

Research Article

Mean length of utterance in words (MLUw) and lexical diversity among Urdu speaking children between 3 to 3.11 years of age

Sidra Shabbir^{1*}, Rabia Zubair², Sajeela Aatif³, Amina Iqbal⁴, Mahnoor Javed⁵

ABSTRACT

Background: Language is an excellent and convenient means of communication. Mean Length of Utterance (MLU) measures linguistic productivity in children.

Objectives: to find the mean length of utterance in words and lexical diversity among children aged 3-3.11 years typically developing Urdu-speaking.

Methods: The data of n=100 children of both genders of age between (3 to 3.11 years) who performed on blank levels 1 and 2 was collected through a cross-sectional survey. The data was collected from day-care centres and schools in Rawalpindi and Islamabad from November 2022 to June 2023. MLU calculation procedure was carried out through activities from ICW and Blank Level. For the calculation of MLU, the formula was used.

Results: The Mean Length of Utterance at word level (MLUw) of children aged between 3 to 3.11 years was 3.43. In lexical diversity (LD) children have used Nouns (NN) most abundantly as 55% of Verbs (VB) with 35% Tense Auxiliaries (TA)30%.

Conclusion: It is found that children's MLUw of 3.43 between the ages of 3 and 4 indicates an increase in speech complexity. Nouns, verbs, and tense auxiliaries have been recognized as the main elements via lexical analysis.

Keywords: diversity; language disorders; lexical marker; speech pathology.

Designation & Affiliation

¹ Lecturer, Faculty of Rehabilitation & Allied Health Sciences, Riphah International University Islamabad

² Assistant Professor, Faculty of Rehabilitation & Allied Health Sciences.

³ Virtual Therapist, Little Learners, Islamabad Pakistan

⁴ Audiologist, Children Hospital, Lahore Pakistan

Citation

Shabbir S, Zubair R, Aatif S, Iqbal A, Javed M. Mean length of utterance in words (MLUw) and lexical diversity among Urdu speaking children between 3 to 3.11 years of age. T Rehabili. J. 2024;08(02); 03-07 doi: 10.52567/trehabj.v8i02.52.

Copyright (c) 2024



Sidra Shabbir, Rabia Zubair, Sajeela Aatif, Amina Iqbal., Mahnoor Javed. This work is licensed under a Creative Commons Attribution 4.0. Authors retain copyright and grant the journal right of first publication and allows others to share the work with an acknowledgment of the work's authorship and initial publication in this journal. No use, distribution or reproduction is permitted which does not comply with these terms.

Article History

Submitted: 22-12-2023

Accepted: 12-05-2024

Published: 20-05-2024

Correspondence*

Sidra Shabbir, Lecturer Faculty of Rehabilitation & Allied Health Sciences, Riphah International University Islamabad Pakistan.

E-mail: sidrashabbir82@gmail.com

INTRODUCTION

Language facilitates social interaction and environmental adaptation and can be in the expressive or receptive form [1, 2]. The mean length of utterance (MLU) is used to assess expressive language as well as recognize the language deficiencies among children, individual differences in language development, and therapeutic efficacy [3-5].

The mean language of utterance at the word level is already measured in Spanish, Irish, Russian, English Southern Banto, etc. Lexical development can be evaluated by the MLUw and ranges from 3 to 5 words in 3-4-year-olds. [6-10]. MLUw typically rises as the skill of languages advances, indicating their capacity to construct increasingly difficult sentences and articulate themselves in more detail [11, 12].

Lexical diversity is a measure of how many different words a child learns and uses. Children's vocabulary develops rapidly by the age of three to four, and they may already know several hundred to more than a thousand words [13]. They may still have a somewhat narrow vocabulary, nevertheless, when compared to older children as well as adults. To communicate, they frequently rely on a core lexicon and basic phrase structures [14]. Acquisition of lexical diversity in children depends on the language and is impacted by individual input, language structure, and cultural norms as well as age. Usually, common terminology from the surrounding area is learned first [15].

A significant literature gap exists since the research on Mean Length of Utterance (MLU) and lexical variety that has been conducted so far has not specifically focused on preschoolers who speak Urdu and are between the ages of three and four. Furthermore, it has practical application for speech-language pathologists (SLPs) in evaluating language proficiency and identifying possible delays or abnormalities. Compiling normative data for MLUW and lexical variation in children speaking Urdu facilitates clinical evaluations by enabling comparisons with age-appropriate reference points.

METHODOLOGY

Cross Sectional study was conducted on n=100 typically developing children from schools and day-care centres after the approval from the research ethical committee of Riphah International University (Ref # RIPHAH/RCRS/REC Letter 01457). The sampling technique used for this research was convenient Sampling. The data was collected from day-care centres and schools in Rawalpindi and Islamabad from November 2022 to June 2023.

A total of 109 children were screened for language impairments utilizing blank levels 1 and 2 to gather data. Children, ages 3 to 3.11 years, who demonstrated satisfactory performance on blank levels 1 and 2, regardless of gender, were included in the study. The n=9 children with documented language disorders were not included in the study.

MLU calculation procedure was carried out through activities from information carrying words (ICW) and Blank Level. Demographics and utterances of the research participants were documented using age-appropriate Blank Level 1 and 2 activities. These activities were designed through noun cards, verb cards, picture description cards, object toy animals, and pretend play toys.

After obtaining consent from the site and parents of the children, audio recordings were used for recording the data. Speech sample was collected by using age-appropriate designed activities. A total of 100 utterances were collected for analysis across 2 to 3 sittings. The recorded data was

transcribed for the calculation of MLU. The formula for the calculation was the Total number of words/100 utterances [16, 17]. The rule book was established through an extensive literature review between 2009 and 2023 to determine the inclusion and exclusion of utterances at the word level in Urdu. The rule book was formulated after an extensive literature review [16-20].

Data was documented and analysed in Excel software. For the mean length of utterance mean standard deviation was calculated. While for identified lexical markers frequency and percentages were mentioned.

RESULTS

The mean age of Urdu-speaking children is 3.52 ± 0.27 years. The percentage of males is $n=28$ and that of females is $n=72$. The mean length of utterance among the participants was 3.47 ± 0.45 , which means that on average, the participants' sentences contained approximately 3.47 words.

Moreover, $n=40$ diverse lexical markers were used with a total frequency of 17183 from 100 utterances. The frequently used lexical markers, more than 1000 frequency, were Noun (30.41%), Verb (20.39%), Tense Auxiliaries (18.01%), Personal Demonstratives (8.90%) Personal Pronouns (8.27%), and Aspectual Auxiliaries (7.72%). All other lexical markers were not frequently used as their frequency was >1000 . (Figure 1)

DISCUSSION

The present study investigated children's conversational language abilities. In a group of typically developing Urdu-speaking children in Pakistan, the mean length of utterance (MLU) and lexical diversity (D) were initially determined.

In terms of language development for children between the ages of 3 and 3.11 year, Children with this age are usually still in the early phases of learning a language [21]. With an MLU of 3.47, children in this age range develop utterances that are, on average, just slightly longer than three words [3]. So children at this age develop from single words to longer and more complicated monologues. There can be possibilities of exhibiting longer and shorter utterances than the mean due to individual differences [21]. On the other hand, an MLU of 3.47 for children aged 3-3.11 years who speak Urdu often shows a trend towards more advanced speech patterns in language development.

Research has shown that children's MLUW tends to improve as they learn more words, suggesting a developmental progression in their language abilities [22]. The increased MLU could be explained by the children's language abilities' maturation stage in addition to the linguistic richness and complexity of Urdu. To make definitive decisions about MLU variances among children of this age range across languages, more comparative studies across various language groups would be required. Scholars have observed that there is an important relationship between the breadth and diversity of a child's vocabulary and the growth in MLUW, in addition to chronological age [22-24].

Certain lexical indicators stand out as extremely common and essential for the linguistic evolution of 3- to 3.11-year-old Urdu-speaking children. For appropriate assessment and intervention in speech-language pathology/therapy, it is essential to comprehend the usual acquisition and development of lexical markers in Urdu-speaking children [25] between the ages of 3 and 4 years. Nouns are among the most easily learned words; they are used in speech regularly and describe tangible objects. By regularly categorizing objects, caregivers reinforce the child's

comprehension and play a key role. Next in terms of speed of acquisition are verbs, which are important for communicating experiences and actions. Like nouns, verbs are learned

through caregiver modeling and the development of storytelling abilities [26, 27].



Figure 1: Lexical diversity among 3-3.11 years old children

Next, acquired tense and aspectual auxiliaries facilitate the expression of temporal linkages and are impacted by cultural storytelling traditions and cognitive development[28, 29]. Alongside other markers, personal

pronouns, which are intimately associated with social interactions and personal identity, emerge, underscoring the complex process of language development[30].

This understanding is used by speech-language pathologists (SLPs) to evaluate language competence, identify areas of strength and weakness, and create specialized intervention strategies. Understanding how social norms affect language development necessitates cultural sensitivity. Including caregivers encourages interaction and the generalization of skills. To address speech and language impairments, improve the potential for language development, and minimize long-term communication difficulties, early intervention is essential. In general, an evidence-based approach that supports language development and communication skills in early Urdu-speaking children is driven by an understanding of lexical marker development.

Grey literature was not included in the rule book. Gender differences were not calculated for the mean length of utterances at the word level. The sample size is low and from an urban setting.

CONCLUSION

The mean length of utterance among the participants was approximately 3.47 words, which showed higher language skills in 3-to 3.11-year-old old children. The study also emphasizes how crucial lexical markers such as verbs, nouns, tense and aspectual auxiliaries, and personal pronouns were to the linguistic development of 3-to 3.11-year-old Urdu-speaking children. Effective evaluation and intervention in speech-language pathology and treatment depend on an understanding of the normal acquisition and development of these markers.

DECLARATIONS & STATEMENTS

Author's Contribution

SS and RZ: substantial contributions to the conception and design of the study.

SS: acquisition of data for the study.

SS: analysis of the data for the study.

SS and RZ: interpretation of data for the study.

SS: drafted the work.

SS and RZ: revised it critically for important intellectual content.

SS and RZ: final approval of the version to be published and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All authors contributed to the article and approved the submitted version.

Ethical Statement

The study was conducted in schools and daycare centers after the approval from the research ethical committee of Riphah International University (Ref # RIPHAH/RCRS/REC Letter 01457)

Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

We sincerely thank the Sky Bird School System, Islamaia Foundation School, Federal Public School, Seven Stream School System, and the parents of their students for their invaluable support and cooperation in facilitating our

research and providing essential insights.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

REFERENCES

1. Sirbu A. The significance of language as a tool of communication. *Scientific Bulletin" Mircea Cel Batran" Naval Academy.* 2015;18(2):405. [CrossRef] [PubMed]
2. Dethorne LS, Johnson BW, Loeb JW. A closer look at MLU: what does it measure? *Clin Linguist Phon.* 2005;19(8):635-48. [CrossRef] [PubMed]
3. Ezeizabarrena MJ, Garcia Fernandez I. Length of Utterance, in Morphemes or in Words?: MLU3-w, a Reliable Measure of Language Development in Early Basque. *Front Psychol.* 2017;8:2265. [CrossRef] [PubMed]
4. Gabig CS. The mean length of utterance (ml). *Encyclopedia of autism spectrum disorders.* 2021:2828-9. [CrossRef] [PubMed]
5. DeMaris A, Smith AB. Relationships among measures of longest utterances, MLU, age, and number of utterances in child language samples. *Speech Lang. Hear.* 2017;20(2):84-90. [CrossRef] [PubMed]
6. Brookes H, Makaure P, Yalala S, Danvers H, Mössmer M, Little F, et al. Mean Length of Utterance: A study of early language development in four Southern Bantu languages. *J Child Lang.* 2024;1-26. [CrossRef] [PubMed]
7. Tomas E, Dorofeeva S. Mean Length of Utterance and Other Quantitative Measures of Spontaneous Speech in Russian-Speaking Children. *J Speech Lang Hear Res.* 2019;62(12):4483-96. [CrossRef] [PubMed]
8. Matiasovitsová K, Čechová P, Sláma J, Homolková K, Smolík F. Mean Length of Utterance in Czech Toddlers: Validity Estimates and Comparison of Words, Morphemes, and Syllables. *J Speech Lang Hear Res.* 2024;67(3):837-52. [CrossRef] [PubMed]
9. Potratz JR, Gildersleeve-Neumann C, Redford MA. Measurement Properties of Mean Length of Utterance in School-Age Children. *Lang Speech Hear Serv Sch.* 2022;53(4):1088-100. [CrossRef] [PubMed]
10. Méndez LI, Perry J, Holt Y, Bian H, Fafulas S. Same or different: Narrative retells in bilingual Latino kindergarten children. *"Biling. Res. J."*. 2018;41(2):150-66. [CrossRef] [PubMed]
11. Rosselli M, Ardila A, Matute E, Vélez-Urbe I. Language Development across the Life Span: A Neuropsychological/Neuroimaging Perspective. *Neurosci J.* 2014;2014:585237. [CrossRef] [PubMed]
12. Lester NA, Moran S, Küntay AC, Allen SEM, Pfeiler B, Stoll S. Detecting structured repetition in child-surrounding speech: Evidence from maximally diverse languages. *Cognition.* 2022;221:104986. [CrossRef] [PubMed]
13. Wood CL, Bustamante KN, Schatschneider C, Hart SA. Relationship Between Children's Lexical Diversity in Written Narratives and Performance on a Standardized Reading Vocabulary Measure. *Assess Eff Interv.* 2019;44(3):173-83. [CrossRef] [PubMed]

14. Montag JL, Jones MN, Smith LB. Quantity and Diversity: Simulating Early Word Learning Environments. *Cognitive Science*. 2018;42(S2):375-412. [[CrossRef](#)] [[PubMed](#)]
15. Tang TP, Lau DK, Leung MT. Corpus of Mandarin Child Language: a preliminary study on the acquisition of semantic content categories in Mandarin-speaking preschoolers. *Front Psychol*. 2023;14:1234525 [[CrossRef](#)] [[PubMed](#)]
16. Oosthuizen H, Southwood F. Methodological issues in the calculation of mean length of utterance. *S Afr J Commun Disord*. 2009;56:76-87. [[PubMed](#)]
17. Santos ME, Lynce S, Carvalho S, Cacela M, Mineiro A. Mean length of utterance words in children with typical language development aged 4 to 5 years. *Revista CEFAC*. 2015;17:1143-51. [[CrossRef](#)]
18. Pert S, Stow C. Bilingual Assessment of Simple Sentences (BASS): An expressive language assessment for children of Pakistani heritage who speak Mirpuri, Punjabi or Urdu. 2019. [[CrossRef](#)]
19. Khan SN, Khan K, Khan A, Khan A, Khan AU, Ullah B. Urdu word segmentation using machine learning approaches. *Int. j. adv. comput. sci. appl*. 2018;9(6):193-200. [[CrossRef](#)]
20. Durrani N, Hussain S, editors. Urdu Word Segmentation. North American Chapter of the Association for Computational Linguistics; 2010; [528–536](#).
21. Majorano M, Rainieri C, Corsano P. Parents' child-directed communication and child language development: a longitudinal study with Italian toddlers. *J Child Lang*. 2013;40(4):836-59. [[CrossRef](#)] [[PubMed](#)]
22. Cadime I, Moreira CS, Santos AL, Silva C, Ribeiro I, Viana FL. The development of vocabulary and grammar: a longitudinal study of European Portuguese-speaking toddlers. *J Child Lang*. 2019;46(4):653-81. [[CrossRef](#)] [[PubMed](#)]
23. Montanari S, Mayr R, Subrahmanyam K. Bilingual Speech Sound Development During the Preschool Years: The Role of Language Proficiency and Cross-Linguistic Relatedness. *J Speech Lang Hear Res*. 2018;61(10):2467-86. [[CrossRef](#)] [[PubMed](#)]
24. Voniati L, Armostis S, Tafiadis D. Measures of Naturalistic Language for Dialect-Speaking Children: The Case of Cypriot Greek. *Languages*. 2023;8(1):21. [[CrossRef](#)]
25. Butt GA, Mumtaz N, Saqulain G. Development & Validation of Urdu Receptive Language Scale (URLS). *Pak J Med Sci*. 2021;37(7):1924-9. [[CrossRef](#)] [[PubMed](#)]
26. Smith LB. Children's noun learning: How general learning processes make specialized learning mechanisms. *The emergence of language: Psychology Press*; 2013. p. 277-303. [[CrossRef](#)] [[PubMed](#)]
27. Waxman S, Fu X, Arunachalam S, Leddon E, Geraghty K, Song HJ. Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child development perspectives*. 2013;7(3):155-9. [[CrossRef](#)] [[PubMed](#)]
28. Desclés J-P. A cognitive and conceptual approach to tense and aspect markers. 2016. [[CrossRef](#)]
29. Zhang M, Hudson JA. The Development of Temporal Concepts: Linguistic Factors and Cognitive Processes. *Front Psychol*. 2018;9:2451. [[CrossRef](#)] [[PubMed](#)]
30. Orvell A, Gelman SA, Kross E. What “you” and “we” say about me: How small shifts in language reveal and empower fundamental shifts in perspective. *Social and Personality Psychology Compass*. 2022;16(5):e12665. [[CrossRef](#)]