

The validity of the Language Environment Analysis system in two neonatal intensive care units

Eva Ståhlberg-Forsén^a, MA, Anette Aija^{bd}, MD, Birgit Kaasik^b, MA, Reija Latva^c, MD, PhD, Sari Ahlqvist-Björkroth^d, PhD, Liis Toome^b MD, PhD, Liisa Lehtonen^{de}, MD, PhD, Suvi Stolt^a, PhD

Affiliations: ^aUniversity of Helsinki, Helsinki, Finland, ^bTallinn Children's Hospital, Tallinn, Estonia, ^cTampere University Hospital, Tampere, Finland ^dUniversity of Turku, Turku, Finland and ^eTurku University Hospital, Turku, Finland.

Short running title: The validity of the LENA system in the NICU

Corresponding author: Eva Ståhlberg-Forsén, Department of Psychology and Logopedics, unit of Logopedics, Haartmaninkatu 3, PO Box 21, 00014 University of Helsinki, Finland, eva.stahlberg-forsen@helsinki.fi, +358-40-7347 964

Accepted manuscript for publication in Acta Paediatrica

DOI: 10.1111/apa.1580

ABSTRACT

Aim: To evaluate the validity of the Language Environment Analysis (LENA) system's automatic measures in two neonatal intensive care units supporting parent-infant closeness, and in two Finno-Ugric languages: Finnish and Estonian. **Methods:** The sound environment of 70 very preterm infants was recorded for 16 hours in the neonatal intensive care units with the LENA system roughly at the gestational age of 32 (+2) weeks. Of these, the recordings of 14 infants (20%, two five-minute samples with a high percentage of speech, totally 140 minutes) were analyzed in detail and in two different ways. Parental closeness diaries were used to document the presence of the parents. Agreements between LENA system and human coder estimates were analyzed. **Results:** Findings showed a high variation in agreements. The highest agreements were found in female and adult word counts ($r = .91$ and $.95$). The agreements for child vocalization count, conversational turns, and silence were modest or low ($r = -.03$ — 0.64). **Conclusion:** Our study provides novel information on the validity of the LENA system in the neonatal intensive care unit. Findings show that the LENA system provides valid information on adult words, but LENA estimates for child vocalizations were less valid at this early age.

Key words: language acquisition, Language Environment Analysis, neonatal intensive care unit, parental speech, preterm infant.

KEY NOTES

- The validity of the Language Environment Analysis system (LENA) has not previously been studied in the neonatal intensive care environment.
- When the values of the automated LENA system and human coders were compared, the results showed that the LENA system provides valid information regarding female and adult word counts.
- LENA estimates for early infant vocalizations were shown to be less valid.

Preterm children have an elevated risk for weak language skills (1, 2). The impact of preterm birth is complex, and the language development of preterm children can be influenced by multiple factors, including brain injury and environment. The developing brain of the preterm newborn is vulnerable, but has a plasticity that enables the infant to compensate for and benefit from environmental factors (3).

During a typical full term pregnancy, the fetus experiences filtered sounds of low sound level and frequency, including prosodic characteristics of speech (4, 5). The mother's voice is a prominent part of the sound environment *in utero* and the fetus can detect and respond to the maternal voice from about 24 weeks of gestation (4). Fetuses studied at 36 weeks of gestation responded to the mothers' voices and to prosodic changes in the speech, indicating that learning of voices and prosodic features starts prenatally (6). Event-related potentials measured in full term newborns have shown that differences in responses to syllables can be identified 1—7 days after birth and may predict later language development (7).

The sound environment in the neonatal intensive care unit (NICU) differs from the intrauterine environment and consists of human voices, silence, and repetitious or short-duration sounds, including sounds from medical equipment (5). The infant is not continuously exposed to sounds of maternal cardiac and digestive functions and the acoustic features of the mother's voice are different than *in utero* (4). Recommended standards for NICU contain instructions for acoustics, including aims to reduce harmful noise and provide speech privacy for families (8). Growing evidence show that parent-infant closeness in the NICU, and an environment that promotes this, is beneficial to the parent-infant relationship and for the development of the preterm infant (9). Parental speech at moderate sound levels, combined with skin-to-skin care, is considered favorable for the preterm infant's language development (5). The amount of parental talk at 32 weeks' gestational age has been associated with the number of vocalizations of preterm infants at the same age (10) and language skills at 7 and 18 months' corrected age (11). Further knowledge is needed to understand the

language environment and the optimal acoustic conditions in the NICU to support language development (5).

The Environment Analysis (LENA) system (LENA Research Foundation) is a tool for recording and analyzing language environments, and has been used to investigate preterm infants' early vocalizations (10, 12) and language environments (10, 11, 13). However, there is a need to obtain validity information on this measure in the NICU context, since it is currently lacking. Furthermore, the environment may also differ in different NICUs due to, e.g., parental presence (9, 14), which may influence the amount of parental talk in the NICU unit. This study provides information from NICUs that support parent-infant closeness (14).

The LENA system was developed for the American-English language context and validated from age 2 months to 48 months (15-17). Its validity has been studied in some non-English languages, including Chinese (18), European French (19) and Vietnamese (20), but validity information on other languages is still needed (21). Finnish and Estonian are two separate languages from the Finno-Ugric group of languages. The languages have linguistic features in common, such as rich inflectional morphology. Words are relatively long, since suffixes are added to the word stem to express grammatical functions. The primary word stress usually falls on the first syllable of the word (22). Further, the fundamental speech frequency of Finnish female speakers may differ from that of English female speakers. Female Finnish university students tend to use a lower fundamental speech frequency, compared with values reported in international literature (23). The LENA system in the Finnish language setting has been studied only with 6—12 month old children (24) and it has not previously been evaluated in the Estonian language context.

The objective of the present study was to evaluate the validity of the LENA system in the NICU environments in two countries. The research questions were: 1. How accurately does the LENA system identify segments of female and male adult,

segments of key child (the infant with the recording processor) and silence in the Finnish and Estonian NICU environments? 2. How valid is the information provided by the LENA system in the settings regarding female word count, male word count, adult word count, child vocalization count, conversational turns and duration of silence?

METHODS

Participants

The participants were very preterm born (<32 gestational weeks) infants participating in an ongoing longitudinal research project in the NICUs in Turku University Hospital, Finland, and in Tallinn Children's Hospital, Estonia. The recruitments started in March—April 2017. Parents of the infants were contacted in the NICU when the infant's medical condition was stable. Infants with life-threatening conditions and considerable congenital anomalies or syndromes were excluded.

At the time of the present study, 29 infants from Turku and 41 infants from Tallinn, with Finnish or Estonian as their primary language, were recruited to the project. From this sample, 7 infants from both units (a total of 14, 20% of recruited participants) were randomly selected for the validation study. Twins were excluded to verify that only vocalizations from the key child were counted. Please see Table 1, for background characteristics of the participants in the present study.

The Ethics Committees of the Hospital District of Southwest Finland and the University of Tartu have approved the study protocol. The participating families received verbal and written information about the study and gave their signed informed consent.

(INSERT TABLE 1)

Analysis

The sound environment of each participant was recorded for 16 hours continuously in the NICU with the LENA system roughly at the gestational age of 32 (+2) weeks. The

LENA system consists of a digital processor for recording and computer software for segmentation and analysis. The processor was kept as near the infant's head as possible (roughly 10 cm in the bed, roughly 30 cm during Kangaroo care) over the entire recording time. A parental closeness diary (14) was maintained to document the parents' presence.

From the information derived from the LENA system and the parental closeness diaries, two five-minute chunks of the highest 10% of production of adult speech from each recording (total number of minutes: 140 minutes), when a parent was present, were selected for analysis. The samples were analyzed in detail, in two different ways, based on definitions by the LENA Foundation (25). The validation procedure consisted of the following parts: In part A, the human coder listened to the samples and checked, based on the human ear, if the LENA-provided labels for female adult, male adult, key child and silence were correct or not. In part B, the following variables were analyzed: female words, male words, adult words, child vocalizations and conversational turns. In addition, the validity of the silence estimate was investigated.

The LENA system labels segments and estimates counts directly. The same definitions as the LENA system uses in the automatic analysis (15-17, 25, 26) were used. Female, male and adult word counts consisted of words spoken in the environment of the infant. Unclear or overlapping speech was excluded. Child vocalizations were counted when the vocalization was surrounded by more than 300 milliseconds of silence or sounds that were not the infant's vocalization. Cries or vegetative sounds were not counted as child vocalizations. Conversational turns were counted when an infant vocalized as a response to adult speech or when an adult responded to the infant within 5 seconds (26, 27). As LENA software does not differentiate child-directed responses from overheard speech (25), both were counted as parts of conversational turns. The LENA system categorizes speech labels into near and far classes (17). In this study, these labels were combined. Silence is defined by the LENA system as a segment 800 ms or longer with scant or no acoustic

information, or with an acoustical energy of 32 dB or less (25). All segments of no sound or with very faint background sound measured as 1 s or longer were included in the human estimate.

The human coder analyses were conducted with the transcriber software (28). In part A, the human coder listened to the segment derived from the LENA system and coded the label provided by the LENA system as correct or false, based on the human ear. In part B, the recorded speech was first transcribed. Then, the words, vocalizations and conversational turns were manually counted, and values were compared with LENA counts. Human coder estimates of silence were manually measured in seconds with a digital stopwatch. All segments of silence measured as 1 second or longer were included in the estimate.

One coder in Turku and one coder in Tallinn acted as principal coders. The principal coders analyzed most of the samples (Turku: 93%, Tallinn: 86%) and the rest of the samples were analyzed by independent coders. The coders were trained in the coding principles and the consensus of the principles was agreed upon. To assess interrater reliability, 29% of the data (4 samples) was double-scored, and Krippendorff's alpha (α) values were calculated between the scorings (29). In part A, the interrater reliability was as follows: female adult $\alpha = .72$ (confidence interval .21;.1.0), male adult $\alpha = .15$ (-0.69;.95), key child $\alpha = .98$ (.93;1.0), and silence $\alpha = .29$ (-.37;.65). In part B, the values were: female word count $\alpha = .92$ (.92;.92), male word count $\alpha = .94$ (.92;.98), adult word count $\alpha = .78$ (.56;.92), child vocalization count $\alpha = .77$ (.41;.98), conversational turns $\alpha = .81$ (.68;.95), and silence $\alpha = .94$ (.92;.98).

In part A, the agreement between LENA and human coders was reported by calculating agreement percentages. The agreement percentages were first calculated from the total number of LENA labels for each variable. Secondly, mean values for each variable were calculated. In part B, Spearman correlations were calculated to describe the associations between LENA and human provided estimates. Kalpha values were calculated to measure the agreement between LENA and human values

in both part A and B. Statistical analyses were conducted with IBM SPSS Statistics for Windows, Version 25.0 (Armonk, NY: IBM Corp.)

RESULTS

Regarding the results of part A, the total number of female LENA labels noted in the sample analyzed was 1421. From those, 86% were coded as correct based on the human ear. Correspondingly, the samples analyzed included 263 key child labels, but only 39% of those were coded as correct, when analyzed by the human ear (see Table 2). The highest agreement between LENA labels and human coders was found in female labels. The agreements for the following labels were modest or fair: key child, male, and silence. Kappa values were as follows; female $\alpha = .80$ (confidence interval .62;.92), male $\alpha = .28$ (-.36;.81), key child $\alpha = .25$ (-.28;.72) and silence $\alpha = .30$ (-.51;.89).

(INSERT TABLE 2)

Regarding the findings of part B, the mean value for female words based on LENA estimates was 609, and the mean value based on human transcription was 698. Furthermore, based on the LENA system the mean value of child vocalizations was 14 and the same value based on human calculations (i.e. transcription) was 6 (Table 3). Significant correlations were found for the following counts: female words ($r = .91$, $p < 0.001$) and adult words ($r = .95$, $p < 0.001$). Kappa values were as follows; female word count $\alpha = .88$ (confidence interval .74;.96), male word count $\alpha = -.33$ (-.89;.19), adult word count $\alpha = .92$ (.88;.95), child vocalization count $\alpha = .09$ (-.51;.60), conversational turns $\alpha = .25$ (-.24;.72) and silence $\alpha = -0.18$ (-.64;.26).

(INSERT TABLE 3)

DISCUSSION

This study assessed the validity of the LENA system in two different NICUs. Two different methods of analysis were used, and they provided mainly comparable information. The findings showed that the LENA system provides valid information in the NICU settings on adult words, especially on female words. However, the validity of the following LENA values was modest or weak: child vocalization count, conversational turns and silence.

The present study showed that LENA provides valid information on adult words in the NICU setting. This information is important, since maternal, family and caregiver voices are an essential part of the sound environment in the NICU, and have a beneficial effect on the development of preterm newborns (5). The high agreement for adult word count is consistent with findings from other non-English studies (18, 19). The high validity of adult words supports the use of the LENA system when investigating maternal talk in the NICU. However, a lower agreement for male words was found in this study. It is possible that the LENA system could not differentiate all male and female words due to language-specific or prosodic features, such as the lower fundamental voice frequency of females in this Finnish and Estonian sample, compared with English samples. The lower agreement for male words may also have been influenced by lower amounts of male speech in this sample.

The agreements for key child labels, child vocalization counts, and conversational turns were only fair. Agreements for child vocalization counts were lower than in the previous Finnish study (24), where the participant age was within the normative range for LENA. The development of very preterm infants' early vocalizations proceeds based on maturation (i.e. gestational age), rather than chronological age (30). Thus, very early vocalizations of preterm infants are very likely to differ from very early vocalizations of full term infants. Still, in a previous study conducted in the NICU, human coders found early precursors to speech, protophones, including vocants,

squels and growls, at 32 weeks' gestational age to be more frequent than cries (12). In this study, LENA- and human-estimated amounts of child vocalizations were low. One explanation for the differences between the studies may be different sample selection principles. The possible challenge of distinguishing early vocalizations from vegetative sounds, based only on auditory assessment, also needs to be considered. Furthermore, in this study low validity was found for conversational turns. This finding may be influenced by the fact that child vocalizations, which are part of conversational turns, were not identified by the LENA system as reliably as female words. The immature vocalizations of preterm infants may present a challenge in analyzing conversational turns at this age.

The agreement between LENA and human coder estimates of silence were low. LENA automatic measurements differ from human perception and evaluation of silence. It is challenging for a human coder to manually measure the brief durations of silence automatically measured. The initial focus of the LENA system is on language input (15) and the measurement of adult words (17), which may explain why LENA estimates of silence have not been widely investigated. However, it is relevant to further evaluate methods to measure the validity of LENA estimates of silence in NICU settings. Silence, low-level sounds and noise are basic elements of the NICU sound environment and can affect the development of preterm infants (5).

The present study was conducted in two NICUs that support parent-infant closeness (14). At the time of the recordings, most participants in Turku, but a smaller proportion of participants in Tallinn, stayed in single-family rooms. The unit policy and the physical environment may influence the sound environment in different NICUs and, consequently, the data obtained and the validity of measured variables in different studies. Acoustically-designed single family rooms are associated with lower sound levels (5). Overlapping noise or speech can influence the correlation between LENA and human-provided measures (17, 19).

Our study provides information from two less-studied languages, Finnish and Estonian. It is important to gather knowledge from different language contexts since language-specific linguistic features, prosodic features and fundamental frequency of the speaker's voice may influence LENA results. The good agreement for female and the lower agreement for male labels are comparable to findings from the previous Finnish study (24).

This study provides novel validity information on the LENA system in the NICU environment. Another strength is that the information is analyzed in two ways, from two countries and two languages. Further, the participants were randomly selected from a representative sample of very preterm infants. The number of participants in this study may be considered small. However, our sample size is comparable to that in a previous study (21). In addition, the percentage used in the present study for the definition of the sample size (20% of the total sample) is a normal, even good, proportion in reliability analysis. Furthermore, in this study, over two hours' worth (140 minutes in total) of recorded data was analyzed in detail in two different ways. This kind of detailed analysis would not have been possible to accomplish with a larger data set. Still, a larger data set would have provided even more validity information on the variables assessed in the present study.

The results of the present study can be clinically applied when investigating the amount of very early caregiver talk on the development of preterm infants in NICU settings. This could be done, e.g., by studying the effects of developing NICU structures, such as single-family rooms, to support parent-infant closeness and interaction, or for investigating very early intervention in the NICU. LENA estimates for child vocalizations were less valid, indicating that automatic measuring of early infant vocalizations and conversational turns in this population is challenging. Further research is needed to evaluate and develop tools for analyzing early vocalizations of preterm infants.

ACKNOWLEDGEMENTS

We wish to thank the participating families. We are grateful to Ella Hämäläinen, Minna Paaso, Jana Rojak and Kati Siirilä for help with data collection or coding.

LIST OF ABBREVIATIONS

CPAP, continuous positive airway pressure; Kalpha, Krippendorff's alpha; LENA, Language Environment Analysis; NICU, neonatal intensive care unit

STATEMENT OF CONFLICT OF INTEREST

None declared.

FUNDING

This research was financed by the Doctoral Programme in Psychology, Learning and Communication, University of Helsinki, and the Swedish Cultural Foundation in Finland.

REFERENCES

1. Barre N, Morgan A, Doyle LW, Anderson PJ. Language abilities in children who were very preterm and/or very low birth weight: a meta-analysis. *J Pediatr* 2011; 158 5:766–74.
2. van Noort-van der Spek IL, Franken MC, Weisglas-Kuperus N. Language functions in preterm-born children: a systematic review and meta-analysis. *Pediatrics* 2012; 129 4:745–54.
3. DeMaster D, Bick J, Johnson U, Montroy JJ, Landry S, Duncan AF. Nurturing the preterm infant brain: leveraging neuroplasticity to improve neurobehavioral outcomes. *Pediatr Res* 2019; 85 2:166–75.
4. Moon C. Prenatal Experience with the Maternal Voice. In: Filippa M, Kuhn P, Westrup B, eds. *Early Vocal Contact and Preterm Infant Brain Development: Bridging the Gaps Between Research and Practice*: Springer, 2017.
5. Philbin MK. The Sound Environments and Auditory Perceptions of the Fetus and Preterm Newborn. In: Filippa M, Kuhn P, Westrup B, eds. *Early Vocal Contact and Preterm Infant Brain Development Bridging the Gaps Between Research and Practice*: Springer, 2017.
6. Voegtline KM, Costigan KA, Pater HA, DiPietro JA. Near-term fetal response to maternal spoken voice. *Infant Behav Dev* 2013; 36 4:526–33.
7. Guttorm TK, Leppänen PHT, Poikkeus A-M, Eklund KM, Lyytinen P, Lyytinen H. Brain Event-Related Potentials (ERPs) Measured at Birth Predict Later Language Development in Children with and Without Familial Risk for Dyslexia. *Cortex* 2005; 41 3:291–303.
8. White RD, Smith JA, Shepley MM, Committee to Establish Recommended Standards for Newborn ICUD. Recommended standards for newborn ICU design, eighth edition. *J Perinatol* 2013; 33 Suppl 1:S2–16.
9. Flacking R, Lehtonen L, Thomson G, Axelin A, Ahlqvist S, Moran VH, et al. Closeness and separation in neonatal intensive care. *Acta Paediatr* 2012; 101 10:1032–7.
10. Caskey M, Stephens B, Tucker R, Vohr B. Importance of parent talk on the development of preterm infant vocalizations. *Pediatrics* 2011; 128 5:910–6.
11. Caskey M, Stephens B, Tucker R, Vohr B. Adult talk in the NICU with preterm infants and developmental outcomes. *Pediatrics* 2014; 133 3:e578–84.
12. Oller DK, Caskey M, Yoo H, Bene ER, Jhang Y, Lee CC, et al. Preterm and full term infant vocalization and the origin of language. *Sci Rep* 2019; 9 1:14734.
13. Caskey M, Vohr B. Assessing language and language environment of high-risk infants and children: a new approach. *Acta Paediatr* 2013; 102 5:451–61.
14. Raiskila S, Axelin A, Toome L, Caballero S, Tandberg BS, Montiroso R, et al. Parents' presence and parent-infant closeness in 11 neonatal intensive care units in six European countries vary between and within the countries. *Acta Paediatr* 2017; 106 6:878–88.
15. Gilkerson J, Richards JA, Warren SF, Montgomery JK, Greenwood CR, Kimbrough Oller D, et al. Mapping the Early Language Environment Using All-Day Recordings and Automated Analysis. *Am J Speech Lang Pathol* 2017; 26 2:248–65.
16. Gilkerson J, Richards AJ. The LENA Natural Language study (Technical report LTR-02-02). Technical Report: LENA Foundation; 2008

17. Xu D, Yapanel U, Gray S, Gilkerson J, Richards J, Hansen J. Signal Processing for Young Child Speech Language Development. The 1st workshop on child, computer and interaction. Chania, Crete, Greece.2008
18. Gilkerson J, Zhang Y, Xu D, Richards JA, Xu X, Jiang F, et al. Evaluating language environment analysis system performance for Chinese: a pilot study in Shanghai. *J Speech Lang Hear Res* 2015; 58 2:445–52.
19. Canault M, Le Normand MT, Foudil S, Loundon N, Thai-Van H. Reliability of the Language ENvironment Analysis system (LENA) in European French. *Behav Res Methods* 2016; 48 3:1109–24.
20. Ganek HV, Eriks-Brophy A. A Concise Protocol for the Validation of Language ENvironment Analysis (LENA) Conversational Turn Counts in Vietnamese. *Communication Disorders Quarterly* 2018; 39 2:371–80.
21. Ganek H, Eriks-Brophy A. Language ENvironment analysis (LENA) system investigation of day long recordings in children: A literature review. *J Commun Disord* 2018; 72:77–85.
22. Dasinger L. Issues in the Acquisition of Estonian, Hungarian and Finnish: A Crosslinguistic Comparison. In: Slobin DI, ed. *The crosslinguistic study of language acquisition* Laurence Erlbaum Associates, Inc., 1997.
23. Leino T, Laukkanen AM, Ilomaki I, Maki E. Assessment of vocal capacity of Finnish university students. *Folia Phoniatr Logop* 2008; 60 4:199–209.
24. Elo H. *Acquiring Language as a Twin : Twin children's early health, social environment and emerging language skills.* University of Tampere; Tampere University Press; 2016
25. Gilkerson J, Richards AJ. *A Guide to Understanding the Design and Purpose of the LENA® System.* Technical Report Boulder, CO: LENA Foundation; 2020
26. Gilkerson J, Coulter KK, Richards JA. *Transcriptional Analyses of the LENA Natural Language Corpus (Technical report LTR-06-2).* Boulder, CO: LENA Foundation; 2008
27. Oller DK, Niyogi P, Gray S, Richards JA, Gilkerson J, Xu D, et al. Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences* 2010; 107 30:13354–9.
28. Barras C, Geoffrois E, Wu Z, Liberman M. Transcriber: Development and use of a tool for assisting speech corpora production. *Speech Communication* 2001; 33 1:5–22.
29. Hayes AF, Krippendorff K. Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures* 2007; 1 1:77–89.
30. Stolt S, Lehtonen L, Haataja L, Lapinleimu H. Development and predictive value of early vocalizations in very-low-birth-weight children: a longitudinal study. *Clin Linguist Phon* 2012; 26 5:414-27.

TABLE 1 Background characteristics of participants. Mean values and standard deviations (SD) or numbers (N) and percentage of participants (%) are shown.

	Turku N=7	Tallinn N=7	Total N=14
Age (weeks): mean (SD)			
Gestational age at birth	27 (3)	27 (2)	27 (2)
Gestational age at recording day	33 (1)	33 (0)	33 (0)
Birth weight (grams): mean (SD)			
	924 (388)	1166 (305)	1045 (358)
Gender: n (%)			
Female	4 (57)	3 (43)	7 (50)
Male	3 (43)	4 (57)	7 (50)
Respiratory support at recording day: n (%)			
None	2 (29)	5 (71)	7 (50)
Invasive ventilation	1 (14)	0	1 (7)
CPAP ^a	2 (29)	1 (14)	3 (21)
High flow nasal cannula	1 (14)	1 (14)	2 (14)
High flow nasal cannula/CPAP ^a	1 (14)	0	1 (7)
Warmth regulation at recording day: n (%)			
None	3 (43)	1 (14)	4 (29)
Incubator	0	1 (14)	1 (7)
Warming mattress	4 (57)	5 (71)	9 (64)
Type of room at recording day^b: n (%)			
Single family room	6 (86)	1 (14)	7 (50)
Double family room	1 (14)	3 (43)	4 (29)
Room for 3 patients	0	1 (14)	1 (7)
Room for 4 patients	0	2 (29)	2 (14)
Maternal education: n (%)			
High school	1 (14)	2 (29)	3 (21)
Occupational	4 (57)	1 (14)	5 (36)
Lower university	0	2 (29)	2 (14)
Upper university	2 (29)	2 (29)	4 (29)
Paternal education: n (%)			
High school	0	3 (43)	3 (21)
Occupational	5 (71)	1 (14)	6 (43)
Lower university	1 (14)	3 (43)	4 (29)
Upper university	1 (14)	0	1 (7)

^acontinuous positive airway pressure

^bchairs for Kangaroo care/skin-to-skin care were available in all type of rooms

TABLE 2 The descriptive statistics of part A. The total number of labels provided by LENA and the number of correct labels when analyzed by the human ear are presented. Mean values, standard deviations and minimum–maximum values for the LENA labels of each group and each value are shown. Agreement percentage is also presented.

	LENA labels			Human coder agreements			Agreement percentage
	N	Mean (SD)	Min–max	N	Mean (SD)	Min–max	Mean
TURKU (N=7)							
Female	824	118 (13)	100–132	740	106 (13)	82–124	90
Male	355	51 (45)	2–115	176	25 (35)	0–99	42
Key child	141	20 (32)	2–89	54	8 (11)	0–33	54
Silence ^a	191	48 (32)	8–83	42	36 (30)	6–78	72
TALLINN (N=7)							
Female	597	85 (33)	27–118	507	72 (38)	20–118	82
Male	223	32 (20)	7–65	118	17 (22)	0–61	53
Key child	122	17 (21)	3–53	47	7 (16)	0–43	24
Silence ^a	276	46 (46)	4–116	74	12 (26)	0–66	22
TOTAL (N=14)							
Female	1421	102 (29)	27–132	1247	89 (33)	20–124	86
Male	578	41 (35)	2–115	294	21 (29)	0–99	47
Key child	263	19 (26)	2–89	101	7 (13)	0–43	39
Silence ^a	467	47 (39)	4–116	216	22 (29)	0–78	42

^aLENA labels for silence were noted for 4 recordings in Turku and 6 recordings in Tallinn, totally for 10 recordings.

TABLE 3 The descriptive statistics for part B. Mean, standard deviations (SD), minimum and maximum values (min.–max.) for LENA and human estimates are presented. Lengths of silence are presented in seconds. Agreements between LENA and human estimates are presented using Spearman's correlation efficient values (*r*). Significance level (*p*) is also displayed.

	LENA estimates		Human coder estimates		<i>r</i>	<i>p</i>
	Mean (SD)	Min–max	Mean (SD)	Min–max		
TURKU (N=7)						
Female words	712 (272)	263–1105	812 (233)	484–1207	.79	.04
Male words	266 (314)	12–862	157 (207)	0–592	.86	.01
Adult words	978 (220)	708–1288	969 (241)	662–1312	.82	.02
Child vocalizations	13 (21)	1–59	8 (4)	0–12	.13	.79
Conversational turns	7 (9)	1–26	4 (3)	0–7	.73	.06
Silence	133 (166)	28–185	41 (38)	0–109	.61	.15
TALLINN (N=7)						
Female words	505 (452)	12–1249	584 (376)	110–1166	1.00	<0.001
Male words	115 (166)	5–476	81 (146)	0–403	.52	.30
Adult words	620 (438)	156–1331	665 (339)	224–1256	.96	<0.001
Child vocalizations	15 (21)	1–55	4 (5)	0–16	-.10	.83
Conversational turns	8 (9)	1–26	4 (5)	0–14	.08	.87
Silence	267 (106)	139–414	70 (37)	15–106	.86	.01
TOTAL (N=14)						
Female words	609 (374)	12–1249	698 (323)	110–1207	.91	<0.001
Male words	191 (254)	5–862	119 (323)	0–592	.64	.02
Adult words	799 (381)	156–1331	817 (324)	224–1312	.95	<0.001
Child vocalizations	14 (20)	1–59	6 (5)	0–16	-.03	.91
Conversational turns	8 (9)	1–26	4 (4)	0–14	.23	.44
Silence	200 (110)	28–414	56 (39)	0–109	.64	.01