

Reliability of the Language ENvironment Analysis system (LENA™) in European French

Mélanie Canault¹ · Marie-Thérèse Le Normand^{2,6} · Samy Foudil³ · Natalie Loundon⁴ · Hung Thai-Van⁵

Published online: 15 July 2015
© Psychonomic Society, Inc. 2015

Abstract In this study, we examined the accuracy of the Language ENvironment Analysis (LENA) system in European French. LENA is a digital recording device with software that facilitates the collection and analysis of audio recordings from young children, providing automated measures of the speech overheard and produced by the child. Eighteen native French-speaking children, who were divided into six age groups ranging from 3 to 48 months old, were recorded about 10–16 h per day, three days a week. A total of 324 samples (six 10-min chunks of recordings) were selected and then transcribed according to the CHAT format. Simple and mixed linear models between the LENA and human adult word count (AWC) and child vocalization count (CVC) estimates were performed, to determine to what extent the automatic and the human methods agreed. Both the AWC and CVC estimates were very reliable ($r = .64$ and $.71$, respectively) for the 324 samples. When controlling the random factors

of participants and recordings, 1 h was sufficient to obtain a reliable sample. It was, however, found that two age groups (7–12 months and 13–18 months) had a significant effect on the AWC data and that the second day of recording had a significant effect on the CVC data. When noise-related factors were added to the model, only a significant effect of signal-to-noise ratio was found on the AWC data. All of these findings and their clinical implications are discussed, providing strong support for the reliability of LENA in French.

Keywords Adult word count · Child vocalization count · Reliability · Human transcriber · European French · Automatic speech recognition technology · Signal-to-noise ratio

Studies in child language acquisition and disorders generally require reliable audio recordings. Most of the time, these recordings are limited in duration since they must meet physical constraints. The main goal of the experimenter is to develop a data acquisition system that is non-invasive and also preserves the quality of the recordings. Whereas fixed recording systems can restrict the child's movements, mobile systems, usually placed on the child, may affect the quality of recordings by adding friction noises. Experimenters, as a result, have to rely upon their own ingenuity to acquire a sufficient amount of high-quality data. Still, the problems do not end there. Once the data collection stage has been completed, data processing remains a time-consuming and tedious task, even when performing simple word counts. Localization of the child's productions, versus those addressed by adults to the child, is mandatory before transcribing them and then analyzing their content. All these constraints are liable to limit the number of tested subjects, and hence the amount of data collected. To overcome these drawbacks, a system was launched in 2004 allowing for large-scale all-day audio recording and

Marie-Thérèse Le Normand and Hung Thai-Van joint authorship.

✉ Marie-Thérèse Le Normand
marielenormand@mac.com

¹ Laboratoire Dynamique du Langage, UMR 5596 CNRS, Université Lumière Lyon 2, Lyon, France

² INSERM & Laboratoire Psychopathologie et Processus de Santé, Université Paris Descartes, Sorbonne, Paris Cité, France

³ Université Lumière Lyon 2, Lyon, France

⁴ Service ORL et Unité d'Audiophonologie, Hôpital Necker, Paris, France

⁵ Service d'Audiologie & d'Explorations Orofaciales, Hospices Civils de Lyon & Centre Recherche en Neurosciences de Lyon, INSERM, U1028, CNRS UMR 5292, Lyon, France

⁶ LPP, Institut de Psychologie, 71 avenue Edouard Vaillant, 92100 Boulogne Billancourt, France

automated vocal analyses of speech segments. This system, known as LENA (Language ENvironment Analysis) has already been used to track communication skills development from 2 to 48 months of age (Christakis et al., 2009; Oller, 2010; Warren et al., 2010; Zimmerman et al., 2009).

The LENA system is an innovative tool, which opens up several perspectives for researchers working on child language development and disorders using naturalistic language samples. The system can collect audio data either in typically developing children or in children with atypical development—for example, language delay, sensory impairment, or autism spectrum disorders.

Aragon and Yoshinaga-Itano (2012) examined LENA measures of the home language environment across English- and Spanish-speaking families of typically developing children and children with hearing loss. Interestingly, despite lower socioeconomic status and average maternal levels of education in Spanish-speaking homes, great similarities were found between child vocalizations, adult word counts, and conversational turns in children with typical development in English-speaking families and in children with hearing impairments in Spanish-speaking families. This good unexpected outcome brought the impact of language input on child language development to the attention of caregivers and highlighted the importance of early intervention.

The scope of automatic sound environment processing for language development is not restricted to deaf children. Other studies have focused on autistic children, since language input and conversational exchanges also seems to be indicators for communicative development in these children (Siller & Sigman, 2002; Warlaumont, Richards, Gilkerson, & Oller, 2014; Xu, Gilkerson, Richards, Yapanel, & Gray, 2009; Yoder, Oller, Richards, Gray, & Gilkerson, 2013). For example, Warren et al. (2010) demonstrated not only that autistic children are engaged in fewer interactions and vocalizations than typically developing children but also that their productions increase with the number of words that are addressed to them. This suggests that strong stimulation, such as that provided during therapy sessions, is a reliable means to increase the language productions of autistic children and provides valuable information for their management. Caskey, Stephens, Tucker and Vohr (2011) and Caskey and Vohr (2013) suggested that LENA showed that a high exposure to parental language is predictive of the vocalizations and interactions addressed to preterm children. Jackson and Callender (2013) used LENA for impoverished migrants who predominantly speak another language at home and attend English-speaking childcare or preschools.

LENA is endowed with a system of automatic processing of the speech signal, providing quantitative information about the language environment of infants, toddlers and preschoolers. This is made possible through an audio recording device that allows for analysis and automatic classification of the speech and linguistic environment of an individual child.

The LENA system consists of a digital language processor that records a full day of language used by a child and his/her communication partners and special software that processes the audio recording to provide automatic reports. From the outset, LENA had the clinical aim of demonstrating to caregivers and to families the positive impact of early language input, particularly in the cases of child developmental disorders. LENA researchers confirmed the longitudinal study by Hart and Risley (1995), which showed the close relationship between the flow of language addressed to children from an early age and their later vocabulary skills, IQ test scores and academic achievement. Forty-two families were followed between the seventh and thirty-sixth month of their child and were recorded at a rate of 1 h per month, resulting in a total of 1,318 h that were manually transcribed and analyzed by the authors over 6 years. Since this very time-consuming kind of study is rare and difficult to replicate, LENA was designed to facilitate and expand such research using fully automated procedures.

The LENA system has several advantages. Because it is a very small, lightweight device, it fits easily onto clothing worn by the target child and enables the acquisition of good quality data for 10 to 16 h at a time (Christakis et al., 2009; Warren et al., 2010; Xu, Yapanel, & Gray, 2009; Zimmerman et al., 2009). Thus, audio recordings of a child in his/her actual sound environment all day long, without an experimenter being present, are feasible. Once data collection is complete, the resulting audio file is transferred, analyzed and processed automatically by the computer program. The program provides viewable reports on the target child's number of vocalizations (CVCs; an estimate of the number of speech or pre-speech productions by the target child per hour or per day), adult word counts (AWCs; an estimate of the number of adult words spoken near a child per hour or per day), conversational turns, and duration of exposure to electronic media (television, radio and other interactive electronic devices).

The LENA system is based on an acoustic model for automatic speech recognition (Gilkerson & Richards, 2008, 2009; Xu et al., 2008). The model first allows a segmentation of audio signals into different categories depending on whether the sounds come from human speech or from the acoustic background. The system tags the speech as having been produced by adult males, adult females, the target child or other children. It labels other stretches as noise, media, overlaps, and silences. In a second step, an estimate of the number of vocalizations or words within the sequences corresponding to speech segments is performed. Both the segmentation and the estimate processes require a fine description of acoustic criteria for voice identification and automated acoustic feature analysis. However, it cannot be excluded that the complexity of the source signal (e.g., a noisy environment) may affect the performance of the system. Environmental effects are thought to be the largest source of variability (Xu, Yapanel, & Gray,

2009; Xu et al., 2008). For instance, reverberation effects (echo) resulting from the size of the room, type of flooring, environmental location as well as the effects of distance may negatively impact the signal integrity, whereas human ears are much less sensitive to these phenomena. The voice characteristics of the speakers i.e., the speed of delivery, pitch, accent and any dialectal variations, may also affect the reliability of the LENA estimates. The degree of accuracy of the LENA system has hence been evaluated for American English. When checking for the segmentation performed by LENA, the degree of consistency reported between LENA and human transcription reached 82 % for AWC, 76 % for CVC, and 71 % for electronic media recognition (Xu, Yapanel, & Gray, 2009). This high degree of accuracy between LENA and human count estimates was confirmed by Oetting, Hartfield, and Pruitt (2009) for AWC, although the same authors found much poorer correlations for estimates of conversational turns ($r = .08-.14, p > .05$).

With regard to AWC, Xu, Yapanel, and Gray (2009) found that LENA counts on average 2 % fewer words than human transcribers. One explanation is that LENA may identify speech productions on the basis of distinct temporal sequences, whereas human ears are well-trained to separate overlapping speech flows, thus resulting in higher word counts. In addition, in the presence of different sound sources—for example, in noisy situations and outdoor recordings, much larger discrepancies between LENA and human counts are reported. This is particularly true in the case of a degraded signal-to-noise ratio (SNR). The complexity of the sound environment is therefore thought to have a negative impact on reliability. Regarding CVC, LENA recognizes all vocalizations quite well while it discards vegetative sounds (e.g., coughs, breathing, digestion), fixed sounds (cry, laughter), and overlapping speech. In sum, LENA algorithms correctly detect 75 % of the Child Utterances Clusters consisting of periods identified as pertaining to the voice of the key child, and that are not ignored or interrupted by the utterances of any other speaker (labeled as “male adult,” “female adult,” or “other child”) or by silence or noise lasting more than 800 ms (Oller et al., 2010).

Most of the studies using LENA were carried out on English-speaking children (Burgess, Audet, & Harjusola-Webb, 2013; Oller, 2010; Sacks et al., 2014; Soderstrom & Whittebolle, 2013; Warren et al., 2010). Given the new diagnostic and therapeutic horizons opened up by LENA mentioned at the outset, it might be of great interest for the international scientific community, especially in countries where more than one language is spoken. This raises the question of whether or not the use of LENA can be generalized to all languages of the world. This is far from obvious since LENA was designed using English language and its speech signal processing model. Since every language has its own phonetic and acoustic features, it is therefore necessary to validate LENA in other languages than English. Despite a high

level of correlation between LENA count estimates and those provided by transcribers, it should be emphasized again that the LENA system was developed and validated mostly in native-English speakers. In a recent study, Weisleder and Fernald (2013) investigated the influence of the amount of child-directed speech on language development between 19 and 24 months of age in 29 low-socioeconomic status Spanish-speaking families living in the United States. For this, they first compared LENA AWCs with those provided by native Spanish-speaking transcribers, and found, on the basis of 60-min recorded samples per child, a high correlation ($r = .80$). Although no normative data using LENA in Spanish have yet been published, this calls for expanding home language environment measures with LENA to Romance languages. Besides, international validation studies of LENA have been initiated in Asian languages, either tonal ones such as Mandarin (Gilkerson et al., 2015; Zhang, 2013), or nontonal, such as Korean (Pae, 2013). As regards the use of LENA in French, only preliminary data have been presented so far (Canault & Thai-Van, 2013). Table 1 summarizes most of these studies on LENA reliability for different languages, showing the moderate to high correlations or likelihood according to the methodologies adopted by the authors. Since too few cross-linguistic reliability studies have been reported in the LENA literature, it is crucial to understand the usefulness of LENA in different languages, not only as an innovative automatic speech recognition technology but also to clarify the controversial debate on universal and language diversity issues.

The speech characteristics of European French are highly different from those of English. The English vowel system includes full diphthongs and oppositions between short and long vowels, whereas these two linguistic features are absent in French. Although this should not play a decisive role in the processing performed by LENA, other acoustic speech cues may affect LENA automated procedures for vocal analysis. Differences in speech rate have been clearly identified between the two languages, with an average rate of 7.18 syllables per second in French versus 6.19 syllables per second in English. This difference is thought to reflect the complexity of the syllabic structure, that is, a larger variety of syllabic components for English than French (Pellegrino, Coupé, & Marsico, 2011). Differences in prosody also exist since English and French have different rhythms. English is defined as a “stress-timed” language, with stresses occurring at regular intervals. English syllables may thus be more or less compressed to maintain a fixed duration of the stressed group. By contrast, French is defined as a “syllable-timed” language, with syllables of almost fixed duration (Abercrombie, 1967). In other words, the dominant acoustic speech cues are stress-related and thus spectral in English (Bolinger, 1985), whereas they are temporal in French (Wenk & Wioland, 1982). There is another reason why English is known as a stress-timed language: That is, depending on whether a given word starts

Table 1 Summary of studies on LENA reliability in different languages

Source	Language	Recording	Age	LENA–Human	Estimates	<i>r</i>	Likelihood Ratio ¹
Gilkerson & Richards, 2008	American English	<i>N</i> = 70	2–48 mo	six 10-min chunks of recording (1 h) 32 boys 38 girls	AWC CVC		88 % 91 %
Xu, Yapanel & Gray, 2009	American English	<i>N</i> = 2	10 mo 31 mo	12 h	AWC CVC CTC	.92	
Oetting et al., 2009	American English	<i>N</i> = 17		30-min chunks	AWC CTC	.71–.85 .08–.14	
VanDam & Silbert, 2013a, b	Canadian English	<i>N</i> = 26	30 mo	2,340 chunks	AWC CVC		70 %
Weisleder & Fernald, 2013	Low SES Spanish	<i>N</i> = 10	19–24 mo	six 10-min chunks of recording (1 h)	AWC	.80	
Canault & Thai-Van, 2013	French	<i>N</i> = 18 9 boys 9 girls	3–48 mo	six 10-min chunks of recording (1 h per day)	AWC CVC	.64 .71	
Gilkerson et al., 2015	Chinese SDM (Shanghai Dialect and Mandarin)	<i>N</i> = 22 <i>N</i> = 19 10 boys 20 girls	3–23 mo	15-min chunks of recordings 5.5 h	AWC CVC		72 % 73 % 22 % 72 %

Likelihood ratio test is used to compare the goodness of fit of two models, one of which (the null model) is a special case of the other (the alternative model).

with a strong or a weak syllable (accentual organization), its meaning can differ (e.g., 'record vs. re'cord). Since 85 %–90 % of lexical words have a strong initial syllable (Cutler & Carter, 1987), this may serve as a key criterion for speech segmentation in English (Cutler & Carter, 1987; Cutler & Norris 1988). In French, it is rather the group of words as a whole that is stressed (Delattre, 1962), although the role of cues related to word boundaries, such as the lengthening of the final syllable (Adda-Decker, Gendrot, & Nguyen, 2008) or F0 initial raising, cannot be ruled out (Vaissière, 2010). Language diversity can be an obstacle to defining universal acoustic parameters for word counting. Yet the LENA designers have to face this challenge.

In the present study, we aimed to examine the accuracy of the LENA system in European French investigating the relations between LENA and human AWCs and CVCs. Three research questions were raised: What would be the correlations between the LENA and the human AWC and CVC estimates in the selected total recording (6 ages × 3 children in each × 3 days × 6 10-min chunks of recordings—i.e., 324 samples)? Second, because each participant was recorded three days over a week to ensure the collection of enough data, would 1 h of the selected recording per participant be sufficient to obtain a reliable sample (6 10-min chunks of recordings × 18 participants—i.e., 108 samples)? Finally, would the validity between LENA and human count estimates remain accurate in the AWC and CVC data when the noise-related factors, such as the SNR, and the two LENA categories NON (noise near) and OLN (overlap near) were added in the linear mixed model? Distance factors related to the intense activity

in a noisy context may have an impact on the validity between LENA and human counts as Xu, Yapanel, and Gray (2009) have suggested.

Method

Participants

The participants were selected by e-mail. They were mainly voluntary middle to high-class families working at the Edouard Herriot Hospital in Lyon. Eighteen native typically developing French-speaking children (nine girls, nine boys) ranging from 3 to 48 months of age, without any auditory and developmental neurocognitive disorders were selected for this study. They were divided into six age groups, each of them corresponding to a crucial stage of language development: vocalizations from 3 to 6 months, babbling between 6 to 12 months, first words between 12 to 18 months, vocabulary spurt between 18 to 24 months, grammatical spurt between 24 to 36 months and stabilization of grammar between 36 to 48 months. Each age group included three participants. Written consent was obtained from parents with legal responsibility for the child.

Data collection

Each participant was recorded for a minimum of 10 h (up to 16 h) per day, three days a week, using the LENA digital language processor recording device. The DLP was easily fitted into the child's clothing or placed nearby when the child

could not wear it—for example, during a bath or nap. To avoid a potential methodological bias related to the quality of the audio recording, all parents were instructed to use appropriate clothing provided by the LENA foundation and to switch on the DLP all day long. The data were collected in children’s natural environments: home, outside, nursery and anywhere else the children went, over one week with a high ecological validity as the recording situation closely approximates real-life situations with one or multiple speakers at home or in a daycare context.

Recording selection

A total of 324 samples were selected for the 18 participants. For each participant, six chunks of 10-min recordings were selected per day, resulting in eighteen chunks for the three recording days spread over less than one week. Each audio recording was selected independently at random by two volunteer research assistants. All the types of activity engaged in by the child and different times of day were included, for instance, mealtime, bathtime, storytime, playtime, and time outside with different levels of noise. We mainly selected chunks in which the number of productions by the child and the adult was the highest. Thus, the chunks of recordings in which we observed no productions, such as naptime, were excluded.

Data transcription and coding

To assess the reliability of LENA in European French, we focused on AWC and CVC estimates, two variables that have been reported to be measured by LENA with a high degree of accuracy in American English (Oetting et al., 2009; Xu, Yapanel, & Gray, 2009). The objective, here, was to compare AWC and CVC estimates generated by the LENA system to the output that we obtained from transcribing the 324 selected samples (human count estimate). First, CVC estimates the number of any speech-like babbling or vocalizations within a child utterance cluster. Fixed signals (e.g., cries, screams) and vegetative noises (e.g., burping) were not count as vocalizations. For instance, if the child said “ba” or “bababababa” this was counted as one vocalization, whereas if the child said “bababa # baba,” this was counted as two vocalizations. During the single-word period and the two-to-three-word combinations, a word is counted as one vocalization. Thus, if the child said “bababa#papa#parti” (baba#daddy#gone), this was counted as three vocalizations. We extracted all the automated vocalization segments from the key child for comparison with the transcriptions. The AWCs were estimated for each segment identified as an adult speaker on the basis of the LENA speech-processing algorithms. The software does not attempt to segment or label specific words or word boundaries; instead, the software uses statistical models to estimate

the number of words per speaker segment. The sum of meaningful speech segments by female adults and male adults is reported as the AWCs. Figure 1 shows an audio example, with the human transcription of the AWCs and CVCs and the LENA labels.

A total 324 selected recordings, representing a total of 54 h, were orthographically transcribed by two native French speakers, each hour of transcription resulting from the concatenation of six chunks of 10-min each. The two transcribers used the *FREQ* and the *MLU* program from Computerized Language ANalysis (CLAN) to report the word count (MacWhinney, 2000) consistently following these rules:

A word was transcribed orthographically as meaningful speech if it contained at least one syllable. Words such as *chien* (“dog”), *maman* (“mummy”), *aller* (“gonna”), and *bleu* (“blue”) were counted as one word. Free morphemes such as the determiners *le*, *la*, and *les* (“the”); the prepositions *à* (“to”), *de* (“of”), and *par* (“by”); and the pronouns *je* (“I”), *il* (“he”), and *elle* (“she”), and so forth, were counted as one word, similarly to bound morphemes, including prefixes and suffixes, because these words cannot be broken down into two or more morphemes. Thus, the word *défaire* (“to undo”) or *malheureuses* (“unhappy”) was counted as one word.

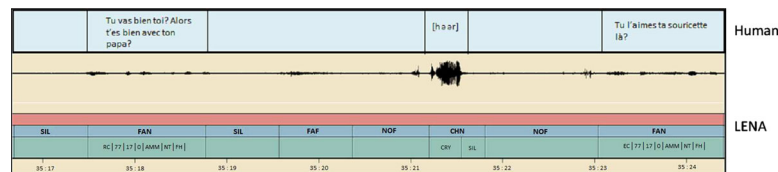
Elided forms such as “c,” “d,” “j,” “l,” “m,” “n,” “s,” “t,” and “qu” (for *ce*, *de*, *je*, *le*, *me*, *ne*, *se*, *te* or *tu*, and *que*) are grammatical words containing an apostrophe in the written code. In the spoken form, these elisions are mainly related to the fact that in most of the written forms the “e” is silent. Consequently, the chunk *l’chien* (“the dog”) was counted as one word, whereas *je sais pas* (“I don’t know”), *t’as vu* (“you see”), and *j’vais l’faire* (“I am going to do it”) were counted as two words.

Compound words containing independent elements (whether separated or not by a hyphen in the written code) were broken down into meaningful subunits; for instance, the compound words *après-midi* (“afternoon”) and *petit-déjeuner* (“breakfast”) were counted as two words.

Every onomatopoeia, defined as a sound associated with what is named, was counted as one word. For instance, *boum parti* (“boom gone”) were counted as two words.

Data analyses

To assess the reliability of the LENA system, we extracted the AWCs and CVCs generated by LENA (Software Version: V3.0.1) for each recording. LENA generates a segmentation map of a recording stream. All segments are labeled by being matched statistically to one of the following eight categories: CH (child), CX (other child), FA (female adult), MA (male adult), OL (overlap), TV (electronic media), NO (noise), and SIL (silence). The seven categories other than SIL are further divided into two types, depending on how N (near) or F (far) each segment is from the statistical model for that category (for instance, MAN stands for male adult near). The intensity



(SIL= *silence*, FAN = *female adult near*, FAF = *female adult far*, NOF = *noise far*, CHN = *child near*)

Excerpt from participant 1, aged 3 mo, 8.55a.m, Home recording day 3

@Begin

*FAN: tu vas bien toi? alors t'es bien avec ton papa?

(‘hey are you okay? so do you feel good with your dad?’)

*CHN: heer

*FAN: tu l'aimes ta “souricette” là?

(‘do you like your little mouse there?’)

@End

Fig. 1 Audio sample representing human transcription and LENA labels of adult word counts and child vocalization counts

of the segments of child and adult language productions was therefore compared to that of segments labeled by the LENA software as NON/NOF (i.e., subsegments corresponding to noise both proximal and distal) and SIL (i.e., subsegments with a sound level below 32-dB). The SNRs for the AWCs and CVCs were based on the comparison of the intensity of the following LENA segments CHN versus NON + NOF + SIL (child near vs. noise near + noise far + silence) for the computation of CVC and MAN + FAN versus NON + NOF + SIL (male adult near + female adult near + silence vs. noise near + noise far + silence) for the computation of AWCs. The SNR [$SNR = 20 \log_{10}(x_{rms}) - 20 \log_{10}(b_{rms})$] was computed using root mean square (RMS). The RMS formula was $x_{rms} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \dots + x_n^2)}$. RMS was defined as the ratio of signal power to the noise power corrupting the signal, with a ratio higher than 1.1 (greater than 0 dB) indicating more signal than noise. LENA and human count estimates and SNR values for adult and child were statistically analyzed per participant. Correlation coefficients were calculated to ensure the consistency between the LENA AWCs and CVCs and the human AWCs and CVCs.

Results

Reliability

A third of the audio recording files—that is, 108 samples—were transcribed independently by a second expert (57,487

words for Expert 1 and 59,918 words for Expert 2). They each received the same instructions for transcription i.e., to follow the CHILDES manual and to use the CLAN and the KidEVAL program. They marked start- and endpoints of the utterance and counted the words in each utterance. Tables 2 and 3 show the reliability between the two transcribers for the intelligible and unintelligible words. The correlations derived from these data are very high ($r = .99, p < .001$).

LENA–human correlations

Figure 2 displays scatterplots between the LENA and human AWCs and CVCs for the selected dataset—that is, 324 samples. The green lines represent the points at which LENA estimates were equal to human-transcribed estimates. The scatterplots show that both the LENA AWC and CVC were significantly correlated with their corresponding human-transcribed estimates ($r_s = .64$ and $.71$, respectively; $p < .001$). However, both LENA AWC and CVC were underestimated in comparison to human-transcribed vocalizations and word counts. This does not preclude a good reliability between LENA and human counts in French language.

Figure 3 displays scatterplots between the LENA and human AWC and CVC estimates by age groups. The scatterplots show that both the LENA AWCs and CVCs were significantly correlated with their corresponding human-transcribed estimates, with correlations ranging from $.61$ to $.87$ ($p < .001$) and from $.39$ to $.83$ ($p < .001$), respectively. All of these correlations indicate a good reliability of LENA and human counts

Table 2 Reliability between the two transcribers for intelligible words

Participant	Gender	Age Group	Transcriber1	Transcriber2	Word Differences
P1	girl	0–6 mo	3,383	3,399	16
P2	boy	0–6 mo	2,149	2,147	2
P3	girl	0–6 mo	2,898	2,863	35
P4	girl	7–12 mo	4,776	4,902	126
P5	boy	7–12 mo	3,735	3,839	104
P6	boy	7–12 mo	2,913	2,915	2
P7	boy	13–18 mo	2,159	2,237	78
P8	girl	13–18 mo	2,993	3,062	69
P9	girl	13–18 mo	3,250	3,400	150
P10	boy	19–24 mo	2,982	3,058	76
P11	girl	19–24 mo	2,767	2,837	70
P12	boy	19–24 mo	2,426	2,485	49
P13	boy	25–36 mo	3,703	3,851	148
P14	girl	25–36 mo	3,695	3,893	198
P15	boy	25–36 mo	4,288	4,355	67
P16	boy	37–48 mo	3,241	3,329	88
P17	girl	37–48 mo	2,442	2,645	203
P18	girl	37–48 mo	3,687	3,901	214
TOTAL			57,487	59,118	1,631

according to child's age. Furthermore, the correlations between age and LENA counts and between age and human counts were significantly similar on the CVC

data ($r_s = .37$ and $.49$, $p < .001$, respectively). This is not the case for the correlations between age and LENA AWC data ($r = .01$, $p > .05$) nor between age and

Table 3 Reliability between the two transcribers for unintelligible words

Participant	Gender	Age Group	Transcriber1	Transcriber2	Differences
P1	girl	0–6 mo	201	191	10
P2	boy	0–6 mo	87	77	10
P3	girl	0–6 mo	192	157	35
P4	girl	7–12 mo	607	597	10
P5	boy	7–12 mo	1,077	1,038	39
P6	boy	7–12 mo	583	544	39
P7	boy	13–18 mo	151	136	15
P8	girl	13–18 mo	365	297	68
P9	girl	13–18 mo	290	279	11
P10	boy	19–24 mo	113	111	2
P11	girl	19–24 mo	175	172	3
P12	boy	19–24 mo	143	137	6
P13	boy	25–36 mo	517	493	24
P14	girl	25–36 mo	103	91	12
P15	boy	25–36 mo	194	164	30
P16	girl	37–48 mo	147	84	63
P17	boy	37–48 mo	185	126	59
P18	boy	37–48 mo	170	138	32
Total			5,300	4,832	468

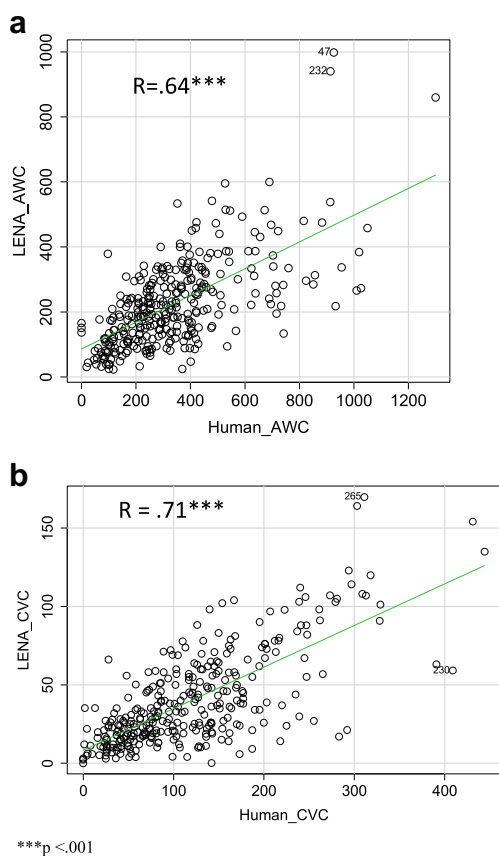


Fig. 2 Scatterplots between LENA and human adult word counts (AWCs, top) and child vocalization counts (CVCs, bottom) (324 samples). $***p < .001$

human AWC data ($r = -.22$, $p < .001$) indicating a certain variability concerning talking to young children.

Figure 4 displays scatterplots between the LENA and human AWC and CVC estimates by recording days. Both the LENA AWCs and CVCs were significantly correlated to the human AWCs and CVCs, ranging from .57 to .73 ($p < .001$) and from .57 to .80 ($p < .001$), respectively indicating that the LENA and human counts were reliable.

LENA versus human count estimates

LENA and human count estimates when adding the six age groups and the three recording days Because the datasets have a nested structure (6 ages \times 3 children in each \times 3 days \times 6 10-min chunks of recordings), two linear mixed models were constructed using the R statistical package (version 3.02; R Development Core Team, 2013). The LENA AWCs and CVCs were the dependent measures, and the six recordings and participants were the random factors (i.e., 108 samples). The fixed factors were a combination of age (six age groups) and of day of recording (three levels). The rationales for conducting such analyses were to obtain a more robust estimate per participant and to eliminate the problem of

nonindependence of observations. Overall, the linear mixed models show that 1 h of recording was sufficient to obtain a reliable sample. When examining the age groups, a main effect was found in the AWC data for the 7- to 12-month-old and the 13- to 18-month-old groups. When examining the recording day, a main effect was found in the CVC data. Analyses of deviance of Type II (Wald χ^2 test) between human and LENA count estimates confirm these two effects. The results of this analysis are shown in Table 4.

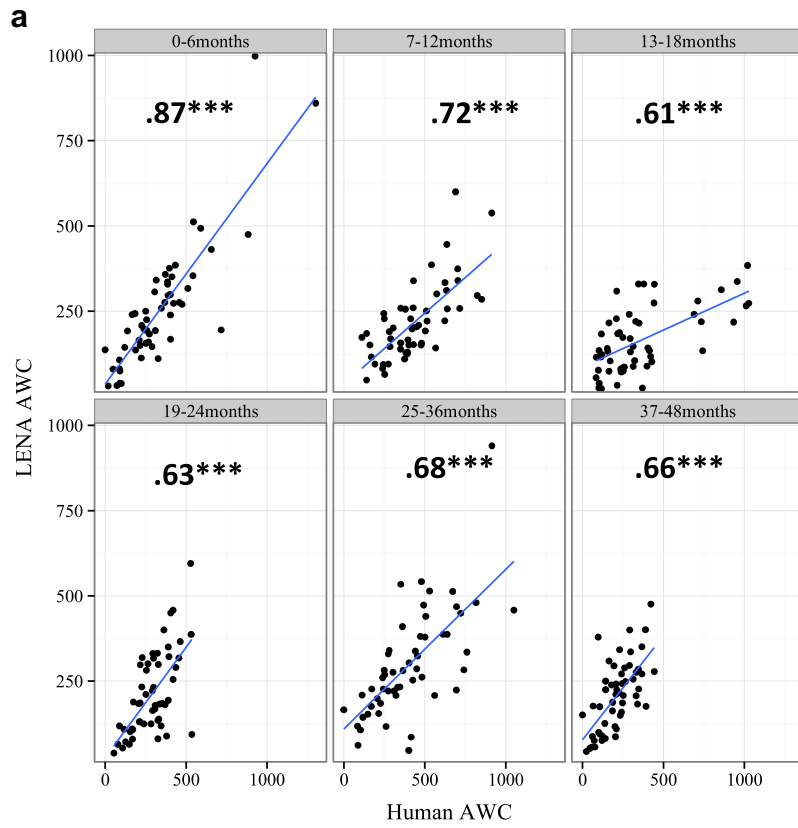
LENA and human counts estimates when adding age groups, recording days, SNR, NON, and OLN When the SNR, the NON (noise near), and the OLN (overlap near) factors were added to the linear mixed model, the coefficient correlations between the human and LENA counts remained significant for the AWC and CVC data, providing evidence of the reliability of the LENA system. Furthermore, a significant effect of SNR was found on the AWC data, indicating that the Distance factor had an impact on both the LENA and the human counts. Analyses of deviance of Type II (Wald χ^2 test) between the human and LENA count estimates confirmed all these effects. The results of this analysis are shown in Table 5.

LENA versus human counts in raw scores for the selected recording sample Figure 5 shows the bar plots between the LENA and human AWC and CVC estimates. The raw scores of the human AWC estimates were greater than the LENA AWC estimates (110.318 vs. 73.274 total words, the average ratio of the two estimates was 1.56). Similarly, the human CVC estimates were much greater than the LENA CVC estimates (38.409 vs. 12.881 total words, with an average ratio of 2.86).

Figure 6 display a series of bar plots comparing the ratio between the LENA and human count estimates for each participant. The human AWC was three times greater than the LENA AWC estimate in Participants 8 and 9. The human CVC estimate was eight times greater than the LENA CVC estimate in Participant 13. These results indicate a certain amount of variability between participants.

Discussion

The reliability of the LENA system was examined in European French in 18 children aged 3 to 48 months for the three full audio recording days. Such a validation is important because spoken French, a syllable-timed language, differs in many phonetic and acoustic features, relative to English. A total of 324 10-min chunks of recordings with their corresponding human transcriptions were analyzed, yielding the first validation of the accuracy of LENA in European French for both LENA adult word and child vocalization counts.



*** $p < .001$

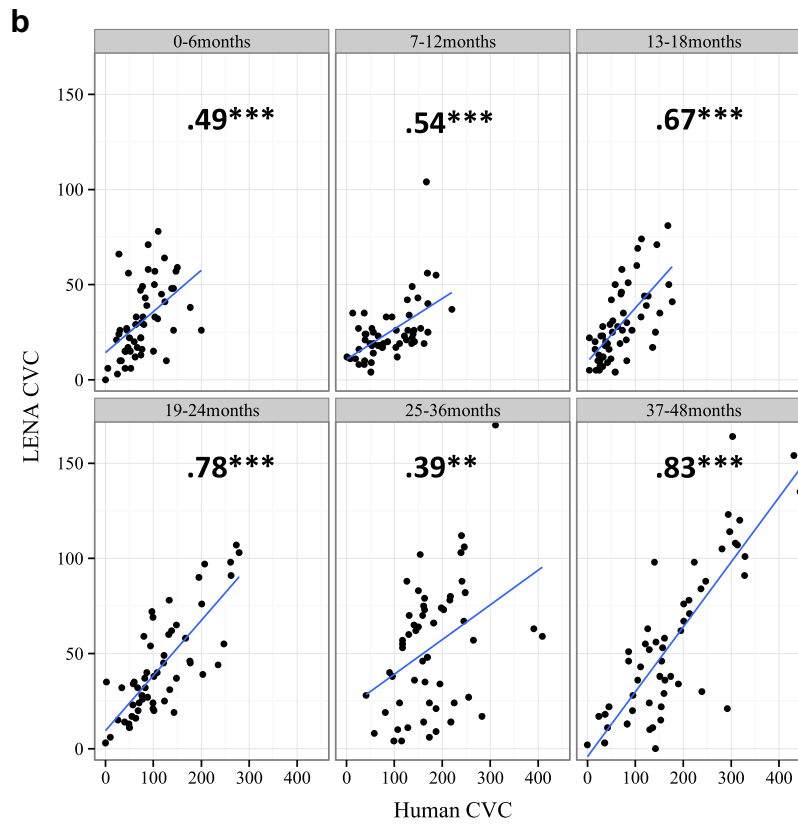


Fig. 3 Scatterplots between LENA and human AWCs and CVCs by age groups. *** $p < .001$

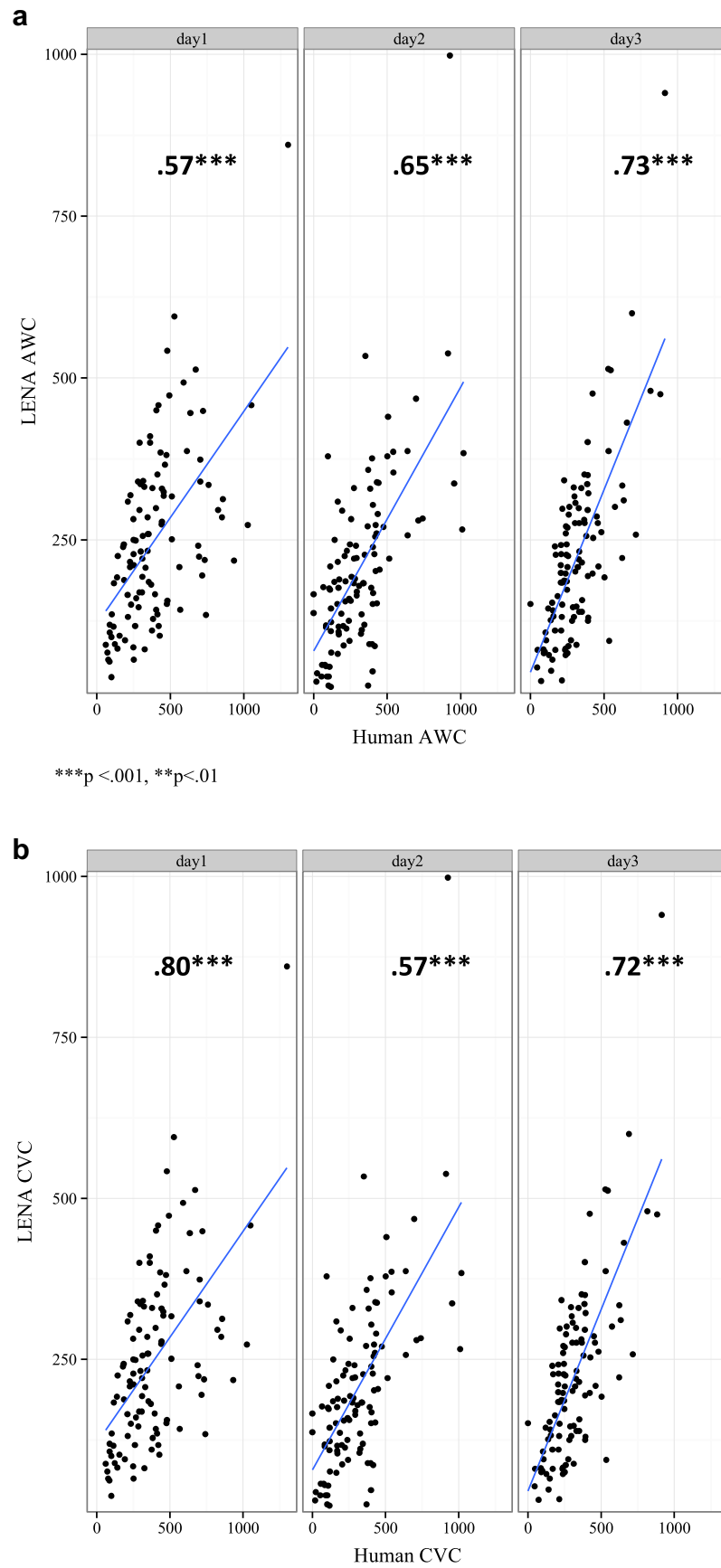


Fig. 4 Scatterplots between LENA and human AWCs and CVCs by recording days. *** $p < .001$, ** $p < .01$

Table 4 LENA and human count estimates when adding age groups and recording days

Fixed Effects	Estimate (β)	SE (β)	<i>t</i> Value ¹
(Intercept AWC)	85.18211	30.32964	2.809
Human	0.49293	0.02739	17.997***
Age 7–12 mo	−85.32887	38.953	−2.191*
Age 13–18 mo	−102.12669	38.8692	−2.627*
Age 19–24 mo	−18.17183	38.88826	−0.467
Age 25–36 mo	11.54772	38.90204	0.297
Age 37–48 mo	14.80594	39.01909	0.379
Recording of day2	−3.49249	12.30249	−0.284
Recording of day3	12.53313	12.29115	1.02
(Intercept CVC)	13.75435	8.49349	1.619
Human	0.25659	0.01614	15.893***
Age 13–18 mo	0.20674	11.69384	0.018
Age 19–24 mo	2.25038	11.70954	0.192
Age 25–36 mo	−3.4102	11.80271	−0.289
Age 37–48 mo	−0.28266	11.80496	−0.024
Age 7–12 mo	−9.17133	11.69475	−0.784
Recording of day2	−6.75675	2.35638	−2.867*
Recording of day3	−1.28804	2.35847	−0.546

Analysis of Deviance. Table Type II. Wald. χ^2 tests

Response: LENA AWC

Predictors	Chisq	Df	Pr(> χ^2)
Human	323.897	1	<.0001***
Age	17.117	5	.004284**
Day	1.934	2	.380218

Response: LENA CVC

Predictors	Chisq	Df	Pr(> χ^2)
Human	252.5963	1	<.0001***
Age	1.2123	5	.94369
Day	9.3004	2	.00956**

108 observations, 18 participants. ¹ Absolute *t* values of greater than 2 can be interpreted as indicating statistical significance (Baayen, Davidson, & Bates, 2008; Bates, Maechler, & Bolker, 2013).

Simple correlational analyses revealed a very good reliability in the selected chunks of recordings i.e., 324 samples. Overall, the correlations were .64 on the AWCs data and .71 on the CVCs data. This indicates that the LENA system does a fairly good job of estimating adult word productions and child vocalizations. This good reliability found in French between LENA and human count estimates is consistent with other reliability studies done in the English and Spanish languages, where the correlations between the two methods ranged from .71 to .85 on the AWCs data (Oetting et al., 2009; Weisleder & Fernald, 2013).

When controlling the random effects of participants and of recordings, 1 h of recording was found to be sufficient to obtain a reliable sample for both estimates. It is important to note that the LENA device was never intended for 1-h recordings, six 10-min chunks of recordings. There is a reason that 10–16 h of recordings serve as the basis for the statistical

Table 5 LENA and human count estimates when adding age groups, recording days, SNR, NON, and OLN

Fixed Effects	Estimate (β)	SE(β)	<i>t</i> Value
(Intercept AWC)	30.85408	34.13262	0.904
Human	0.53582	0.03059	17.514***
NON	0.1392	0.32387	0.43
OLN	−0.27942	0.22864	−1.222
SNR	4.88039	1.12263	4.347
Age 7–12 mo	−86.69945	42.08752	−2.060*
Age 13–18 mo	−90.47402	42.45134	−2.131*
Age 19–24 mo	−12.62928	41.85494	−0.302
Age 25–36 mo	18.95025	41.67831	0.455
Age 37–48 mo	27.60698	41.83007	0.66
recording of day2	1.04661	13.14954	0.08
recording of day3	12.97432	12.57776	1.032
(Intercept CVC)	6.71668	9.87251	0.68
Human	0.25079	0.01731	14.489***
NON	−0.08843	0.06647	−1.33
OLN	−0.06497	0.04778	−1.36
SNR	0.49433	0.26731	1.849
Age 7–12 mo	−7.87953	11.73761	−0.671
Age 13–18 mo	−0.80177	11.80145	−0.068
Age 19–24 mo	4.63633	11.74724	0.395
Age 25–36 mo	−1.92294	11.80407	−0.163
Age 37–48 mo	2.40328	11.82861	0.203
recording of day2	−5.95153	2.64007	−2.254*
recording of day3	−0.15612	2.55014	−0.061

Analysis of Deviance. Table Type II. Wald. χ^2 tests

Response: LENA AWC

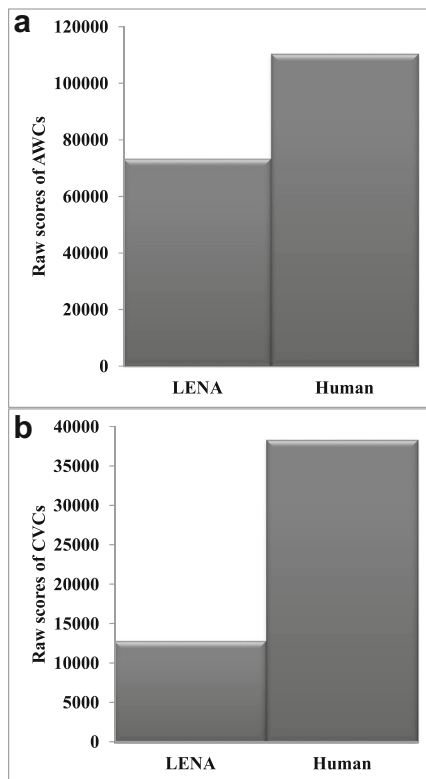
Predictors:	Chisq	Df	Pr(> χ^2)
Human	306.7395	1	<2.2e16***
NON	0.1847	1	.66734
OLN	1.4935	1	.22167
SNR	18.899	1	<.0001***
Age	15.0539	5	.01013
Day	1.3837	2	.50064

Response: LENA CVC

Predictors:	Chisq	Df	Pr(> χ^2)
Human	209.9174	1	<.0001***
NON	1.7698	1	.1834
OLN	1.849	1	.1739
SNR	3.4198	1	.06442
Age	1.3246	5	.93238
Day	6.8102	2	.0332

108 observations, 18 participants.

analyses. Therefore, if the relations between French human counts and the LENA automated counts for 1 h of 10–16 h of recording yields a good reliable sample, this relations should continue to strengthen with the greater volume of data. The fact that the human counting and the LENA automated



Panel a: Raw scores between LENA and Human AWCs and CVCs

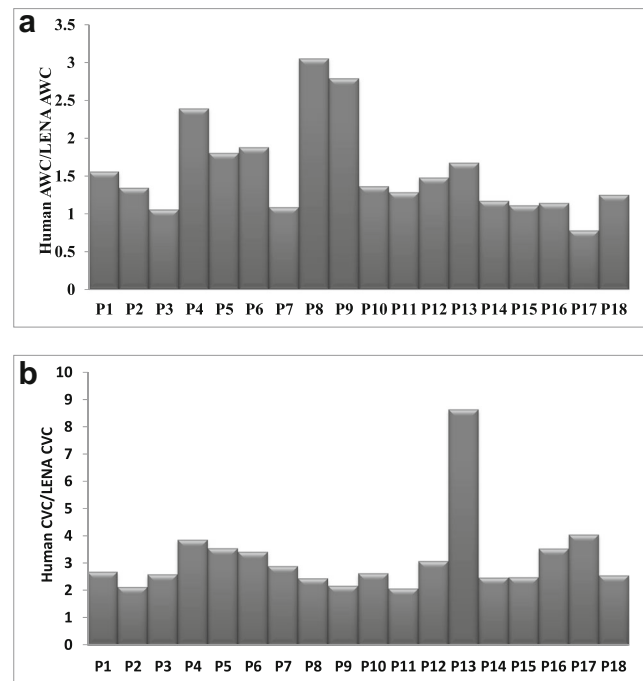
Panel b: Raw scores between LENA and Human AWCs and CVCs

Fig. 5 Bar plots between LENA and human AWCs and CVCs (324 samples). Top: Raw scores between LENA and human AWCs and CVCs. Bottom: Raw scores between LENA and human AWCs and CVCs

analysis algorithms remain reliable encourages the use of the LENA system in French for tracking the sequential skill and the developmental changes in learning to talk (vocal play, babble, first words, expressive jargon, intonational sentences, and word combinations).

An effect of two of the age groups (7–12 and 13–18 months) was, however, found on the AWC data, and an effect of the second day of recording was found on the CVC data. One reason to get such different patterns of results in these two estimates might be the lack of sufficient data points for age within each participant: only three children by age group were examined in this French study. Another reason might be the contextual factors that could influence the amount of speech heard and vocalizations produced by a young child under naturalistic conditions. This is consistent with Soderstrom and Wittebolle (2013) who found significant effects of both activity and time of day on the LENA and the human AWC and CVC data.

With regard to the effect related to noise-related factors, the results of the linear mixed model showed that LENA-based prediction was not affected by OLN (overlap near) and by NON but rather by SNR on the LENA AWC data. The impact of SNR found in this study is not surprising because it is



The ratios (Human AWC/LENA AWC) per participant are presented in the top bar chart, while the ratios (Human CVC/LENA CVC) per participant are presented in the bottom bar chart.

Fig. 6 Bar plots between LENA and human AWCs and CVCs per participant. The ratios (human AWC/LENA AWC) per participant are presented in the top bar chart, whereas the ratios (human CVC/LENA CVC) per participant are presented in the bottom bar chart

challenging to isolate adult words from noisy environment. These results are also consistent with previous LENA studies showing that the LENA system and human-transcribed counts essentially deviate during chunks of recording containing substantial noise (Xu, Yapanel, & Gray, 2009).

Although overall average difference ratio per participant was 1.5 on the AWC data and 2.8 on the CVC data, all coefficient correlations remained significant providing strong evidence of using LENA in European French. The differences found between LENA and human counts were particularly great in P8, P9 and P13. The average difference ratio between the LENA and human AWCs were found for Participants 8 and 9, 13 months of age, and in LENA and human CVCs for Participant 13, 33 months of age. Listening carefully to all of the 10-min chunks for these three participants, it is shown that their recordings were done in a very noisy environment, mostly at the daycare center, where overlap sounds, external conversations, and background noises were predominant. Twelve recordings of Participants 8 and 9 during the first day and the second day were in a noisy environment, whereas the other six other sessions during the third day were recorded in a quiet home, explaining the smaller differences between the LENA and human counts. When listening to the recording of Participant 13, three factors appeared to be involved: (i) outdoor recording session, (ii) overlapping speech segments, and (iii) clothing noise.

It is important to note that the DLP recording device cannot capture language productions outside an approximate 6-foot radius from the key child (Warren et al., 2010). In some situations (e.g., bath time, hot weather conditions), the DLP device could not be attached to the child's clothing as recommended, making voice identification of the key child *versus* other vocal categories more problematic for the LENA algorithms whereas this remained still easy for the human coders. To give an example, we noticed that, during an outdoor recording session, LENA algorithms miscategorized the chirping of birds as a female adult voice, but not the human coders. Human and LENA AWCs were likewise reported whenever the child was involved in outdoor activities, with the LENA AWCs again following the human-identified estimates once quieter activity resumed; this accounted for an average error rate of 27 % over a 12-h recording session (Xu, Yapanel, & Gray, 2009). According to these researchers, the effects of reverberation and echo resulting from environmental location, its acoustic characteristics (room size, flooring, etc.) and far-field effects appear to be the main factors that distort the performance of the system. Last but not least, LENA measures of language environment during outdoor activities often have to deal with noisy recording conditions. When a linguistic message is delivered in a noisy environment, human listeners are capable of extracting its content from the background noise. Outdoor recordings are highly susceptible to errors in counting. This is in total agreement with the findings of the LENA system devisers, who extensively studied the impact of noisy recording conditions on data speech quality (Xu, Yapanel, & Gray, 2009).

In the case of overlapping speech, humans are able to separate sources, counting all of the intelligible words, but the LENA system counts none. Whereas the automated LENA speech-processing algorithms systematically eliminate overlapping speech segments, the human listener is able to process context-related information associated to each sound, with far more vocal events and human talk being identified. The number of overlapping speech segments detected by the machine will proportionally increase the differences in counts observed between the LENA and the human transcriptions. For instance, Gilkerson & Richards (2009) reported that word counting within the same typical family can vary by more than 50 % around the mean, due primarily to the presence of multiple overlapping speech segments. When the adult speakers move a lot, this makes transcriptions even more difficult because of the difficulty for the human coder to assess the distance of the speech productions. The transcribers' counts are based on intelligibility rated by qualitative perceptual judgments, whereas the LENA system relies on an automated signal-processing algorithm. This can explain the discrepancy we found between LENA and human counts. Whereas the LENA system systematically labels AWCs and CVCs in the two categories FAR and NEAR, this annotation is much more difficult and

more unpredictable for the human listener than for the LENA system. This observation supports Xu, Yapanel, and Gray's (2009) claims that the AWCs are influenced by the speech quality introduced by different speakers in audio files—in particular, overlap sounds, external conversations, and background noise.

LENA studies on reliability have several limitations in all languages: First and foremost, the AWC and CVC variables provide only a count of child vocalizations or adult words, but information on the type or quality of conversation is not captured. Although the AWC and CVC variables provide an accurate representation of adult or child words, they may underestimate the content words, which are a valuable component of language development. Second, the LENA system has a 6-foot radius in which it captures audio data; therefore, it is possible that adults were vocalizing near the children, but not necessarily directing words toward them. Lastly, LENA provides purely naturalistic audio data for statistical model training. In order to make more meaningful claims about the adult output directed at children and vice versa additional observational measures are needed (e.g., video). Despite these limitations, the present study does provide meaningful reliable information about human and LENA counts and has important research implications on child language development and disorders.

Research implications of LENA on child language development and disorders

LENA can be used for tracking child language development and language disorders: Many studies have shown that the linguistic environment and social interactions influence language acquisition and development (Braine, 1994; Kuhl, 2011; Rowe, 2012; Snow, 1994). Several aspects of language input have been found to predict outcomes: quantity (Hart & Risley, 1995; Huttenlocher et al., 1991; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010), quality (Cartmill et al. 2013; Pan, Rowe, Singer, & Snow, 2005; Pan, Rowe, Spier, & Tamis-LeMonda, 2004), lexico-syntactic diversity (Huttenlocher et al., 2010), word frequency (Weizman & Snow, 2001), and decontextualized language (Snow, Tabors, & Dickinson, 2001). Recent studies conducted with the LENA system have led to similar conclusions. The longitudinal results of Hart and Risley (1995) on the relations between the language addressed to children from an early age and their later academic achievement have been replicated in 30 English-speaking families using LENA (Greenwood, Thiemann-Bourque, Walker, Buzhardt, & Gilkerson, 2011). Zimmerman et al. (2009) showed that interactions have the most positive impact on child language development, whereas Christakis et al. (2009) pointed to the negative impact of TV exposure. Oller (2010) confirmed in another LENA study that language directly addressed to the child has a greater impact

on lexical development than when it is only overheard: A child exposed to three different languages and followed between 11 and 24 months of age was found to learn and use new words mostly in the language that was most often spoken to him.

The use of LENA is also growing in the field of clinical research. It is well-known that degraded sensory inputs will negatively impact oral language development in children with hearing loss (Briscoe, Bishop, & Norbury, 2001; Stoel-Gammon, 1988; Yoshinaga-Itano, Sedey, Coulter, & Mehl, 1998), whereas appropriate auditory stimulations will play a major role in both phonological and lexical development (Desjardin, Ambrose, & Eisenberg, 2009; Farran, Ledesberg, & Jackson, 2009). Recent studies using LENA also point in this direction. The studies by VanDam, Ambrose, and Moeller (2012), VanDam and Silbert (2013a, b), and Vohr, Topol, Watson, St. Pierre, and Tucker (2014) showed, respectively, that the number of interactions and the richness of the linguistic environment will help children to develop language skills in particular at the receptive level.

This study forecasts extensions for further cross-linguistic generalization of an automatic assessment of child–caregiver interactions to a much broader range of populations. Combining LENA data from various contexts across the day with research on the importance of a high-quality preschool language environment will allow researchers, practitioners and other stakeholders to advance professional development efforts and optimize interventions for clinical populations.

Conclusion

LENA offers a reliable and efficient method for collecting data related to language development and the language environment. This reliability study offers a starting point for describing the language environment of French-speaking children and determining how word counts could be used to assist researchers and practitioners to provide optimal assessments or interventions. Using the LENA system for research in natural settings will contribute to a deeper knowledge of the language environment in clinical settings.

Author note The authors thank François Pellegrino for providing valuable help with the SNR analysis, and D. Kimbrough Oller and A. Cristia for their constructive comments on an earlier version of the manuscript. They are also grateful to Sébastien Flavier for his technical assistance and Florence Loisel for her help with data collection.

References

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh, UK: Edinburgh University Press.
- Adda-Decker, M., Gendrot, C., & Nguyen, N. (2008). Contributions du traitement automatique de la parole à l'étude des voyelles orales du français. *Revue Traitement Automatique des Langues*, 49, 1–34.
- Aragon, M., & Yoshinaga-Itano, C. (2012). Using Language ENvironment Analysis to improve outcomes for children who are deaf or hard of hearing. *Seminars in Speech and Language*, 33, 340–353. doi:10.1055/s-0032-1326918
- Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossedrandom effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Bates, D., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using S4 classes (R package version 1.0-4) [Software]. Retrieved from <http://CRAN.R-project.org/package=lme4>
- Bolinger, D. (1985). *Intonation and its parts*. London, UK: Edward Arnold.
- Braine, M. D. S. (1994). Is nativism sufficient? *Journal of Child Language*, 21, 9–32.
- Briscoe, J., Bishop, D. V., & Norbury, C. F. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 42, 329–340.
- Burgess, S., Audet, L., & Harjusola-Webb, S. (2013). Quantitative and qualitative characteristics of the school and home language environments of preschool-aged children with ASD. *Journal of Communication Disorders*, 46, 428–439. doi:10.1016/j.jcomdis.2013.09.003
- Canault, M., & Thai-Van, H. (2013, April). *Validation of the Language ENvironment Analysis System in French: Data quality and associated factors*. Paper presented at LENA International Conference, Denver, CO.
- Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences of United States of America*, 100, 11278–11283. doi:10.1073/pnas.1309518110
- Caskey, M., Stephens, B., Tucker, R., & Vohr, B. (2011). Importance of parent talk on the development of preterm infant vocalizations. *Pediatrics*, 128, 910–916. doi:10.1542/peds.2011-0609
- Caskey, M., & Vohr, B. (2013). Assessing language and language environment of high-risk infants and children: A new approach. *Acta Paediatrica*, 102, 451–461. doi:10.1111/apa.12195
- Christakis, D. A., Gilkerson, J., Richards, J. A., Zimmerman, F. J., Garrison, M. M., Xu, D., & Yapanel, U. (2009). Audible television and decreased adult words, infant vocalizations, and conversational turns: A population-based study. *Archives of Pediatrics and Adolescent Medicine*, 163, 554–558. doi:10.1001/archpediatrics.2009.61
- Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2, 133–142. doi:10.1016/0885-230890004-0
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 113–121. doi:10.1037/0096-1523.14.1.113
- Delattre, P. (1962). Comparing the prosodic features of English, French, German and Spanish. *International Review of Applied Linguistics*, 1, 193–210.
- Desjardin, J. L., Ambrose, S. E., & Eisenberg, L. S. (2009). Literacy skills in children with cochlear implants: The importance of early oral language and joint storybook reading. *Journal of Deaf Studies and Deaf Education*, 14, 22–43.
- Farran, L. K., Ledesberg, A. R., & Jackson, L. A. (2009). Maternal input and lexical development: The case of deaf pre-schoolers. *International Journal of Language and Communication Disorders*, 44, 145–153. doi:10.1080/13682820801973404
- Gilkerson, J., & Richards, J. A. (2008). The LENA natural language study (Technical Report LTR-02-2). Boulder, CO: LENA

- Foundation. Retrieved from www.lenafoundation.org/wp-content/uploads/2014/10/LTR-02-2_Natural_Language_Study.pdf
- Gilkerson, J., & Richards, J. A. (2009). *The power of talk: Impact of adult talk, conversational turns, and TV during the critical 0–4 years of child development* (Technical Report LTR-01-2, 2nd ed.). Boulder, CO: LENA Foundation. Retrieved from www.lenafoundation.org/wp-content/uploads/2014/10/LTR-01-2_PowerOfTalk.pdf
- Gilkerson, J., Zhang, Y., Xu, D., Richards, J. A., Xu, X., Jiang, F., . . . Topping, K. (2015). Evaluating Language Environment Analysis system performance for Chinese: A pilot study in Shanghai. *Journal of Speech Language and Hearing Research, 58*, 445–452. doi:10.1044/2015_JSLHR-L-14-0014
- Greenwood, C. R., Thiemann-Bourque, K., Walker, D., Buzhardt, J., & Gilkerson, J. (2011). Assessing children's home language environments using automatic speech recognition technology. *Communication Disorders Quarterly, 32*, 83–92. doi:10.1177/1525740110367826
- Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236–248. doi:10.1037/0012-1649.27.2.236
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology, 61*, 343–365. doi:10.1016/j.cogpsych.2010.08.002
- Jackson, C. W., & Callender, M. F. (2013, April). *A comparison of reported language use and observed percentage of Spanish-English use as measured by LENA recordings of young ELLs*. Paper presented at LENA International Conference, Denver, CO.
- Kuhl, P. K. (2011). Who's talking. *Science, 333*, 529–530. doi:10.1126/science.1210277
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.
- Oetting, J. B., Hartfield, L. R., & Pruitt, S. L. (2009). Exploring LENA as a tool for researchers and clinicians. *ASHA Leader, 14*, 20–22.
- Oller, D. K. (2010). All-day recordings to investigate vocabulary development: A case study of a trilingual toddler. *Communication Disorders Quarterly, 31*, 213–222. doi:10.1177/1525740109358628
- Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., . . . Warren, F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences, 107*, 13354–13359.
- Pae, S. (2013, April). *An investigation of the language environment of young Korean children and the impact of feedback on language development*. Paper presented at LENA International Conference, Denver, CO.
- Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development, 76*, 763–782. doi:10.1111/j.1467-8624.2005.00876.x
- Pan, B. A., Rowe, M. L., Spier, E., & Tamis-LeMonda, C. (2004). Measuring productive vocabulary of toddlers in low-income families: Concurrent and predictive validity of three sources of data. *Journal of Child Language, 31*, 587–608.
- Pellegrino, F., Coupé, C., & Marsico, E. (2011). Across language perspective on speech information rate. *Language, 87*, 539–558. doi:10.1353/lan.2011.0057
- R Development Core Team. (2013). R: A language and environment for statistical computing (Version 3.02). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org
- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development: Child-directed speech and vocabulary. *Child Development, 83*, 1762–1774. doi:10.1111/j.1467-8624.2012.01805.x
- Sacks, C., Shay, S., Repplinger, L., Leffel, K. R., Sapolich, S. G., Suskind, E., . . . Suskind, D. (2014). Pilot testing of a parent-directed intervention (project ASPIRE) for undeserved children who are deaf or hard of hearing. *Child Language Teaching and Therapy, 30*, 91–102.
- Siller, M., & Sigman, M. (2002). The behaviors of parents of children with autism predict the subsequent development of their children's communication. *Journal of Autism and Developmental Disorders, 32*, 77–89.
- Snow, C. E. (1994). Beginning from baby talk: Twenty years of research on input and interaction. In C. Gallawa & B. Richards (Eds.), *Input and interaction in language acquisition* (pp. 3–12). Cambridge, UK: Cambridge University Press.
- Snow, C. E., Tabors, P. O., & Dickinson, D. K. (2001). Language development in the preschool years. In D. K. Dickinson & T. O. Tabors (Eds.), *Beginning literacy with language: Young children learning at home and school* (pp. 1–26). Baltimore, MD: Brookes.
- Soderstrom, M., & Wittebolle, K. (2013). When do caregivers talk? The influences of activity and time of day on caregiver speech and child vocalizations in two childcare environments. *PLoS ONE, 8*, e80646. doi:10.1371/journal.pone.0080646
- Stoel-Gammon, C. (1988). Prelinguistic vocalizations of hearing-impaired and normally hearing subjects: A comparison of consonantal inventories. *Journal of Speech and Hearing Disorders, 53*, 302–315.
- Vaissière, J. (2010). Le français, langue à frontières par excellence. In D. Delomier & M. A. Morel (Eds.), *Frontières du linguistique au sémiotique* (pp. 39–62). Limoges, France: Lambert Lucas.
- VanDam, M., Ambrose, S. E., & Moeller, M. P. (2012). Quantity of parental language in the home environments of hard-of-hearing 2-year-olds. *Journal of Deaf Studies and Deaf Education, 17*, 402–420. doi:10.1093/deafed/ens025
- VanDam, M., & Silbert, N. H. (2013a). Precision and error of automatic Speech recognition. *Journal of Acoustical Society of America, 133*, 3245.
- VanDam, M., & Silbert, N. H. (2013b). Precision and error of automatic speech recognition. *Proceedings of Meetings on Acoustics, 19*, 060006. doi:10.1121/1.4798466
- Vohr, B. R., Topol, D., Watson, V., St. Pierre, L., & Tucker, R. (2014). The importance of language in the home for school-age children with permanent hearing loss. *Acta Paediatrica, 103*, 62–69. doi:10.1111/apa.12441
- Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science, 25*, 1314–1324. doi:10.1177/0956797614531023
- Warren, S. F., Gilkerson, J., Richards, J. A., Oller, D. K., Xu, D., Yapanel, U., & Gray, S. (2010). What automated vocal analysis reveals about the vocal production and language learning environment of young children with autism. *Journal of Autism and Developmental Disorders, 40*, 555–569. doi:10.1007/s10803-009-0902-5
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 24*, 2143–2152. doi:10.1177/0956797613488145
- Weizman, Z. O., & Snow, C. E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology, 37*, 265–279.
- Wenk, B. J., & Wioland, F. (1982). Is French really syllable-timed. *Journal of Phonetics, 10*, 193–216.
- Xu, D., Gilkerson, J., Richards, J., Yapanel, U., & Gray, S. (2009). Child vocalization composition as discriminant information for automatic autism detection. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology*

- Society* (pp. 2518–2522). Piscataway, NJ: IEEE Press. doi:[10.1109/IEMBS.2009.5334846](https://doi.org/10.1109/IEMBS.2009.5334846)
- Xu, D., Yapanel, U., Gray, S., Gilkerson, J., Richards, J. A., & Hansen, J. (2008, October). *Signal processing for young child speech language development*. Paper presented at the First Workshop on Child, Computer and Interaction, Chania, Crete, Greece. Retrieved from www.lenafoundation.org/DownloadFile.aspx/pdf/SignalProcessing_ChildSpeech
- Xu, D., Yapanel, U., & Gray, S. (2009). *Reliability of the LENA Language Environment Analysis System in young children's natural home environment* (Technical Report LTR-05-2). Boulder, CO: LENA Foundation. Retrieved 6 December, 2013, from www.lenafoundation.org/TechReport.aspx/Reliability/LTR-05-2
- Yoder, P. J., Oller, K., Richards, J. A., Gray, S., & Gilkerson, J. (2013). Stability and validity of an automated measure of vocal development from day-long samples in children with and without autism spectrum disorder. *Autism Research*, *6*, 103–107.
- Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998). Language of early- and later-identified children with hearing loss. *Pediatrics*, *102*, 1161–1171.
- Zhang, Y. (2013, April). *The correlation of family responsiveness with language and cognitive development in infants and toddlers*. Paper presented at the LENA International Conference, Denver, CO.
- Zimmerman, F. J., Gilkerson, J., Richards, J. A., Christakis, D. A., Xu, D., Gray, S., & Yapanel, U. (2009). Teaching by listening: The importance of adult-child conversations to language development. *Pediatrics*, *124*, 342–349. doi:[10.1542/peds.2008-2267](https://doi.org/10.1542/peds.2008-2267)