Canonical Babbling Ratio Extracted from Day-long Audio Recordings: A preliminary report from India

Shoba S. Meera[1], Divya Swaminathan[1], Sri Ranjani V[2], Malavi Srikar[1], Reny Raju[1]

[1]Department of Speech Pathology and Audiology, National Institute of Mental Health and Neurosciences, Bangalore, India, [2]Department of Speech-Language Pathology, Sri Ramachandra Institute of Higher Education and Research, Chennai, India

[1] ssmeera@nimhans.ac.in ssmeeras@gmail.com

ABSTRACT

Presence of canonical babbling (CB) is a crucial vocal developmental milestone in the later part of an infant's first year of life. CB can be measured as a ratio (CBR) – the proportion of canonical syllables to other syllables produced. In this study we present CBR data and consonant inventory from thirteen 6–10-month-old typically developing infants from a culturally and linguistically diverse background. CBRs were computed based on speech samples collected using day-long audio recorders (Language ENvironment Analysis) – a technology being used and reported for the first time in India. All infants above 8 months of age had a CBR of >0.15 indicating they were in the CB stage. Further, all infants in the CB stage produced bilabial sounds, which are some of the earliest consonants present in an infant's consonant inventory. These findings are in-line with literature reported globally, thus generalizing our understanding to previously understudied populations.

Keywords: Canonical babbling ratio, day-long audio recording, infants, Indian languages

1. INTRODUCTION

A critical milestone in an infant's vocal development in the first year of life is the use of adult-like consonant-vowel (CV) combinations with a rapid transition between the CV (e.g., /ba/, /du/)[1], [2]. These CV combinations are called canonical syllables. When canonical syllables are produced together in a repetitive sequence it is called canonical babbling (e.g., /baba/, /dududu/). One way to measure CB is to compute a canonical babbling ratio (CBR). CBR is the proportion of canonical syllables to other syllables produced. A CBR of 0.15 and above has been used to determine if an infant is in the CB stage[3]. Infants produce CB before producing words. CB in typically developing infants is well documented and is present between 6-10 months of age[1], [4], [5]. Some of the earliest sounds infants produce are bilabials and hence CB too comprises bilabial consonant-vowel combinations like /bababa/[6]–[8]. CB and CBR have been shown to predict language outcomes in typically developing infants[3].

Although a perceived notion is that emergence of canonical syllables is independent of the languages children are exposed to, and perhaps the cultures they are brought up in, studies have been primarily conducted in high-income countries (HICs) with largely monolingual samples (see discussion in Cychosz et al., 2021[5]). This bias in our field not only creates serious issues in terms of coverage of human languages and generalizability across human populations (as discussed for instance in Kidd & Garcia, 2022[9]) but also represents an unfair situation, since a majority of the world's infants are growing up in low- and middle-income countries (LMICs). India alone accounts for over 20% of the world's children (https://data.unicef.org/how-many/how-many-children-under-18-are-there-in-india/).

Only recently a paper reported vocal development including canonical syllables from across cultures including some data from LMICs [5]. The study found that canonical syllables were present in most infants by 7 months of age. Further, most infants had a CBR of 0.15 and above by 10 months of age. Although the study measured vocal development in five culturally diverse settings, it did not include a sample of infants from a LMIC that is naturally bi/multilingual, like India. In fact, only a handful among included infants who had exposure to a second language, and no analyses were made to check for potential differences. More broadly, to our knowledge, no previous study has investigated CB development in a naturally bi/multilingual Indian sample, a gap we address in this work.

Most studies that have evaluated early vocal development, including canonical babbling, have used traditional methods like short lab-based[10], [11] recordings or short home recordings[12], [13]. However, these methods do not represent the naturalistic environment of the child, and concern has

been raised about the extent to which observations from short recordings generalize to children's behavior and experiences [14]. A more contemporary approach to measuring CB is using technology that records the entire waking hours of the child e.g., LENA™ (Language ENvironment Analysis)[15], providing a more naturalistic sample of an infant's speech production. These recordings are called day-long audio recordings. Only a few recent studies have used day-long audio recordings to capture and analyse CB[3], [5]. Such technology is relatively new and has seldom been used in LMICs, including India. Also, no such previous study investigated the consonants present in infants' production.

Thus, the current study aimed to add to existing literature on CBR from a culturally and linguistically diverse sample of 6–10-month-old typically developing infants using day-long audio recordings. The study also assessed the consonant inventory that was present in these infants.

2. METHOD

Data from thirteen 6–10-month-old typically developing infants were included in the present study.[i] This sample size is similar to sample sizes found for individual corpora (for instance in Cychosz et al., 2019[5]). Parents provided written consent to participation. The study included infants exposed to one or a combination of the following Indian languages - Hindi, Kannada, Tamil, Malayalam, Telugu, or Indian English.[ii] All infants were screened on two developmental behavioural assessments: (a) The ComDEALL Developmental Checklist (CDDC)[16] which screens for delays in domains such as motor, activities of daily living, language, cognition, social, and emotional skills and (b) The Communication and Symbolic Behavior Scales Developmental Profile – Infant-Toddler Checklist (CSBS – ITC)[17], which rules out potential communication and language delay. Infants were excluded from the study if there was evidence of a genetic condition or syndrome, significant medical or neurological condition affecting development, significant vision or hearing impairment, birth weight <2500 g or gestational age <36 weeks or perinatal brain injury secondary to birth complications.

Canonical babbling was evaluated based on speech samples collected using the LENA digital language processors (DLP)/recorders. Once an infant was recruited to the study, the participating families were given a study kit containing (a) the recorder, (b) a specially designed vest with a front pocket to secure the recorder and, (c) an instruction sheet on how to use the recorder. The researcher also provided instructions on how to use the recorder in-person and over a video call just prior to the start of recording. Families were asked to complete one day of recording when all members of the family were at home and no social events were planned.

2.1. Generating high child vocalization sample to annotate CB

Upon receiving the LENA-DLP/recorder from the family, the recording was uploaded on the LENA Pro 3.4.1 software. The LENA ADEX (Advanced Data EXtractor)[18] program was used to download a .csv file of the 16-hour recording. This file was used to identify three 5-minute segments that contained highest number of child vocalization segments and similarly, three 5-minute segments that contained highest number of adult word segments.[iii] Audio/speech samples from these six segments were compiled into one wave file (total of 30-minutes) that was then annotated using a study specific Graphical User Interface (GUI) developed in Python (Figure-1)

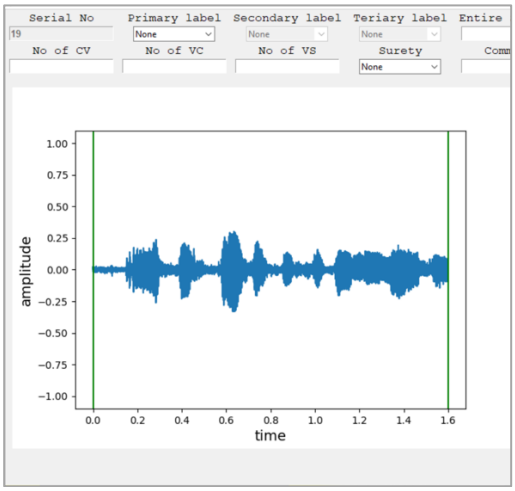

Figure 1: Graphical User Interface (GUI) to annotate CB

2.2. Annotation process:

Two annotators (speech-language pathologists) were trained to tag segments into various primary labels e.g., CH= child segment, FAW= female adult segment, NOIS= noise, etc. Upon successful training (i.e., maintaining an IRR above 0.80 on two successive practice recordings), the annotators annotated a total of 16,260 segments (~390 minutes) in random order.

Inter-rater reliability, Cohen's kappa (k), for CH segments was 0.80 (strong agreement)[19]. Each primary label that was annotated as CH was further annotated into secondary labels of speech-like or non-speech categories. Speech-like category included

vocalizations characterized by the production of consonants and/or vowels. Non-speech included infant cries, raspberries, or squeals. If the secondary label was speech-like, the segment was transcribed. These transcriptions were used to compute CBRs and to describe the consonant inventory. A consonant was considered to be part of the inventory even if it was produced once in the 30-minute sample that was annotated.

3. RESULTS

Data from 13 infants (n=6 females) with a mean age of 8.19 months were analysed. The sample consisted of n=6 (46.15%) monolinguals, n=3 (23.07%) bilinguals and n= 4 (30.76%) multilinguals. Some infants were exposed to languages that were not part of the inclusion criteria (e.g., Marathi). However, these languages were less spoken in the house compared to the primary languages (e.g., Hindi, Kannada). All infants were at, or above par, in their developmental abilities on the CDDC. None of the infants failed the CSBS-ITC indicating no potential risk for language or communication delay. Further, parents reported no concerns in their infant's development. Demographic data of participants and their families are presented in table 1.

Characteristic	Value
	Mean (SD)
Infant age in months	8.19 (1.44)
Maternal age in years	29.92 (2.59)
Paternal age in years	32.46 (1.89)
	n (%)
Infant Sex	
Females	6 (46.15)
Males	7 (53.85)
Maternal Education	
Graduate or post-grad	9 (69.23)
Professional Degree	4 (30.76)
Paternal Education	
Graduate or post-grad	9 (69.23)
Professional Degree	4 (30.76)
Socio Economic Status	
Upper	3 (23.07)
Upper middle	10 (76.92)
Residential area: Urban	13 (100)

Table 1: Demographic data of infant participants and their families

Prior to analysing the number of infants in the CB stage, we ran a Spearman's correlation to determine correlation between age of the infant and total child vocalizations ($r=0.78$, $p<0.001$), and infant age and CBRs ($r=0.64$, $p=0.017$). There was a positive correlation in both cases. See Figure-2

We further found that 10 infants (76.92%) had a CBR of > 0.15 i.e., they were in the CB stage. The three infants (n=1 6mo/female, n=2 7mo/male) who were not yet in the CB stage, had several speech-like vocalizations but no CV combinations. The consonant inventory identified in the 10 infants who had a CBR of >0.15 is presented in Table 2.

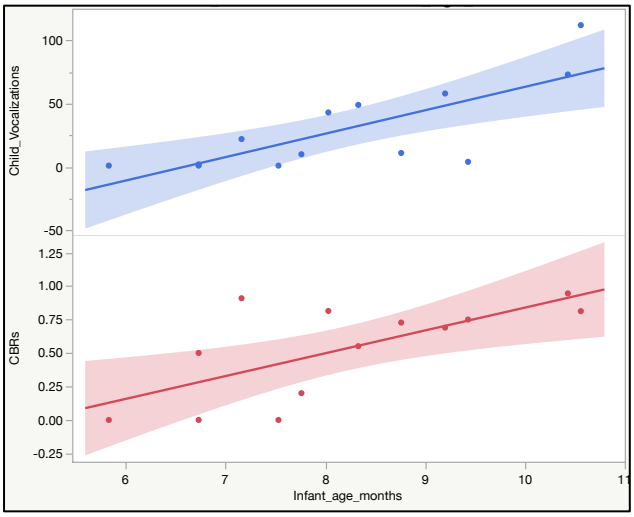

Figure 2: Correlation between infant age and CBRs and child vocalizations

We further found that 10 infants who were in the canonical babbling stage produced bilabial sounds (/m/ as in "mat" 7/10, /b/ as in "bat" 7/10, /p/ as in "pat" 4/10 infants). Other consonants that were present include /j/ as in "yes" (5/10), /n/ as in "no" (4/10), /t/ as in "tap" (4/10), /g/ as in "go" (4/10), /d/ as in "dog" (3/10), /w/ as in "water" (2/10), /h/ as in "hat" (2/10), /k/ as in "king" (2/10), /v/ as in "van" (3/10), /l/ as in "lamp" (1/10).

4. DISCUSSION

The current study from a culturally and linguistically diverse sample from India adds evidence to existing literature on CB – a huge population that is understudied. There was a positive correlation in total child vocalizations and CBRs with infants' age indicating overall increased vocal production with age like previously reported literature[20], [21]. All infants 8 mo and older had a CBR of >0.15 indicating they were in the CB stage. This finding is in line with several other studies that have evaluated CBR to determine the presence of CB stage e.g., Cychosz et al, 2021 [5].

Despite potential concerns for non-generalizability of previous results, this study also converged with previous ones in demonstrating that the first sounds infants produce are bilabial sounds[6]–[8]

CBR was computed from speech samples collected through recorders that collect day-long audio recordings thus providing a more naturalistic sample that accounts for fluctuations in vocalizations within the recording day. Use of day-long audio recorders (LENA) was new for participating families. They asked several questions about the recorder, potential harmful electronic radiation, data storage, etc. Eventually, all families that were approached provided written consent for participation and completed 16 hours of recording.

ID	Age	Language(s)	Consonant inventory
7	7 mo	Kannada	/m/
18	7 mo	Hindi English	/m/, /b/, /p/
9	8 mo	Kannada	/b/, /j/
3	8 mo	Telugu English	/m/, /n/, /b/, /h/ /t/, /k/, /g/, /j/
10	8 mo	Kannada Hindi Marathi English	/m/, /n/, /l/, /g/, /j/
15	9 mo	Telugu	/m/, /t/, /v/
5	9 mo	Hindi Gujarati English	/b/, /w/, /d/, /p/, /h/, /t/, /k/ /g/, /j/
21	9 mo	Hindi Kannada English	/b/
22	10 mo	Malayalam	/m/, /n/, /b/, /d/, /p/, /t/, /g/, /v/, /j/
30	10 mo	Telugu Tamil English	/m/, /n/, /b/, /w/, /d/, /p/, /v/

Table 2: Consonant inventory

These data could be useful for other endeavors: A total of 16,260 segments (~390 minutes) were annotated by two independent rates thus developing an initial pool of well annotated segments that can be used for future studies – e.g., building algorithms to automatically count number of canonical syllables.

In the future, we hope to include a larger sample of infants, more homogeneity within mono/bi/multilingual participants, and representation from different socio-economic strata. An approach that involves citizen scientists [5] may be beneficial for studies such as these, which are labour intensive that make it particularly harder for researchers from LMICs to contribute to literature since resources are limited. That said, there is at present no alternative to careful laboratory annotations to assess infants' consonant inventory, which is one of the strengths of our contribution. More broadly, we hope more diverse researchers contribute similar data, so that together we can start seeing similarities and differences across children as a function of the language(s) they are exposed to and their sociocultural background.

5. REFERENCES

[1] D. K. Oller, "The emergence of the sounds of speech in infancy," in *Child Phonology*, 1980. doi: 10.1016/b978-0-12-770601-6.50011-5.

[2] C. Stoel-Gammon, "D. K., Oller, The emergence of the speech capacity. Mahwah, NJ: Lawrence Erlbaum Associates, 2000. ISBN 0-8058-2629-7.," *J Child Lang*, vol. 30, no. 3, pp. 731–734, Aug. 2003, doi: 10.1017/S0305000903005737.

[3] C. C. Lee, Y. Jhang, G. Relyea, L. mei Chen, and D. K. Oller, "Babbling development as seen in canonical babbling ratios: A naturalistic evaluation of all-day recordings," *Infant Behav Dev*, vol. 50, p. 140, Feb. 2018, doi: 10.1016/J.INFBEH.2017.12.002.

[4] R. E. Stark, "Stages of speech development in the first year of life," *Child Phonology*, pp. 73–92, Jan. 1980, doi: 10.1016/B978-0-12-770601-6.50010-3.

[5] M. Cychosz *et al.*, "Vocal development in a large-scale crosslinguistic corpus," *Dev Sci*, vol. 24, no. 5, p. e13090, Sep. 2021, doi: 10.1111/DESC.13090.

[6] T. Másdóttir, S. McLeod, and K. Crowe, "Icelandic children's acquisition of consonants and consonant clusters," *Journal of Speech, Language, and Hearing Research*, vol. 64, no. 5, 2021, doi: 10.1044/2021_JSLHR-20-00463.

[7] C. K. S. To, Cheung, Pamela S.P, McLeod, and Sharynne, "A population study of children's acquisition of Hong Kong Cantonese consonants, vowels, and tones," *Journal of Speech, Language, and Hearing Research*, vol. 56, no. 1, 2013, doi: 10.1044/1092-4388(2012/11-0080).

[8] L. McCune and M. M. Vihman, "Early Phonetic and Lexical Development: A Productivity Approach," *Journal of Speech, Language, and Hearing Research*, vol. 44, no. 3, 2001, doi: 10.1044/1092-4388(2001/054).

[9] E. Kidd and R. Garcia, "How diverse is child language acquisition research?," *First Lang*, vol.

42, no. 6, 2022, doi: 10.1177/01427237211066405.
[10] M. R. Talbott, C. A. Nelson, and H. Tager-Flusberg, "Maternal Vocal Feedback to 9-Month-Old Infant Siblings of Children with ASD," *Autism Res*, vol. 9, no. 4, pp. 460–470, Apr. 2016, doi: 10.1002/AUR.1521.
[11] E. Patten, K. Belardi, G. T. Baranek, L. R. Watson, J. D. Labban, and D. K. Oller, "Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency," *J Autism Dev Disord*, vol. 44, no. 10, 2014, doi: 10.1007/s10803-014-2047-4.
[12] B. Hart and T. R. Risley, "Meaningful differences in the everyday experience of young American children," p. 268, 1995.
[13] L. D. Lopez, E. A. Walle, G. M. Pretzer, and A. S. Warlaumont, "Adult responses to infant prelinguistic vocalizations are associated with infant vocabulary: A home observation study," *PLoS One*, vol. 15, no. 11 November, 2020, doi: 10.1371/journal.pone.0242232.
[14] E. Bergelson, A. Amatuni, S. Dailey, S. Koorathota, and S. Tor, "Day by day, hour by hour: Naturalistic language input to infants," *Dev Sci*, vol. 22, no. 1, 2019, doi: 10.1111/desc.12715.
[15] J. Gilkerson and J. A. Richards, "The LENA natural language study, LENA Technical Report: LTR-02-2," 2009.
[16] P. Karanth, S. Shaista, and N. Srikanth, "Efficacy of communication DEALL--an indigenous early intervention program for children with autism spectrum disorders," *Indian J Pediatr*, vol. 77, no. 9, pp. 957–962, Sep. 2010, doi: 10.1007/S12098-010-0144-8.
[17] A. Wetherby and B. Prizant, *Communication and symbolic behavior scales: Developmental profile*. 2002. Accessed: Jan. 04, 2023. [Online]. Available: https://psycnet.apa.org/record/2002-17060-000
[18] D. Xu, U. Yapanel, S. Gray, and C. T. Baer, "The LENA(TM) language environment analysis system: The interpreted time segments (ITS) file," *LENA Foundation Technical Report*, no. September 2008, 2008.
[19] M. L. McHugh, "Interrater reliability: The kappa statistic," *Biochem Med (Zagreb)*, vol. 22, no. 3, 2012, doi: 10.11613/bm.2012.031.
[20] D. K. Oller, R. E. Eilers, R. Urbano, and A. B. Cobo-Lewis, "Development of precursors to speech in infants exposed to two languages," *J Child Lang*, vol. 24, no. 2, 1997, doi: 10.1017/s0305000997003097.
[21] A. S. Warlaumont and H. L. Ramsdell-Hudock, "Detection of total syllables and canonical syllables in infant vocalizations," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2016, vol. 08-12-September-2016. doi: 10.21437/Interspeech.2016-1518.

[i] Data presented here are from infants who have participated in a larger study.

[ii] Only infants exposed to these languages were selected since (a) these are languages widely spoken in the study site (b) clinician-researchers/annotators in the lab spoke these languages only (c) these languages were part of the inclusion criteria of the main study these infants were recruited from

[iii] Requirement for the bigger study.