

Original articles

Relationship between mean length of utterance and vocabulary in children with typical language development

Letícia Arruda Nóro¹<https://orcid.org/0000-0002-4410-5184>Helena Bolli Mota¹<https://orcid.org/0000-0001-9505-0964>

¹ Universidade Federal de Santa Maria - UFSM, Santa Maria, Rio Grande do Sul, Brasil.

Research support source: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

Conflict of interests: Nonexistent



Received on: May 1, 2019
Accepted on: October 25, 2019

Corresponding address:

Letícia Arruda Nóro
Rua Felipe de Oliveira, 601, apartamento
201, Bairro Centro
CEP: 97015-250 - Santa Maria,
Rio Grande do Sul, Brasil
E-mail: titamecmat@gmail.com

ABSTRACT

Purpose: to investigate the existence of a relationship between vocabulary and measures of mean length of utterance in children in their language development phase.

Methods: the sample consisted of 72 children aged 2 to 4 years, 11 months and 29 days, 36 boys and 36 girls, with typical language development, evenly distributed into age groups, enrolled in kindergartens with the public school system, in Santa Maria, RS, Brazil. Videos of the spontaneous speech of each subject were made, and then, the analysis of the vocabulary and Mean Length of Utterance took place. Statistical analysis was performed using the Statistical Analysis System program, version 9.2 and Spearman correlation coefficient, with a significance level of $p < 0.05$.

Results: the influence of gender in the Mean Length of Utterance correlation and vocabulary was observed. There was a difference between the ages of 2 and 4 years.

Conclusion: vocabulary development promotes mean length utterance, indicating positive correlation between gender and age range.

Keywords: Vocabulary; Language Development; Child Development; Speech; Child.

INTRODUCTION

One of the indexes for monitoring language development is vocabulary production^{1,2}. During normal language development, the child acquires words at around 12 months, following a period of slow and gradual vocabulary development, which occurs at a rate of approximately 10 words per month, up to 18 months of age. As the child approaches the 50-word milestone, this growth rate increases, characterizing the so-called “vocabulary explosion”. Already at school age this growth presents peaks of accelerated development, remaining until around 16 years. In adulthood vocabulary continues to increase, but it depends largely on the environment^{3,4}.

Children in normal language development are known to have an early phase of slow lexical acquisition and then a rapid phase, during which many words are incorporated daily. There is some individual variation regarding the ages at which such milestones⁵ occur.

This may be related to the referential of the words learned. According to studies, children are more likely to acquire open-class words (nouns, verbs, adjectives, adverbs, and numerals) that have concrete references (which is easier by context) than closed-class words (articles, prepositions, conjunctions, pronouns, interjections)^{6,7}.

In Brazilian Portuguese, the results of vocabulary research are still controversial. In a quantitative and cross-sectional study⁸ of children aged 1 to 5 years, the authors found a prevalence of verbs.

Other authors⁹ have concluded that the hypothesis that nouns prevail over verbs in the initial lexical acquisition period was confirmed in the 18-month age group, and in the 24- and 32-month age groups, there is a tendency for verbs to be match or exceed nouns, with virtually all grammatical classes present in the 32-month age group.

In another study¹⁰ it was verified how the initial lexical acquisition of children with typical development occurs, in terms of types and occurrences of the lexical items and also if the nominal bias hypothesis really occurs. The results of the analysis indicated that the number of nouns was higher than the number of verbs during the lexical acquisition period studied, but the production of nouns was not exclusive, even in this very early period of linguistic acquisition.

The expansion of the expressive vocabulary of nouns is related to the expansion of the use of longer sentences and the use of closed-class words (such as prepositions, pronouns and conjunctions), confirming

that the process of word acquisition is fundamental for syntactic development^{4,11}.

Thus, word acquisition is a crucial factor for further syntactic development, and marks the beginning of the possibility of effective oral communication between the developing child and the surrounding world. Knowing a word is not only based on knowing its meaning and its phonic form. Therefore, knowing a word involves implicitly knowing the class of words to which it belongs, since this knowledge determines the positions it can occupy in a sentence, and also involves knowing what conditions it imposes on the syntactic context in which it belongs may occur^{4,12}.

The *Extensão Média do Enunciado (EME)*, referred to in Brazilian research, derived from Mean Length Utterance (MLU) which studied aspects of grammatical development and outlined the phrasal structures and constituent elements of each segment in various age groups. The main objective was to obtain data on the performance of morphological and syntactic aspects of typically developing children with communication disorders¹³⁻¹⁶.

The Mean length of utterance-morpheme (MLU-m) was proposed as an index for verifying grammatical development. In the international literature, some authors state that there is a relationship between chronological age and MLU-m. Age and vocabulary may interact in predicting grammatical development. Thus, the child first demonstrates sensitivity to grammatical principles and regularities in comprehension, then can use them in the production^{13,17-20}.

In addition to the MLU-m calculation, some studies suggest the calculation of the MLU in words (MLU-w). Such a measure would provide data on the child's overall language development. From a study in which a high correlation was found between MLU-m and MLU-w, the use of MLU-w was indicated as a more reliable measure for calculating the length of utterance and more sensitive to the child's language complexity^{21,22}.

In this context, we sought to investigate the hypothesis that as vocabulary increases, MLU increases in the same proportion. Based on the above, the aim of this study was to investigate the existence of a relationship between the mean length of utterance and vocabulary in children in their language development, enrolled in kindergartens of the Santa Maria municipal public system (RS, Brazil).

METHODS

This research is characterized as quantitative, descriptive and cross-sectional data collection, being part of a project approved by the Research Ethics Committee of the Universidade Federal de Santa Maria, under number 0219.0.243.000-11. As a mandatory condition for participation in the study, those responsible for the subjects signed the Informed Consent Form.

The sample of this article was selected by convenience and consists of 72 children aged 2 years to 4 years 11 months and 29 days, 36 boys and 36 girls, evenly distributed among age groups. They were matched for gender. All children were members of Brazilian Portuguese-speaking monolingual families with typical language development. Exclusion factors were: hearing loss; neurological, emotional and / or cognitive impairment; presence of oral motor or organic alterations; or children who had / were doing speech therapy. These questions were observed during a previous interview and confirmed according to the interview with parents or guardians.

The sample selection was carried out in municipal preschools of kindergarten in a city of Rio Grande do Sul. The speech-language evaluation included a questionnaire designed for those responsible, orofacial and oral language evaluation and hearing screening.

To assess the orofacial aspects, the adapted "Protocol orofacial myofunctional evaluation with scores (OMES)"²¹ was used. Language assessment was performed using the "Behavioral Observation Protocol"²² for children up to 4 years old. With older children, attention was given to spontaneous oral narratives, answers to questions and observation of play.

The hearing screening for children up to 2 years 6 months and 29 days old was Visual Reinforcement Audiometry²³ and for children aged 2 years and 7 months up to 5 years 11 months and 29 days, the audiological evaluation was performed with conditioned ludic audiometry or pure tone audiometry²⁴. In case of failure in the answers, in one or more frequencies, and in two consecutive screenings, the child was referred for complete otorhinolaryngological and audiological evaluation.

Thus, with the children who met the inclusion criteria, the vocabulary was assessed through spontaneous speech and the naming of miniature objects and toys. Video recordings were made for 20 minutes. For speech transcription of children up to 3; 3; 29, the consensus method was used^{25,26}. For children of the

other age groups, who present the most stable productions, the reliability method between the transcripts was used^{27,28}.

For vocabulary analysis, two criteria were used: "types and occurrences" or "types and tokens" and "open class and closed class" of words.

- Types and occurrences or types and tokens: children's speech was separated by words, and the types and occurrences (types and tokens) of each subject were accounted for. For the classification of types were considered all the different words used by the child. The count of occurrences followed the same criteria, from the identification of the number of repetitions of each type of word in the corpus. Thus, it was possible to verify the production frequencies of each word class: open class (content words): nouns, adjectives, verbs, adverbs and numerals; closed class (functional words): articles, prepositions, conjunctions, pronouns, interjections.

For MLU analysis, the subject's speech sample was divided into statements, until the number of 100 of them was reached. When this number was reached, the remainder of the transcript was discarded. However, transcripts that obtained less than 100 statements were considered. Large speech samples are significantly related, but not significantly different, and may dramatically reduce sample size²⁹. For the scoring were used the criteria proposed by Araújo and Befi-Lopes (2004):

- Articles: one (1) point to mark gender (female: one (1) point; male: one (1) point) and one (1) point to mark number (singular: one (1) point; plural: one (1) point);
- Noun: one (1) point to mark the gender, one (1) point to mark the number, one (1) point to mark the augmentative and one (1) point to mark the diminutive;
- Verbs: one (1) point to mark number-person and one (1) point to time morphemes mode;
- Pronouns: one (1) point for each occurrence;
- Prepositions: one (1) point for each occurrence;
- Conjunctions: one (1) point for each occurrence;

Contractions of prepositions with articles, pronouns, or other elements were counted as one morpheme. For example, "*no, da, neste, daqui*".

Grammatical morphemes (GM) were grouped into two groups for better data visualization: GM-1: nouns, verbs and articles and GM-2: conjunctions, pronouns and prepositions. The sum of GM-1 and

GM-2 constituted the total MLU-m. For the calculation of MLU – w all words were counted by dividing them by the total number of utterances.

All exact segment repetitions have been included. Disfluencies as word repetitions will be scored once.

Finally, the Statistical Analysis System Program, version 9.2 was used. With the same program, the Spearman correlation test was used. In both tests, the significance level adopted was 5% ($p < 0.05$).

RESULTS

To verify whether the vocabulary and mean length of utterance measures of the statement correlate, the Spearman correlation coefficient was used, as shown in Table 1. The results indicated the existence of a positive correlation between all variables. Thus, it was possible to verify that with the expansion of vocabulary there is an increase in the mean length of the utterance, both in words and in morphemes.

Table 1. Results of correlation between mean length of utterance and vocabulary

	r	P-value
MLU-m x OC	0.440	0.0001*
MLU-m x CC	0.496	0.0001*
MLU-w x OC	0.454	0.0001*
MLU-w x CC	0.450	0.0001*

Caption: MLU-m: mean length of utterance in morphemes; MLU-w: mean length of utterance in words. OC: open class; CC: closed class. r: correlation of Spearman. P values < 0.05 were considered statistically significant. * indicates positive correlation

Table 2 presents the correlation analysis of the vocabulary and MLU (morphemes and words) in each age group. From the data it was found that there is a

statistically significant difference in the age group of 2 and 4 years, but mainly in the 2 years.

Table 2. Results of correlation between mean length of utterance and vocabulary, in age groups

Age groups		r	Value of P
2 years	MLU-m x OC	0.553	0.005*
	MLU -m x CC	0.581	0.002*
	MLU -w x OC	0.470	0.020*
	MLU -w x CC	0.362	0.081
3 years	MLU -m x OC	0.257	0.225
	MLU -m x CC	0.268	0.204
	MLU-w x OC	0.255	0.227
	MLU-w x CC	0.309	0.141
4 years	MLU-m x OC	0.276	0.191
	MLU -m x CC	0.184	0.388
	MLU -w x OC	0.461	0.023*
	MLU -w x CC	0.323	0.123

Caption: MLU-m: mean length of utterance in morphemes; MLU-w: mean length of utterance in words. OC: open class; CC: closed class. r: correlation of Spearman. P values < 0.05 were considered statistically significant. * indicates positive correlation.

Table 3 presents the analysis of the influence of gender on the correlation of MLU and vocabulary.

Male children had significant results on the influence of vocabulary and mean length of utterance.

Table 3. Analysis of the influence of gender on mean length of utterance and vocabulary

		r	Value of P
Females	MLU-m x OC	0.184	0.281
	MLU-m x CC	0.275	0.103
	MLU-w x OC	0.197	0.247
	MLU-w x CC	0.187	0.273
Males	MLU-m x OC	0.630	0.0001*
	MLU-m x CC	0.557	0.0004*
	MLU-w x OC	0.614	0.0001*
	MLU-w x CC	0.512	0.0014*

Caption: MLU-m: mean length of utterance in morphemes; MLU-w: mean length of utterance in words. OC: open class; CC: closed class. r: correlation of *Spearman*. P values <0.05 were considered statistically significant. * indicates positive correlation

DISCUSSION

In the results found in Table 1, it can be seen that the increase in vocabulary implies a greater extension of the utterance, due to a greater need to refine the phonological and lexical representations, indicating a high correlation.

The highest correlation was found between the MLU-m and the closed class, confirming that for the syntactic development the word acquisition process is fundamental⁴, although the difference is not large.

This result is in line with a study¹⁰ conducted with Brazilian Portuguese-speaking children with specific language impairment, which aimed to verify the influence of age on lexical and grammatical performance, as well as to investigate the existence of correlation between expressive vocabulary and language measures average length of utterance. The results indicated that the vocabulary expansion favors the extension of the phrasal extension.

An international study showed⁷ that early in the process of language acquisition children are more likely to acquire open-class words (comprising nouns, verbs, adjectives, adverbs, and numerals). This is due to its more concrete content and its greater use in the language. As children need intense exposure to new lexical items in order to increase vocabulary, they begin to acquire closed-class words due to the syntactic requirement for phrasal linking elements^{13,30}.

In this sense, as the vocabulary expands, the length of the children's utterance tends to increase concomitantly and, thus, the introduction of closed-class words in the utterance becomes necessary for the connection between the sentence words and between the words to occur own sentences³¹.

Due to time, frequency and linguistic maturity, the vocabulary is progressively refined, generating the

development of networks of relationships between words through semantic categorizations, such as the word "dog". The child first learns the basic concept, then learns the concept that is "animal", and finally a subordinate concept, "labrador dog". Thus, the discourse and the formation of ideas by the children accompany this development, gradually occurring the predominance of more elaborate and articulated characteristics, implying a high correlation between vocabulary and mean length of the utterance³²⁻³⁴.

In the present research it was possible to notice that the increase of vocabulary is of great importance for the learning process, as it will contribute to the reading and word comprehension process. This knowledge covers various aspects of words, such as pronunciations, definitions of their meanings, syntactic rules of their uses, and the way words are written. All of these elements are connected to a network of mental connections and their quality will relate to the proper use of vocabulary^{35,36}.

Regarding age (Table 2), it was found that the age of 2 years was the one that obtained the highest correlation indexes. These results may lead us to reflect that in this period shorter and simpler sentences are produced, since at this stage the child uses mainly verbs, adjectives and nouns, which does not require the use of pronouns, prepositions, conjunctions and adverbs. Thus, the combination of few words occurs and the elements are ordered correctly. At the age of 3 years, no significant correlation was observed. This can be justified by the fact that at this age children use phrasal structures in which there is a higher occurrence of words with lower semantic load, thus being used in combination with nouns and verbs.

Finally, at the age of 4 years, only a significant correlation was observed between open class and MLU-w.

At this age, phrasal structures become more complex and well-developed semantic-syntactic relations can be observed using a large number of closed-class words¹², which cannot be observed in the present study.

This result is similar to the research cited above. Thus, it is inferred that age is not sufficient to promote grammatical development, but it will be necessary to integrate linguistic knowledge.

Table 3 shows the relationship of gender influence in the correlation of the mean length of utterance and vocabulary. There is a correlation in males between the variables. This result does not demonstrate the tendency for girls to acquire words at a slightly faster rate than boys in general studies. However, this may be justified by the fact of individual differences^{36,37}.

What may also be a justification is the possible interference of the socio-cultural level of origin of children, which was not the purpose of the study. But it can be inferred that high-level children would produce more complex statements^{38,39}.

CONCLUSION

From the study with 72 children, it is emphasized that the vocabulary development promotes the extension of utterance, indicating the positive correlation between the variables; Regarding the age group, a significant difference was observed for the age groups of 2 and 4 years, especially in the 2-year group; It can be inferred that linguistic knowledge results from an expansion of vocabulary, and, consequently, favors the extension of utterance and the learning of closed-class words.

Although the number of children in each age group was relatively small, the data obtained may serve as a reference for other research, since the number of statements was high and the inclusion criteria of participants, speech elicitation and analysis of verbal production were rigorous.

ACKNOWLEDGEMENTS

We thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for their support in this research.

REFERENCES

1. Befi-Lopes DM, Galea DES. Análise do desempenho lexical em crianças com alteração no desenvolvimento da linguagem. *Pró-Fono R Atual Cient.* 2000;12(9):31-7.
2. Befi-Lopes DM, Puglisi ML, Rodrigues A, Giusti E, Gândara JP, Araújo K. Perfil comunicativo de crianças com alterações específicas no desenvolvimento da linguagem: caracterização longitudinal das habilidades pragmáticas. *Rev Soc Bras Fonoaudiol.* 2007;12(4):265-73.
3. Benedict KS. Early lexical development: comprehension and production. *J Child Lang.* 1979;6:183-200.
4. Gândara JP, Befi-Lopes DM. Tendências da aquisição lexical em crianças em desenvolvimento normal e crianças com Alterações Específicas no Desenvolvimento da Linguagem. *Rev Soc Bras Fonoaudiol.* 2010;15(2):297-304.
5. Bassano D, Maillachon I, Eme E. Developmental changes and variability in the early lexicon: a study of French children's naturalistic productions. *J Child Lang.* 1998;25(3):493-531.
6. Bloom P. Précis of how children learn the meanings of words. *Behav Brain Sci.* 2001;24(6):1095-103.
7. D'odorico L, Fasolo M. Nouns and verbs in the vocabulary acquisition of Italian children. *J Child Lang.* 2007;34(4):891-907.
8. Sherer S, Souza APS. Types e tokens na aquisição típica de linguagem por sujeitos de 18 a 32 meses falantes do português brasileiro. *Rev. CEFAC.* 2011;13(5):838-46.
9. Nóro LA, Silva DD, Wiethan FM, Mota HB. Initial lexical acquisition and noun bias hypothesis verification. *Rev. CEFAC.* 2015;17(suppl.1):52-9.
10. Befi-Lopes DM, Nuñez CO, Cáceres AM. Correlation between expressive vocabulary and mean length utterance in children with language disorder. *Rev. CEFAC.* 2013;15(1):51-7.
11. Felix J, Santos ME, Benitez-Burraco A. Spontaneous language of preterm children aged 4 and 5 years. *Rev. CEFAC.* 2017;19(6):742-8.
12. Brown R. *A first language.* Cambridge MA: Harvard University Press; 1973.
13. Araujo K, Befi-Lopes DM. Extensão média do enunciado de crianças entre 2 e 4 anos de idade: diferenças no uso de palavras e morfemas. *Rev Soc Bras Fonoaudiol.* 2004;9(1):156-63.
14. Fensterseifer A, Ramos AP. Extensão média de enunciados em crianças de 1 a 5 anos. *Pró-Fono R. Atual. Cient.* 2003;15(3):251-8.
15. Thordardottir ET. Early lexical and syntactic development in Quebec French and English: implications for cross-linguistic and bilingual

- assessment. *Int J Lang Commun. Disord.* 2005;40(3):243-78.
16. Rice ML, Redmond SM, Hoffman L. Mean length of utterance in children with specific language impairment and in younger control children shows concurrent validity and stable and parallel growth trajectories. *J Speech Lang Hear Res.* 2006;49(4):793-808.
 17. Dixon JA, Marchman VA. Grammar and the lexicon: developmental ordering in language acquisition. *Child Dev.* 2007;78(1):190-212.
 18. Rice ML, Smolik F, Perpich D, Thompson T, Rytting N, Blossom M. Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *J Speech Lang Hear Res.* 2010;53(2):333-49.
 19. Hickey T. Mean length of utterance and the acquisition of Irish. *J Child Lang.* 1991;18(3):553-69.
 20. Malakoff ME, Mayes LC, Schottenfeld R, Howell S. Language production in 24-month-old inner-city children of cocaine and-otherdrug- using mothers. *J Appl Dev Psychol.* 1999;20(1):159-80.
 21. Felício CM, Ferreira CL. Protocol of orofacial myofunctional evaluation with scores. *Int J Ped Otorhinolaryngol.* 2008;72(3):367-75.
 22. Zorzi JL, Hage SRV. PROC - Protocolo de observação comportamental: avaliação de linguagem e aspectos cognitivos infantis. São José dos Campos: Pulso Editorial; 2004.
 23. Lidden G, Kankkonen A. Visual reinforcement audiometry. *J Acta Oto-Laryngologica.* 1961;67:281-92.
 24. Northern JL, Downs MP. Avaliação auditiva comportamental. 5ª ed. Rio de Janeiro: Guanabara Koogan; 2005.
 25. Shriberg LD, Kwiatkowski J, Hoffmann KA. A procedure for phonetic transcription by consensus. *J Speech Hear Res.* 1984;27(3):456-65.
 26. Morris SR. Test-Retest Reliability of Independent Measures of Phonology in the Assessment of Toddlers' Speech. *Lang. Speech Hear. Serv. Schools.* 2009;40:46-52.
 27. McLeod S, Harrison LJ, McCormack J. The intelligibility in context scale: validity and reliability of a subjective rating measure. *J Speech Lang. Hear. Res.* 2012;55(2):648-56.
 28. Sosa AV, Stoel-Gammon C. Lexical and phonological effects in early word production. *J Speech Lang Hear Res.* 2012;55(2):596-608.
 29. Casby MW. An examination of the relationship of sample size and mean length of utterance for children with development language impairment. *Child Lang TeachTher.* 2011;27(3):286-93.
 30. Castilho A. Fundamentos teóricos da gramática do português culto falado no Brasil: sobre o segundo volume, classes de palavras e as construções gramaticais. *Alfa.* 2007;51(1):99-135.
 31. Tonietto L, Siqueira M, Parente MA. Da aquisição das primeiras palavras aos significados literal e metafórico. In: Eisenberg Z, Parente MA (orgs). *Psic da linguagem - da constituição da fala às primeiras narrativas.* São Paulo: Vetor; 2010. p. 57-82.
 32. Santos MT, Befi-lobes DM. Vocabulário, consciência fonológica e nomeação rápida: contribuições para a ortografia e elaboração escrita. *J. Soc. Bras. Fonoaudiol.* 2012;24(3):269-75.
 33. Souza D, Eisenberg Z. Desenvolvimento lexical: aprendendo palavras abstratas. In: Eisenberg Z, Parente MA (orgs). *Psic da linguagem - da constituição da fala às primeiras narrativas.* São Paulo: Vetor; 2010. p. 109-30.
 34. 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. *Rev. CEFAC.* 2015;17(4):1143-51.
 35. Ferracini F, Capovilla AGS, Dias NM, Capovilla FC. Avaliação de vocabulário expressivo e receptivo na educação infantil. *Rev. Psicopedagog.* 2006;23(71):124-33.
 36. Barret MD. Desenvolvimento lexical inicial. In: Fletcher P, Macwhinney B (orgs). *Compêndio da linguagem da criança.* Porto Alegre: Artes Médicas; 1997. p. 299-322.
 37. Carvalho AMA, Befi-Lopes DM, Limongi SCO. Mean length utterance in Brazilian children: a comparative study between Down syndrome, specific language impairment, and typical language development. *CoDAS.* 2014;26(3):201-7.
 38. Le Normand MT, Parisse C, Cohen H. Lexical diversity and productivity in French preschoolers. *Clin Linguist Phon.* 2008;22(1):47-58.
 39. Walker D, Greenwood C, Hart B, Carta J. Prediction of school outcomes based on early language production and socioeconomic factors. *Children and Poverty.* 1994;65(2):606-21.