitive functions, such as attentional allocation, speed of information processing, coordination of audio and visual information, and working memory measures.

Another limitation that may impact the generalizability of our findings was constituted by the presence of a relatively small sample size, particularly given the uneven distribution of the two groups, and the composition of our sample, that however is representative of the distribution of E-LTS and R/E-LTs in the general population, as R/E-LTs are much fewer than E-LTs (Desmarais et al., 2010). Future research should expand the sample size, particularly focusing on R/E-LTs, and including also a control group of children with typical language development to enhance the description of LTs' linguistic and communicative skills and profiles.

Additionally, one more limitation was associated with the cross-sectional nature of the assessment. The present study, analysing the lexical and gestural measures at the same age, did not allow us to investigate how the different linguistic and communication profiles evolved over time and their relationships with later linguistic outcomes. Longitudinal assessments of the profiles of E-LTs and R/E-LTs could be useful to better understand developmental pathways.

Lastly, a further limitation concerns not having investigated the impact of childcare attendance and home language environment on the type of language delay and communicative-linguistic performances of LTs. Previous studies recognized the crucial role of non-

parental childcare quality during the first three years of age in influencing cognitive and language development (Luijk et al., 2015; NICHD ECCRN, 2000), however lacking a specific exploration of gestural skills. Additionally, the differences we found in childcare attendance between E-LTs and R/E-LTs underscore the need for future research to delve into the investigation of how children are stimulated at home and in childcare settings beyond clinical assessments. Thus, future research should explore the nuanced influence of these environmental factors on LTs' linguistic and communicative abilities.

5. Conclusions and clinical implications

The present study highlighted the relevance of assessing receptive skills, beyond expressive skills, to understand the linguistic and communicative profiles of E-LTs and R/E-LTs. Lexical comprehension plays a central role in child development as it serves as the bedrock for developing expressive language skills and maintains a close link to cognitive development (Dale et al., 2003; Thal et al., 1991). Given the potential implications of comprehension deficits in shaping language development in preschool and school-aged children, characterizing their early comprehension skills becomes fundamental in understanding their language profiles. An encouraging avenue for future research would be to explore this clinical population considering a more comprehensive range of comprehension skills, including morphosyntactic and phonological aspects. However, it is essential to acknowledge the considerable challenges in measuring language comprehension skills in very young LTs. Moreover, our findings revealed the importance of delving into a detailed analysis of the specific types of errors produced by LTs in assessing their expressive skills, which could significantly contribute to a deeper understanding of the strategies that characterize language processing across different LTs profiles.

Additionally, our findings confirmed the usefulness of a standardized structured task, besides parental questionnaires and developmental scales, in eliciting and analyzing the production of spontaneous gestures besides verbal responses. Notably, research showed that, during early childhood, both typically and atypically developing children employ deictic and representational gestures in association with or in substitution of verbal productions when prompted to name a target image (Sansavini et al., 2015; Stefanini et al., 2009). In this regard, given the increasing importance of assessing language skills and the underlying communicative skills in explaining LTs' variability and their subsequent language development, results suggest that the administration of a task usually implemented for a standardized language assessment play a valuable role in the clinical assessment of LTs. At the same time, our findings suggest the need to create a comfortable setting for LTs' assessment. Children, particularly those with language difficulties, such as LTs, may struggle to express their competencies in unfamiliar clinical settings, potentially leading to frustration, as evidenced by a notable number of not administered items. This approach could help researchers and clinicians to identify distinct communicative profiles among LTs and those at a higher risk for developmental language disorder, gain insights into the nature of LTs' linguistic and cognitive processes, and plan early multimodal personalised interventions integrating gestures and words in ecological parent-child interactional contexts. Early treatment, by fostering timely interventions and tailored strategies, not only offers targeted support to LTs' specific needs and competencies but could also play a pivotal role in enhancing long-term developmental outcomes (Rinaldi et al., 2021). In this regard, a crucial role is played by the link between healthcare and early intervention services to promote children's well-being and maximise their potential starting from the early stages of life as indicated by the WHO Nurturing Care recommendations (WHO, 2018). In addition, considering the impact of language delay on children's subsequent learning and social development, findings suggest the need of supporting the use of direct assessment tasks by designing tools also for caregivers and other profssional figures (e.g., paediatricians, educators), allowing them daily insights into specific LTs' difficulties that allow to observe early markers to detect children at higher risk for developmental language disorder. Furthermore, by fostering and promoting gestures' use, educators and parents can create an enriched environment supporting children in developing strategies to communicate during crucial stages of language acquisition.

Funding

This work was supported by grants from (a) AlmaIdea 2017 Grant Senior, University of Bologna, "Ritardo di linguaggio nei bambini nati pretermine: screening, valutazione e intervento" ("Language delay in preterm children: Screening, assessment and intervention") awarded to A.S. as main PI, and (b) Italian Ministry of Education, University and Research MIUR PRIN 2017 (2017HRCPE4_004) "Early markers of Language-Learning Impairment" awarded to A.S., as PI of the Unit of Bologna.

CRediT authorship contribution statement

Caterina Verganti: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Chiara Suttora:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis, Data curation, Conceptualization. **Mariagrazia Zuccarini:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Arianna Aceti:** Writing – review & editing, Methodology, Data curation. **Luigi Corvaglia:** Writing – review & editing, Methodology, Data curation. **Arianna Bello:** Writing – review & editing. **Maria Cristina Caselli:** Writing – review & editing. **Annalisa Guarini:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Formal analysis, Conceptualization. **Alessandra Sansavini:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Data availability

The authors do not have permission to share data.

Acknowledgments

We are grateful to the families and infants who participated in this study. We are also grateful to the health professionals of the Neonatology and Neonatal Intensive Care Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna —and, in particular, to Giacomo Faldella, for his support in funding acquisition and supervision of medical aspects of the methodology, Rosina Alessandroni, for medical care of the preterm sample, and Silvia Savini for her support in contacting the parents of preterm children.

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. https://doi.org/10.1016/j.jpeds.2010.10.032

Bello, A., Caselli, M.C., Pettenati, P., Stefanini, S.; 2010. PiNG. Parole in gioco: Una prova per la comprensione e produzione lessicale per la prima infanzia. Giunti, O. S.

Bello, A., Giannantoni, P., Pettenati, P., Stefanini, S., & Caselli, M. C. (2012). Assessing lexicon: Validation and developmental data of the Picture Naming Game (PiNG), a new picture naming task for toddlers: Assessing lexicon. *International Journal of Language & Communication Disorders, 47*(5), 589–602. https://doi.org/10.1111/j.1460-6984.2012.00168.x

Bello, A., Onofrio, D., Remi, L., & Caselli, C. (2018). Prediction and persistence of late talking: A study of Italian toddlers at 29 and 34 months. *Research in Developmental Disabilities, 75*, 40–48. https://doi.org/10.1016/j.ridd.2018.02.006

Benitz, W. E., Committee on Fetus and Newborn, & American Academy of Pediatrics. (2015). Hospital stay for healthy term newborn infants. *Pediatrics, 135*(5), 948–953. https://doi.org/10.1542/peds.2015-0699

BURERT - Bulletin of Emilia Romagna Region; 2019. Linee di indirizzo - Revisione sistematica delle prove Continuità delle cure. Dimissione della mamma e del bambino dal punto nascita. Commissione consultiva tecnico scientifica sul percorso nascita. ⟨https://bur.regione.emilia-romagna.it/bur/area-bollettini/bollettini-in-lavorazione/n-376-del-20-11-2019-parte-seconda.2019-11-20.0598955891/percorso-nascita-indicazioni-alle-aziende-sanitarie-per-garantire-la-sicurezza-del-parto-con-taglio-cesareo-e-la-continuita-delle-cure-tra-ospedale-e-territorio-a-madre-e-bambino/doc-tec-dimissione-madre-bambi.2019-11-20.1574240085⟩.

Capone, N. C. (2007). Tapping Toddlers' evolving semantic representation via gesture. *Journal of Speech, Language, and Hearing Research, 50*(3), 732–745. https://doi.org/10.1044/1092-4388(2007/051)

Carson, D. K., Klee, T., Perry, C. K., Muskina, G., & Donaghy, T. (1998). Comparisons of children with delayed and normal language at 24 months of age on measures of behavioral difficulties, social and cognitive development. *Infant Mental Health Journal, 19*(1), 59–75. https://doi.org/10.1002/(SICI)1097-0355(199821)19:1<59::AID-IMHJ4>3.0.CO;2-V

Caselli, M.C., Bello, A., Rinald, P., Stefanini, S., &Pasqualetti, P.; 2015. Il Primo Vocabolario del Bambino: Gesti, Parole e Frasi. Forme Lunghe e Forme Brevi del Questionario e Valori di Riferimento per la Fascia 8–36 Mesi. Franco Angeli.

Cattani, A., Floccia, C., Kidd, E., Pettenati, P., Onofrio, D., & Volterra, V. (2019). Gestures and words in naming: Evidence from crosslinguistic and crosscultural comparison. *Language Learning, 69*(3), 709–746. https://doi.org/10.1111/lang.12346

Chauhan, S. P., Scardo, J. A., Hayes, E., Abuhamad, A. Z., & Berghella, V. (2010). Twins: Prevalence, problems, and preterm births. *American Journal of Obstetrics and Gynecology, 203*(4), 305–315. https://doi.org/10.1016/j.ajog.2010.04.031

Chilosi, A. M., Pfanner, L., Pecini, C., Salvadorini, R., Casalini, C., Brizzolara, D., & Cipriani, P. (2019). Which linguistic measures distinguish transient from persistent language problems in Late Talkers from 2 to 4 years? A study on Italian speaking children. *Research in Developmental Disabilities, 89*, 59–68. https://doi.org/10.1016/j.ridd.2019.03.005

Collisson, B. A., Graham, S. A., Preston, J. L., Rose, M. S., McDonald, S., & Tough, S. (2016). Risk and protective factors for late talking: An EpiDemiologic Investigation. *The Journal of Pediatrics, 172*, 168–174.e1. https://doi.org/10.1016/j.jpeds.2016.02.020

Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay. *Journal of Speech, Language, and Hearing Research, 46*(3), 544. https://doi.org/10.1044/1092-4388(2003/044)

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language & Communication Disorders, 43*(4), 361–389. https://doi.org/10.1080/13682820701546854

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2010). Three profiles of language abilities in toddlers with an expressive vocabulary delay: Variations on a theme. *Journal of Speech, Language, and Hearing Research, 53*(3), 699–709. https://doi.org/10.1044/1092-4388(2009/07-0245)

D'Odorico, L., & Fasolo, M. (2007). Nouns and verbs in the vocabulary acquisition of Italian children. *Journal of Child Language, 34*(4), 891–907. https://doi.org/10.1017/S0305000907008240

Evans, J. L., Alibali, M. W., & McNeil, N. M. (2001). Divergence of verbal expression and embodied knowledge: Evidence from speech and gesture in children with specific language impairment, 2(3) *Language and Cognitive Processes, 16*, 309–331. https://doi.org/10.1080/01690960042000049.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-Talking toddlers. *Child Development, 83*(1), 203–222. https://doi.org/10.1111/j.1467-8624.2011.01692.x

Ferr,R., Orsini, A., Rea, M., Stoppa, E., & Mascellani, F.; 2015. Bayley Scales of Infant and Toddler Development, 3rd ed. Giunti, O.S.

Fisher, E. L. (2017). A systematic review and meta-analysis of predictors of expressive-language outcomes among late Talkers. *Journal of Speech, Language, and Hearing Research, 60*(10), 2935–2948. https://doi.org/10.1044/2017_JSLHR-L-16-0310

Guarini, A., Marini, A., Savini, S., Alessandroni, R., Faldella, G., & Sansavini, A. (2016). Linguistic features in children born very preterm at preschool age. *Developmental Medicine and Child Neurology, 58*(9), 949–956. https://doi.org/10.1111/dmcn.13118

Hawa, V. V., & Spanoudis, G. (2014). Toddlers with delayed expressive language: An overview of the characteristics, risk factors and language outcomes. *Research in Developmental Disabilities, 35*(2), 400–407. https://doi.org/10.1016/j.ridd.2013.10.027

Iverson, J. M., & Braddock, B. A. (2011). Gesture and motor skill in relation to language in children with language impairment. *Journal of Speech, Language, and Hearing Research, 54*(1), 72–86. https://doi.org/10.1044/1092-4388(2010/08-0197)

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science, 16*(5), 367–371. https://doi.org/10.1111/j.0956-7976.2005.01542.x

Iverson, J. M., Longobardi, E., & Caselli, M. C. (2003). Relationship between gestures and words in children with Down's syndrome and typically developing children in the early stages of communicative development. *International Journal of Language & Communication Disorders, 38*(2), 179–197. https://doi.org/10.1080/1368282031000062891

Kendon, A. (2004). *Gesture Visible Action as Utterance* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511807572

Korpilahti, P., Kaljonen, A., & Jansson-Verkasalo, E. (2016). Population-based screening for language delay: Let's talk STEPS study. *Psychology, 07*(02), 205–214. https://doi.org/10.4236/psych.2016.72023

Lobo, M. A., & Galloway, J. C. (2013). The onset of reaching significantly impacts how infants explore both objects and their bodies. *Infant Behavior and Development, 36*(1), 14–24. https://doi.org/10.1016/j.infbeh.2012.09.003

Lüke, C., Ritterfeld, U., Grimminger, A., Liszkowski, U., & Rohlfing, K. J. (2017). Development of pointing gestures in children with typical and delayed language Acquisition. *Journal of Speech, Language, and Hearing Research, 60*(11), 3185–3197. https://doi.org/10.1044/2017_JSLHR-L-16-0129

Luijk, M. P., Linting, M., Henrichs, J., Herba, C. M., Verhage, M. L., Schenk, J. J., Arends, L. R., Raat, H., Jaddoe, V. W., Hofman, A., Verhulst, F. C., Tiemeier, H., & van IJzendoorn, M. H. (2015). Hours in non-parental child care are related to language development in a longitudinal cohort study. *Child: Care, Health and Development, 41*(6), 1188–1198. https://doi.org/10.1111/cch.12238

Macroy-Higgins, M., & Montemarano, E. A. (2016). Attention and word learning in toddlers who are late talkers. *Journal of Child Language, 43*(5), 1020–1037. https://doi.org/10.1017/S0305000915000379

Manwaring, S. S., Swineford, L., Mead, D. L., Yeh, C.-C., Zhang, Y., & Thurm, A. (2019). The gesture–language association over time in toddlers with and without language delays, 239694151984554 *Autism & Developmental Language Impairments, 4*. https://doi.org/10.1177/2396941519845545.

Marentette, P., Pettenati, P., Bello, A., & Volterra, V. (2016). Gesture and symbolic representation in italian and english-speaking Canadian 2-Year-Olds. *Child Development, 87*(3), 944–961. https://doi.org/10.1111/cdev.12523

Nguyen, T. N., Spencer-Smith, M., Haebich, K. M., Burnett, A., Scratch, S. E., Cheong, J. L. Y., Doyle, L. W., Wiley, J. F., & Anderson, P. J. (2018). Language trajectories of children born very preterm and full term from early to late childhood. *The Journal of Pediatrics, 202*, 86–91. https://doi.org/10.1016/j.jpeds.2018.06.036

NICHD ECCRN. (2000). The relation of child care to cognitive and language development. *Child Development, 71*, 960–980. https://doi.org/10.1111/1467-8624.00202

O'Neill, H., & Chiat, S. (2015). What our hands say: Exploring gesture use in subgroups of children with language delay. *Journal of Speech, Language, and Hearing Research, 58*(4), 1319–1325. https://doi.org/10.1044/2015_JSLHR-L-14-0187

O'Neill, H., Murphy, C.-A., & Chiat, S. (2019). What our hands tell us: A two-year follow-up investigating outcomes in subgroups of children with language delay. *Journal of Speech, Language, and Hearing Research, 62*(2), 356–366. https://doi.org/10.1044/2018_JSLHR-L-17-0261

Pérez-Pereira, M. (2021). Prevalence of language delay among healthy preterm children, language outcomes and predictive factors. *Children, 8*(4), 282. https://doi.org/10.3390/children8040282

Phillips, R. M., Goldstein, M., Hougland, K., Nandyal, R., Pizzica, A., Santa-Donato, A., Staebler, S., Stark, A. R., Treiger, T. M., & Yost, E. (2013). National perinatal association. Multidisciplinary guidelines for the care of late preterm infants. *Journal of Perinatology, 33*(Suppl 2), S5–S22. https://doi.org/10.1038/jp.2013.53

Reilly, S., Wake, M., Bavin, E. L., Prior, M., Williams, J., Bretherton, L., Eadie, P., Barrett, Y., & Ukoumunne, O. C. (2007). Predicting language at 2 years of age: A Prospective community study. *Pediatrics, 120*(6), 441–449. https://doi.org/10.1542/peds.2007-0045

Rinaldi, P., Bello, A., Lasorsa, F. R., & Caselli, M. C. (2022). Do spoken vocabulary and gestural production distinguish children with transient language delay from children who will show developmental language disorder? A pilot study. *International Journal of Environmental Research and Public Health, 19*(7), 3822. https://doi.org/10.3390/ijerph19073822

Rinaldi, S., Caselli, M. C., Cofelice, V., D'Amico, S., De Cagno, A. G., Della Corte, G., Di Martino, M. V., Di Costanzo, B., Levorato, M. C., Penge, R., Rossetto, T., Sansavini, A., Vecchi, S., & Zoccolotti, P. (2021). Efficacy of the treatment of developmental language disorder: A systematic review. *Brain Sciences, 11*(3), 407. https://doi.org/10.3390/brainsci11030407

Rinaldi, P., Pasqualetti, P., Stefanini, S., Bello, A., & Caselli, M. C. (2019). The Italian words and sentences MB-CDI: Normative data and concordance between complete and short forms. *Journal of Child Language, 46*(3), 546–566. https://doi.org/10.1017/S0305000919000011

Sansavini, A., Bello, A., Guarini, A., Savini, S., Alessandroni, R., Faldella, G., & Caselli, C. (2015). Noun and predicate comprehension/production and gestures in extremely preterm children at two years of age: Are they delayed? *Journal of Communication Disorders, 58*, 126–142. https://doi.org/10.1016/j.jcomdis.2015.06.010

Sansavini, A., Bello, A., Guarini, A., Savini, S., Stefanini, S., & Caselli, M. C. (2010). Early development of gestures, object-related-actions, word comprehension and word production, and their relationships in Italian infants: A longitudinal study. *Gesture, 10*(1), 52–85. https://doi.org/10.1075/gest.10.1.04san

Sansavini, A., Favilla, M. E., Guasti, M. T., Marini, A., Millepiedi, S., Di Martino, M. V., , … Vecchi, S., Battajon, N., Bertolo, L., Capirci, O., Carretti, B., Colatei, M. P., Frioni, C., Marotta, L., Massa, S., Michelazzo, L., Pecini, C., Piazzalunga, S., Pieretti, M., & Lorusso, M. L. (2021). Developmental language disorder: Early predictors, age for the diagnosis, and diagnostic tools. a scoping review. *Brain Sciences, 11*(5), 654. https://doi.org/10.3390/brainsci11050654

Sansavini, A., Guarini, A., & Caselli, M. C. (2011). Preterm birth: neuropsychological profiles and atypical developmental pathways. *Developmental Disabilities Research Reviews, 17*(2), 102–113. https://doi.org/10.1002/ddrr.1105

Sansavini, A., Guarini, A., Savini, S., Broccoli, S., Justice, L., Alessandroni, R., & Faldella, G. (2011). Longitudinal trajectories of gestural and linguistic abilities in very preterm infants in the second year of life. *Neuropsychologia, 49*(13), 3677–3688. https://doi.org/10.1016/j.neuropsychologia.2011.09.023

Sansavini, A., Guarini, A., Zuccarini, M., Lee, J. Z., Faldella, G., & Iverson, J. M. (2019). Low rates of pointing in 18-month-olds at risk for autism spectrum disorder and extremely preterm infants: A common index of language delay? *Frontiers in Psychology, 10*, 2131. https://doi.org/10.3389/fpsyg.2019.02131

Sansavini, A., Pentimonti, J., Justice, L. M., Guarini, A., Savini, S., Alessandroni, R., & Faldella, G. (2014). Language, motor and cognitive development of extremely preterm children: Modeling individual growth trajectories over the first three years of life. *Journal of Communication Disorders, 49*, 55–68. https://doi.org/10.1016/j.jcomdis.2014.02.005

Sansavini, A., Zuccarini, M., Gibertoni, D., Bello, A., Caselli, M. C., Corvaglia, L., & Guarini, A. (2021). Language profiles and their relation to cognitive and motor skills at 30 months of age: An online investigation of low-risk preterm and full-term children. *Journal of Speech, Language, and Hearing Research, 64*(7), 2715–2733. https://doi.org/10.1044/2021_JSLHR-20-00636

Stefanini, S., Bello, A., Caselli, M. C., Iverson, J. M., & Volterra, V. (2009). Co-speech gestures in a naming task: Developmental data. *Language and Cognitive Processes, 24*(2), 168–189. https://doi.org/10.1080/01690960802187755

Stefanini, S., Caselli, M. C., & Volterra, V. (2007). Spoken and gestural production in a naming task by young children with Down syndrome. *Brain and Language, 101*(3), 208–221. https://doi.org/10.1016/j.bandl.2007.01.005

Suttora, C., Guarini, A., Zuccarini, M., Aceti, A., Corvaglia, L., & Sansavini, A. (2020). Speech and language skills of low-risk preterm and full-term late Talkers: The role of child factors and parent input. *International Journal of Environmental Research and Public Health, 17*(20), 7684. https://doi.org/10.3390/ijerph17207684

Suttora, C., Guarini, A., Zuccarini, M., Aceti, A., Corvaglia, L., & Sansavini, A. (2022). Integrating gestures and words to communicate in full-term and low-risk preterm late Talkers. *International Journal of Environmental Research and Public Health, 19*(7), 3918. https://doi.org/10.3390/ijerph19073918

Suttora, C., Zuccarini, M., Aceti, A., Corvaglia, L., Guarini, A., & Sansavini, A. (2021). The effects of a parent-implemented language intervention on late-Talkers' expressive skills: The Mediational role of parental speech contingency and dialogic reading abilities. *Frontiers in Psychology, 12*, Article 723366. https://doi.org/10.3389/fpsyg.2021.723366

Thal, D. J., & Tobias, S. (1992). Communicative gestures in children with delayed onset of oral expressive vocabulary. *Journal of Speech, Language, and Hearing Research, 35*(6), 1281–1289. https://doi.org/10.1044/jshr.3506.1289

Thal, D., Marchman, V. A., & Tomblin, J. B. (2013). Late-talking toddlers: Characterization and prediction of continued delay. *In Late talkers. Language, development, interventions and outcome* (pp. 169–201). Baltimore MD: Paul H. Brookes Publishing Co.

Thal, D., Tobias, S., & Morrison, D. (1991). Language and gesture in Late Talkers: A 1 year follow-up. *Journal of Speech, Language, and Hearing Research, 34*(3), 604. https://doi.org/10.1044/jshr.3403.604

World Health Organization, United Nations Children's Fund, W.B.G. 2018. Nurturing care for early childhood development: a framework for helping children survive and thrive to transform health and human potential. Report.

Woodward, L. J., Moor, S., Hood, K. M., Champion, P. R., Foster-Cohen, S., Inder, T. E., & Austin, N. C. (2009). Very preterm children show impairments across multiple neurodevelopmental domains by age 4 years. F339-44 *Archives of Disease in Childhood - Fetal and Neonatal Edition, 94*(5). https://doi.org/10.1136/adc.2008.146282.

Zambrana, I. M., Pons, F., Eadie, P., & Ystrom, E. (2014). Trajectories of language delay from age 3 to 5: Persistence, recovery and late onset. *International Journal of Language & Communication Disorders, 49*(3), 304–316. https://doi.org/10.1111/1460-6984.12073

Zambrana, I. M., Vollrath, M. E., Jacobsson, B., Sengpiel, V., & Ystrom, E. (2021). Preterm birth and risk for language delays before school entry: A sibling-control study. *Development and Psychopathology, 33*(1), 47–52. https://doi.org/10.1017/S0954579419001536

Zuccarini, M., Guarini, A., Gibertoni, D., Suttora, C., Aceti, A., Corvaglia, L., Bello, A., Caselli, M. C., & Sansavini, A. (2023). Describing communication profiles of low-risk preterm and full-term late talkers. *Journal of Communication Disorders, 104*, Article 106336. https://doi.org/10.1016/j.jcomdis.2023.106336