Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

Speaking Rate in 3-4-Year-Old Children: Its Correlation with Gesture Rate and Word Learning

Zhenyang Xi, Yan Gu, Gabriella Vigliocco

Department of Experimental Psychology, University College London, London, UK
zhenyang.xi.17@ucl.ac.uk, yan.gu@ucl.ac.uk, g.vigliocco@ucl.ac.uk

Abstract

Past research has shown that while speaking children before 3-year-old often use gesture to supplement speech while not using gesture as an integrated system with speech, and that the relationship between speech and gesture may relate to vocabulary development. However, such a relationship is unknown in 3-4-year-old children, a period in which we can capture key developmental changes from using gestures alone to using them along with speech. Using a new corpus of semi-naturalistic interaction between caregivers and their 3-4-year-old children (ECOLANG Corpus), this study investigates (1) the effect of age on children's speaking and gesture rate, (2) the relationship between speaking and gesture rates and (3) their correlation with word learning. Specifically, we studied speaking and gesture rates of 32 English-speaking children while talking with their caregivers about sets of pre-selected toys. The children completed a vocabulary test at the time of the experiment and one year later. Results show that there was no effect of age on speaking and gesture rates at this age range, but we found that children with a fast speaking rate also had a higher gesture rate. Additionally, neither speaking rate nor gesture rate correlates with word. Thus, our findings show that by this age, children use gestures that are integrated with speech and their relationship is no longer a predictor of vocabulary learning. We speculate that the transition in the relationship is mainly a result of enhanced conceptual representation ability.

Index Terms: speaking rate, gesture rate, multimodal language development, word learning, children

1. Introduction

People gesture while they talk [1]. However, there are great variations in people's speed of talking or rate of gestures [1], [2]. For example, when speakers repeatedly mention a referent, their gesture rate decreases [3], [4]. In the foreign language context, language learners commonly experience disfluency with a slower speaking rate, and they often use gestures instead of speech [5]. When talking to children, people tend to have a higher gesture rate and slower speaking rate for unknown objects than known objects, e.g., [6], [7]. Speaking rate and gesture rate appear to be closely linked [8] and the relationship between them may vary with age [9].

During the earliest stages of language development, gesture and speech are considered separate systems because they are often produced in isolation and rarely occur simultaneously or share the same semantic meaning [10]. Children at these earliest stages have restricted ability to speak, and gestures often occur without speech and are adopted as a complementary tool to their speech [11]. For example, a child may point to an object of which the word form is unknown to them and then say the words "Mummy, I… that…" to convey the meaning: I want that object. In this circumstance, they use a deictic gesture, i.e., point, to substitute for naming the object. This speech-gesture combination enables children to express elements of a sentence using different modalities, which they would not be able to convey using speech only [9]. Thus, they typically have a rather slower speaking rate but a higher gesture rate when talking about an object. Moreover, at this stage, the early use of such speech-gesture combinations predicts the transition from one-word to two-word utterances, and vocabulary growth [12]. The complementary relationship between speaking rate and gesture rate is consistent with the tradeoff hypothesis, which states that when speakers have difficulty in expressing ideas in one modality, they may rely on the other modality to compensate for the inability [13].

At a later developmental stage when vocabulary continues to grow, there is an age-related increase in both speaking rate and gesture rate [15], [16]. Moreover, speech and gesture start to develop temporal synchrony as well as semantic coherence [12], [17], and they become highly unified both temporally and semantically in adults whose language has already been well established [9]. In adults, the meanings conveyed by speech are often reinforced by concomitant gestures and synchronization is maintained, for instance, by pausing one modality if production in the other modality is delayed [18]. Motivated redundancy between speech and gesture seems to occur more frequently in adult interlocutors for an effective communication [19].

It is unknown when and how such transition occurs. To better understand the co-development of speech and gesture as a function of children's age, and how it correlates with vocabulary development, we focus on children aged from 3 onwards who have entered the later multi-word stage while vocabulary growth continues. Here, we carry out a corpus-based investigation of speaking rate and gesture rate, their association with age, and their relationship to word learning. As in previous studies, e.g., [14], [19], we will examine the rates of speaking and gesturing as a proxy to approach the link between speech and gesture. Specifically, we ask the following questions:

1. Is there any effect of age (in month) on speaking rate and gesture rate among 3-4 years old children?

2. What is the relationship between the speaking rate and gesture rate of 3-4 years old children? Low speaking rate and high gesture rate are expected if children use gesture to supplement their speech, whereas high speaking rate and high gesture rate are expected if speech and gestures are integrated.

3. Do speaking and gesture rates jointly predict 3-4 years old children's concurrent vocabulary size and/or their vocabulary development 1-year later? It is expected that speaking rate and gesture rate will be associated with

10.21437/SpeechProsody.2022-48

vocabulary size if speech and gesture are still complementary to each other within this age range. By contrast, the relationship between speech and gesture may not be predictive of vocabulary development if they have already been integrated well, which is more adult-like. In this case, we assume speaking rate may alone predict vocabulary size as increased speaking rate might be an externalized representation of a larger vocabulary size.

2. Methodology

This study uses the video recordings of caregiver-child interaction from the Ecological Language (ECOLANG) Corpus [20], which is a multimodal corpus of semi-naturalistic interaction between caregivers and their child.

2.1. Participants

Participants were 32 English-speaking children (17 boys and 15 girls) from the ECOLANG Corpus. Their mean age was 42.95 months, range = 36-52 months, *SD* = 4.56 (see Figure 1 for age distribution). They were all typically developing children without any language disorder according to their caregivers' report.

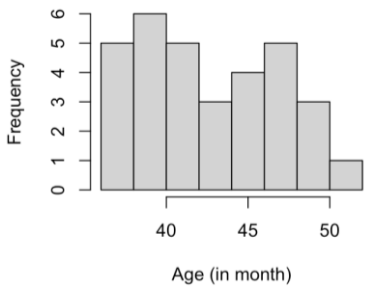

Figure 1: *Histogram displaying the distribution of ages (in month) of the sample*.

2.2. Materials

Four categories of toys (animals, foods, tools, and musical instruments) were used during the corpus collection. Within each category, there were six toys, half of which were unknown to the child and the rest half were known. A British Picture Vocabulary Scale-third edition (BPVS3), which assesses receptive vocabulary and requires no reading, was used to measure children's vocabulary size [21].

2.3. Procedure

The study was carried out in the family home. During the study, children and their caregivers were asked to talk about the 4 categories of toys in different sessions, both when the toy sets were present and when they were absent. The order of the absent and present sessions was counterbalanced so that half of the participants started with the absent sessions first and the other half started with the present condition first. Hence, there were 8 sessions in total and the whole interaction lasted on average 32 minutes. The interactions were recorded by 2 cameras, one facing the caregiver and the other facing the interaction space, along with a separate voice recorder.

The outcome measure of vocabulary size was collected by asking the children to complete the BPVS-3 test both on the day of recording and one year later. During the test, for each question, the children heard a word and chose a picture (out of four options) that they considered to best represent the meaning of the word. It is noteworthy that the 1-year follow-up vocabulary test results were missing for 5 participants due to loss of contact.

2.4. Speech and gesture coding

To study speaking and gesture rate, speech and gestures produced by the children were coded manually. Speech was transcribed by utterance. Gesture coding focused on representational gestures (gestures that in their form symbolize features or aspects of an object or action, e.g., flapping arms to represent birds [22]) and points only, which were the majority gesture used by the children. Notably, in order to be consistent with previous studies, e.g., [23], [24], object manipulations (e.g., holding a toy elephant) were not considered as gestures in this study. Thus, we only annotated the object absent condition. Each utterance and gesture were coded with the topic or toy they refer to and any topic that was irrelevant to our toy sets was coded as "other".

2.5. Computations of speaking rate and gesture rate

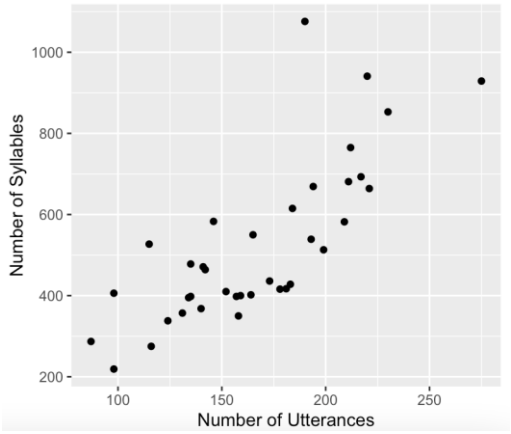

Figure 2: *Scatter plot of the number of syllables versus the number of utterances.*

The length of an utterance can differ between children of our target age. Some children can produce very long utterances whereas some mostly produced short ones. As Figure 2 shows, the number of syllables produced per utterance varied greatly. To be more normalized in the measurement, we used duration of children's speech utterances rather than the number of utterances as the analysis unit for calculating the speaking and gesture rate. Specifically, children's speaking rate was computed by dividing the total number of produced syllables in the absent condition by the corresponding speech duration whereas gesture rate was calculated by dividing the total number of produced gestures by their corresponding speech duration. We computed three types of gesture rate: (1) overall gesture rate, (2) representational gesture rate and (3) point rate.

2.6. Data analysis

For RQ1, pairwise correlations are used to investigate whether speaking rate and gesture rate vary as a function of age. For RQ2, three linear regression models are reported, which predict overall gesture rate, representational gesture rate or point rate from speaking rate, controlling for age and initial vocabulary size. For RQ3, two linear regression models are reported predicting concurrent or 1-year-later vocabulary sizes

from speaking rate and gesture rate, controlling for age. Data is centered before running all regression analyses.

3. Results

As an overview of the data, in the absent sessions, the average total duration of the speech produced by the children was 226.03 seconds, *SD* = 74.41. The children on average produced 510.81 syllables, *SD* = 175.81 and the average total number of gestures per child was 16.13, *SD* = 10.42. Of all gestures, 68.61% of gestures were representational gestures whereas 31.39% were pointing gestures. The mean speaking rate of children was 2.31 syllables/sec, *SD* = 0.51, and the mean overall gesture rate was 0.08 gesture/sec, *SD* = 0.06.

The mean BVPS3 concurrent score was 61.72, *SD* = 14.04, range = 30-87 (N=32), and the mean score for one year later was 83.15, *SD* = 8.66, range= 65-99 (N=27).

3.1. Age effect on speaking rate and gesture rate

Figure 3 presents the pairwise correlations among variables of age, overall gesture rate and speaking rate. Although there was a tendency for speaking rate ($r(30) = .30$, $p = .091$, two-tailed) and gesture rate ($r(30) = .25$, $p = .176$) to increase with age, the correlations were not significant at .05 level. Additionally, a positive correlation between speaking rate and gesture rate was observed ($r(30) = .55$, $p = .001$, 95% CI [0.25, 0.76]), which will be more closely considered in the next sub-section.

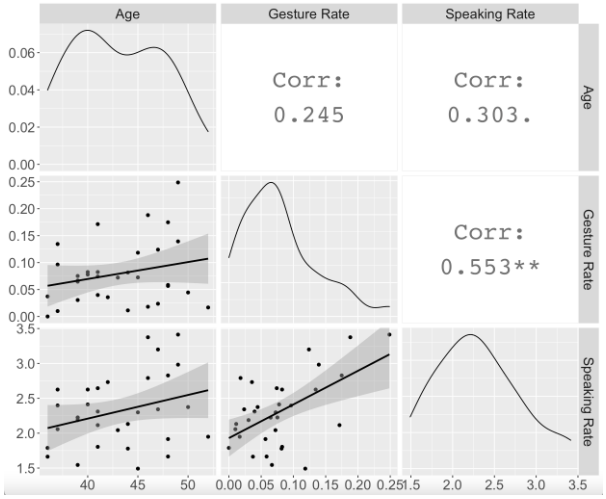

Figure 3: *Pairwise correlations among age, gesture rate and speaking rate (upper panel: Pearson correlation, diagonal: density plots displaying variable distribution, lower panel: scatter plots).*

3.2. Relationship between speaking rate and gesture rate

Table 1: *Linear regression results (t-values and significance) for testing the correlation between speaking rate and gesture rate*

Predictors	Overall Gesture Rate	Representational Gesture Rate	Point Rate
Speaking Rate	3.11**	2.67*	0.94
Age	0.04	-0.32	0.56
Initial BPVS Score	1.04	1.13	-0.05

*** $p < .001$, ** $p < .01$, * $p < .05$, . $p < .1$

Table 1 shows that speaking rate was a significant predictor of both overall gesture rate ($b = 0.06$, $p = .004$, 95% CI [0.02, 0.10]) and representational gesture rate ($b = 0.05$, $p = .013$, 95% CI [0.01, 0.08]). The correlations were positive (see Figure 4 and 5): the more syllables a child speaks per second, the more they gesture overall or make a representational gesture. However, there was no significant correlation between speaking rate and point rate. In addition, age and initial vocabulary size were not significant in all three models. In other words, there is not any correlation between age or initial vocabulary size and gesture rate (all $p > .1$).

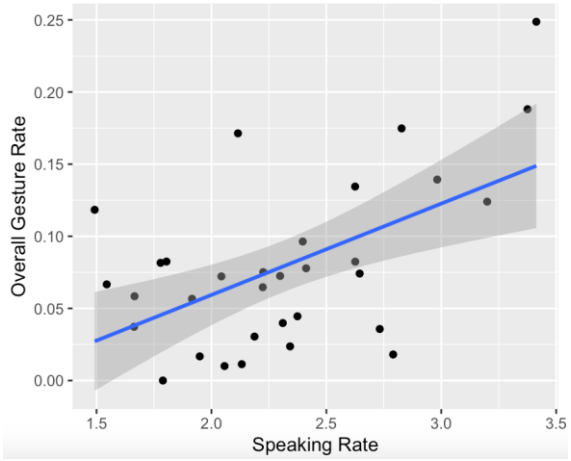

Figure 4: *Overall gesture rate and speaking rate (scattered dots represent raw data points, the solid line represents the model estimate of the slope and the shaded region represents +/- 1 SE mean).*

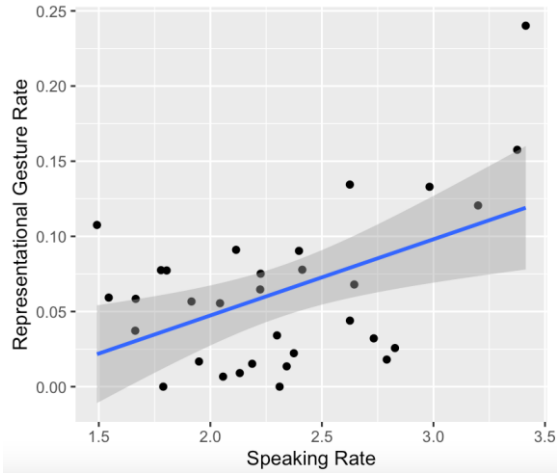

Figure 5: *Representational gesture rate and speaking rate (scattered dots represent raw data points, the solid line represents the model estimate of the slope and the shaded region represents +/- 1 SE mean).*

3.3. The effect of speaking rate and gesture rate on word learning

As shown in Table 2, inconsistent age effects were observed: age can only predict initial vocabulary size ($b = 1.36$, $p = .015$, 95% CI [0.29, 2.43]), but not 1-year-follow-up vocabulary size ($p = .849$), and the older a child is, the larger their

concurrent vocabulary size is. Neither speaking rate nor gesture rate predicted initial or follow-up vocabulary size (all $p > .38$).

Table 2: *Linear regression results (t-values and significance) for testing the correlation between speaking rate, gesture rate and vocabulary size (both concurrently and 1-year-later)*

Predictors	**Initial Vocabulary Size**	**1-year-later Vocabulary Size**
Speaking Rate	0.20	0.84
Gesture Rate	0.90	0.63
Age	2.60*	0.19

*** $p < .001$, ** $p < .01$, * $p < .05$, . $p < .1$

4. Discussion

In this study, we investigated the effect of age on speaking rate and gesture rate among children aged between 3 to 4 years old. We also assessed the relationship between the two rates and examined whether they correlated with word learning. We found that 3-4 years old children who spoke more syllables in a second also tended to make more gestures at the same time and the link between speech and representational gestures seems to potentially underlie this positive correlation. However, we did not find any age effect on speaking rate or gesture rate, nor any correlation between the two rates and word learning.

First, contrary to previous studies that suggest a constant increase in speaking rate and gesture rate with age, from young children to adults [15], [16], this study did not find any age-related effect on both rates (despite a marginal effect on speaking rate). The speaking rate and gesture rate were comparable among children aged between 3 and 4 years old. One possible explanation is that the developmental trajectories of speaking rate are non-linear [25], [26] and/or that the age range in this study is too narrow to observe such an age-related effect. Several other studies which looked at the speaking rates of pre-school children across different ages also found a non-significant age effect [27], [28], [29].

Nevertheless, the positive link between speaking rate and gesture rate indicates that a potential developmental change in the relationship between speech and gesture may have occurred before this age. The findings suggest that there is already a good integration between speech and gesture in children aged between 3 and 4 years old, which is in contrast to the younger children [12], where gesture tends to complement speech. In specific, 3-4 years old children showed the tendency to make more representational gestures when they speak more syllables, i.e., produce more speech.

Remarkably, there is a significant increase in the proportion of representational gestures produced by the 3-4 years old children in our corpus (68.61%) compared to that of the 1-2 years old children (less than 1 %) as in [23], [24]. As age increases, children's imagistic thinking ability seems to improve, which also leads to an improved ability to use gestures to represent such thinking and this enhanced ability seems to be mirrored in the increased use of representational gestures [31]. According to the gesture-for-conceptualization hypothesis, representational gestures may facilitate speaking, possibly resulting in a faster speaking rate, because they help schematize information [32] [33]. The observed positive correlation between speaking rate and representational gesture rate additionally implies that the speaking rate of children may also in turn promotes gesturing representationally.

However, we did not observe such a correlation between point rate and speaking rate. Possibly, pointing rate and speaking rate may intrinsically not hold a positive correlation. At the earliest stages of language development among 1-2 years old children, points and speech were found not to integrate together [24] and this may be retained in 3-4 years old children although it should be noted that the points they [24] looked at were not exclusively when objects were absent. Further evidence is provided by studies on points rate in speaking a second language (L2). For example, during a story-telling task (i.e., object absent), both speakers of intermediate and advanced L2 proficiency produced more points in their second language compared to the first language [5], [30] even though their fluency was reduced in their L2.

Lastly, although it is often assumed that the increase of speaking rate in children may be a result of an enlarged vocabulary size, we did not find any evidence in this sample. Despite the non-significant correlation, we could not fully exclude the possibility that young children's speaking rate and vocabulary size may be positively correlated. It is possible that boosting effect of vocabulary size on speaking rate may reach a ceiling effect when vocabulary size grows to a certain extent and only when it passes the threshold of the next stage will the effect be observed.

5. Conclusions

The current research discovered that speaking rate and gesture rate did not vary significantly within the age range between 3 and 4 years old, which may be attributed to the non-linear development of speed of speaking and gesturing. We have also shown in this study that 3-4 years old children have developed a good integration in their speech and gesture, and that the relationship between speech and gesture at this stage is not predictive of word learning. We speculate that the change in the relationship between speech and gesture (from complementary to integrated) might be attributed to the increased use of representational gestures in the children and that the boosting effect of vocabulary size on speaking rate was not observed because it reaches the ceiling at this stage.

Future studies could investigate the relationship between children's speaking rate and gesture rate for known versus unknown objects, controlling for caregivers' speaking and gesture rates, because caregivers were discovered to adjust their speech when talking to their children depending on the familiarity of the objects, i.e., slowing down for unknown objects compared to known ones [7], [34], which might potentially impact children's speaking and gesturing. Additionally, it will be useful to look at the entire caregiver-child dyads (children aged between 2 to 4 years old) in the ECOLANG Corpus to further examine an age-related effect.

6. Acknowledgements

The work reported here was supported by an ERC Advanced Grant (743035) to GV, and an NWO Rubicon Grant (019.182SG.023) to YG.

7. References

[1] J.M. Iverson and S. Goldin-Meadow, "Why people gesture when they speak," *Nature*, vol. 396, no. 6708, pp. 228-228, Nov. 1998.

[2] Y.C. Tsao and G. Weismer, "Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component," *Journal of Speech, Language, and Hearing Research*, vol. 40, no. 4, pp. 858-856, Aug. 1997.

[3] J. Holler and R. Stevens, "The effect of common ground on how speakers use gesture and speech to represent size information," *Journal of Language and Social Psychology*, vol. 26, no. 1, pp. 4-27, Mar. 2007.

[4] M. Hoetjes, R. Koolen, M. Goudbeek, E. Krahmer, and M. Swerts. "Reduction in gesture during the production of repeated references," *Journal of Memory and Language*, vol. 79, pp. 1-17, Feb. 2015.

[5] M. Gullberg, *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*, Malmö, Sweden: Lund University Press, 1998.

[6] G. Vigliocco, Y. Motamedi, M. Murgiano, E. Wonnacott, C. Marshall, I. Milán-Maillo, and P. Perniss, "Onomatopoeia, gestures, actions and words: How do caregivers use multimodal cues in their communication to children?" In *Proceedings of the 41st Annual Conference of the Cognitive Science Society*, pp. 1171-1177, 2019.

[7] J. Shi, Y. Gu, B. Gryzb, and G. Vigliocco, "Child directed speech: impact of variations in speaking-rate on word learning," In *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*, vol. 42, pp. 1043-1049, Jun. 2020.

[8] S. Kita and A. Özyürek, "What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking," *Journal of Memory and language*, vol. 48, no. 1, pp. 16-32, Jan. 2003.

[9] S. Goldin-Meadow, "The development of gesture and speech as an integrated system," *New Directions for Child Development*, no. 79, pp. 29-42, Jan. 1998.

[10] H. Sowden, M. Perkins, and J. Clegg, "The co-development of speech and gesture in children with autism," *Clinical Linguistics & Phonetics*, vol. 22, no. 10-11, pp. 804-813, Jul. 2009.

[11] M. Morford and S. Goldin-Meadow, "Comprehension and production of gesture in combination with speech in one-word speakers," *Journal of Child Language*, vol. 19, no. 3, pp. 559-580, Oct. 1992.

[12] C. Butcher and S. Goldin-Meadow, "12 Gesture and the transition from one-to two-word speech: when hand and mouth come together," in D. McNeil (Eds.), *Language and gesture*, pp. 235-258. Cambridge, MA: Cambridge University Press, 2000.

[13] J. P. De Ruiter, "Can gesticulation help aphasic people speak, or rather, communicate?" *Advances in Speech Language Pathology*, vol. 8, no. 2, pp. 124-127, Jan. 2006.

[14] E. Pearson, E. Nielsen, S. Kita, L. Groves, L. Nelson, J. Moss, and C. Oliver, "Low speech rate but high gesture rate during conversational interaction in people with Cornelia de Lange syndrome," *Journal of Intellectual Disability Research*, vol. 65, no. 6, pp. 601-607, Jun. 2021.

[15] I.S. Nip and J.R. Green, "Increases in cognitive and linguistic processing primarily account for increases in speaking rate with age," *Child development*, vol. 84, no. 4, pp. 1324-1337, Jul. 2013.

[16] J.M. Colletta, C. Pelleng, and M. Guidetti, "Age-related changes in co-speech gesture and narrative: Evidence from French children and adults," *Speech Communication*, vol. 52, no. 6, pp. 565-576, Jun. 2010.

[17] M.W. Alibali and S. Goldin-Meadow, "Modeling learning using evidence from speech and gesture," *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum, 1993.

[18] D. McNeil, *Hand and Mind: What Gestures Reveal About Thought.* London: The University of Chicago Press, 1992.

[19] J.P. De Ruiter, A. Bangerter, and P. Dings, "The interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis," *Topics in Cognitive Science*, vol. 4, no. 2, pp. 232-248, Apr. 2012.

[20] G. Vigliocco, Y. Gu, Y. Motamedi, B. Gryzb, G. Brekelmans, M. Murgiano, R. Brieke, and P. Perniss, *ECOLANG: A multimodal corpus of semi-naturalistic dyadic (adult-child and adult-adult) interaction*, unpublished.

[21] L.M. Dunn, L.M. Dunn, and C. Whetton, *British Picture Vocabulary Scale*. Windsor, England: NFER-Nelson, 1982.

[22] N.C. Capone and K.K. McGregor, "Gesture Development: A Review for Clinical and Research Practices," *Journal of Speech, Language, and Hearing Research*, vol. 47, pp. 173-186, Feb. 2004.

[23] M.L. Rowe and S. Goldin-Meadow, "Differences in early gesture explain SES disparities in child vocabulary size at school entry," *Science*, vol. 323, no. 5916, pp. 951-953, Feb. 2009.

[24] M.L. Rowe and S. Goldin-Meadow, "Early gesture selectively predicts later language learning," *Developmental Science*, vol. 12, no. 1, pp. 182-187, Jan. 2009.

[25] B.L. Smith and M.K. Kenney, "A longitudinal study of the development of temporal properties of speech production: Data from 4 children," *Phonetica*, vol. 56, no. 1-2, pp. 73-102, 1999.

[26] S. Kowal, D.C. O'Connell, and E.J. Sabin, "Development of temporal patterning and vocal hesitations in spontaneous narratives," *Journal of Psycholinguistic Research*, vol.4, no. 3, pp. 195-207, Jul. 1975.

[27] K.D. Hall, O. Amir, and E. Yairi, "A longitudinal investigation of speaking rate in preschool children who stutter," *Journal of Speech, Language, and Hearing Research*, vol. 42, no. 6, pp. 1367-1377, Dec. 1999.

[28] R.H. Pindzola, M.M. Jenkinsk, and K.J. Lokken, "Speaking rates of young children," *Language, Speech, and Hearing Services in Schools*, vol. 20, no. 2, pp. 133-138, Apr. 1989.

[29] J.F. Walker and L.M. Archibald, "Articulation rate in preschool children: a 3-year longitudinal study," *International Journal of Language & Communication Disorders*, vol. 41, no. 5, pp. 541-565, Sep. 2006.

[30] J. Sherman and E. Nicoladis, "Gestures by advanced Spanish-English second-language learners," *Gesture*, vol. 4, no. 2, pp. 143-156, Jan. 2004.

[31] I.M. Zambrana, E. Ystrom, S. Schjølberg, and F. Pons, "Action imitation at $1^1/_2$ years is better than pointing gesture in predicting late development of language production at 3 years of age," *Child Development*, vol. 84, no.2, pp. 560-573, Mar. 2013.

[32] S. Kita, M.W. Alibali, and M. Chu, "How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis," *Psychological Review*, vol. 124, no. 3, p. 245, Apr. 2017.

[33] S. Kita, "How representational gestures help speaking," in D. McNeil (Eds.), *Language and gesture*, pp. 162-185. Cambridge, MA: Cambridge University Press, 2000.

[34] M. Han, N. H., de Jong, and R. Kager, "Language specificity of infant-directed speech: Speaking rate and word position in word-learning contexts," *Language Learning and Development*, vol. 17, no. 3, pp.221-240, Jul. 2021.