



Universidad Autónoma  
de Madrid

**Biblos-e Archivo**  
Repositorio Institucional UAM

**Repositorio Institucional de la Universidad Autónoma de Madrid**

<https://repositorio.uam.es>

Esta es la **versión de autor** del artículo publicado en:  
This is an **author produced version** of a paper published in:

Language and Education (2024): 28 Feb 2024

**DOI:** <https://doi.org/10.1080/09500782.2024.2312983>

**Copyright:** © 2024 informa UK Limited, trading as Taylor & Francis Group

El acceso a la versión del editor puede requerir la suscripción del recurso

Access to the published version may require subscription

*“This is an Accepted Manuscript version of the following article, accepted for publication in Language and Education. Language and Education (2024): 28 Feb 2024. It is deposited under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.”*

**How teachers emphasize their speech: Gestures and self-repetitions during group  
interaction with toddlers**

Marta Casla

Universidad Autónoma de Madrid, Spain

ORCID ID: <https://orcid.org/0000-0002-4723-773X>

Ana Moreno-Núñez, Spain

Universidad Autónoma de Madrid, ORCID ID: <https://orcid.org/0000-0002-0397-2513>.

Florencia Alam

Centro Interdisciplinario de Investigaciones en Psicología, Matemática y Experimental  
(CIIPME), Argentina

ORCID ID: <https://orcid.org/0000-0002-3893-7873>

Celia Rosemberg

Centro Interdisciplinario de Investigaciones en Psicología, Matemática y Experimental  
(CIIPME), Argentina

ORCID ID: <https://orcid.org/0000-0002-5343-5652>.

**Corresponding Author:**

Marta Casla

[Marta.casla@uam.es](mailto:Marta.casla@uam.es)

Dpto. Psicología Evolutiva y de la Educación

Facultad de Psicología

Universidad Autónoma de Madrid

Spain

## **How teachers emphasize their speech: Gestures and self-repetitions during group interaction with toddlers**

### **Abstract**

When interacting with young children, adults often self-repeat their own utterances that vary in sequences of adjacent utterances called variation sets (VS) (Küntay & Slobin, 1996). These repetitions benefit children's linguistic development because they emphasize form and meaning. This paper analyzes the use of VS during group interaction and from a multimodal point of view. Sixteen teachers were video-recorded during interaction with two-year-old children in Spanish nursery schools. Results show that the use of VS is particularly frequent in these settings and that they are typically combined with gestures. Teachers directed their VS more often to a group of children than to a single child, and those VS directed to the group were more often accompanied by gestures than the VS directed to individuals. The results also show the influence of group size on children's responses. The study thus sheds new light on our understanding of child-directed speech (CDS), as well as the need to adapt the speech to single children during interaction in large groups.

**Keywords:** language development, early childhood education and care, multimodal communication, child-directed speech, variation sets.

## Introduction

In recent years, a number of researchers have analyzed the linguistic characteristics of child-directed speech (CDS) and its influential role in early linguistic development. The quality of CDS is defined by its quantitative and qualitative characteristics (Degotardi et al., 2018; Rowe & Snow, 2020; Soderstorm et al., 2013). Despite the increase in our knowledge of the quality of CDS in promoting children's development, most research has analyzed the linguistic characteristics of the speech directed to young children during dyadic interactions at home, while not so many studies have analyzed the specific characteristics of the speech directed to very young children during group interactions in early educational settings (Degotardi, 2021).

One of the characteristics of CDS is that it is highly repetitive (Kurchiko et al., 2020; Rowe & Snow, 2020; Tamis-LeMonda et al., 2014; Waterfall, 2006). Adults often self-repeat their own words in adjacent utterances that vary in their structure, as in the following example and as shown in Figure 1.

*Mamá: ¡La vaca!*

*Mamá: Vemos la vaquita aquí ((señala la vaca))*

*Mamá: La vaca.*

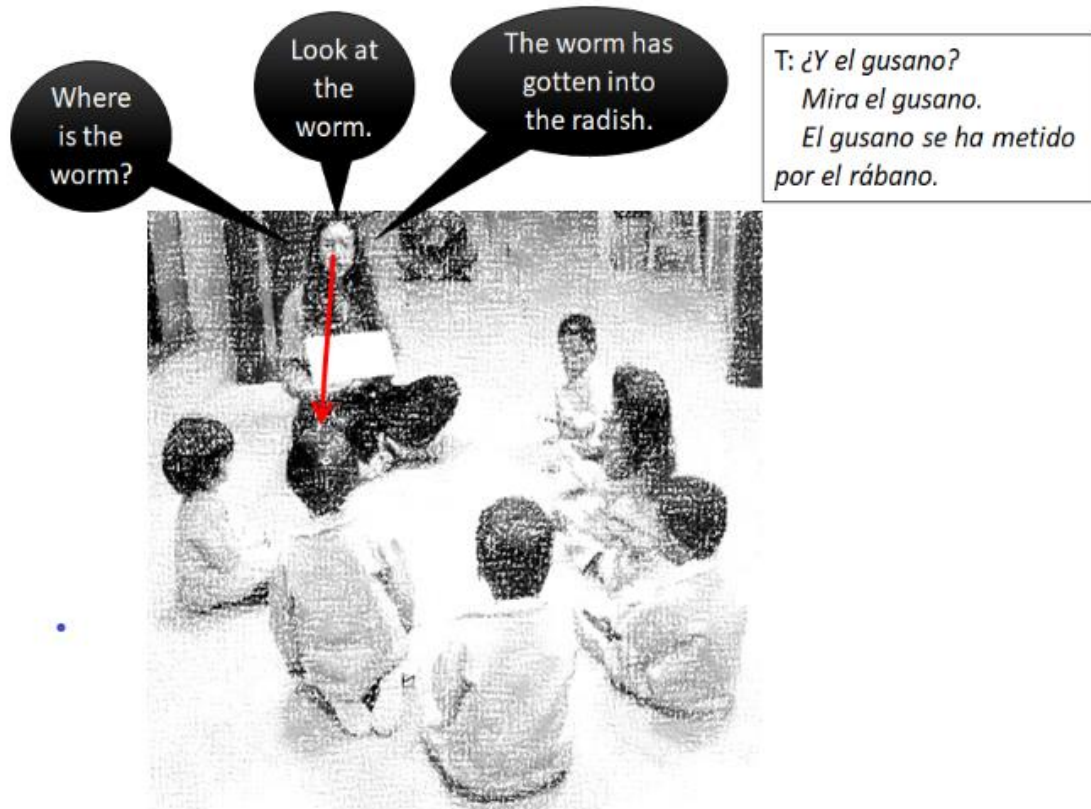
*Mother: The cow!*

*Mother: We see the little cow here ((points to the cow))*

*Mother: The cow*

## Figure 1

Example of VS during group interactions in nursery-schools



Adults' partial self-repetitions in this kind of sequence of utterances were defined by Kuntay & Slobin (1996) as variation sets (VS). Their definition highlights the fact that the same lexical item is repeated across adjacent utterances, engaged in the same event or with the same communicative intention. Furthermore, these repetitions occur within a brief time interval.

In this article, we considered the particularities of group interactions in educational contexts to study the use of VS, a characteristic of CDS that has been very well documented during dyadic interactions with middle income mothers (Alam et al., 2021; Kuntay & Slobin, 1996; Waterfall, 2006). To our knowledge, there are no studies that have specifically analyzed whether teachers use this communicative strategy when they

interact with a group of young children. At the same time, it is important to acknowledge that language development and linguistic input are not restricted to verbal behavior. Adults coordinate their speech with the gestures that they direct to young children, adapting them to the age and the linguistic level of the children (Iverson et al., 1999; Schmidt, 1996; Shaffer et al., 1983). Although non-verbal behaviors have been considered as part of the teachers' conversational interchanges (Cabell et al., 2018; Justice et al., 2018; Perry et al., 2018), co-speech gestures that accompany teachers verbal behavior have rarely been analyzed (Kidd & Rowland, 2018). Therefore, advances in the study of CDS in educational settings should take into account the use of gestures as part of the linguistic environments that young children experience in nursery schools.

The aim of this study is to analyze teachers' use of VS during group interactions with two-year-old children, adopting a multimodal point of view (i.e., analyzing co-speech gestures). In addition, we also contribute evidence that adds to the cultural diversity of the existing literature on English-speaking countries, by analyzing interactions in Spanish nursery schools.

## **Background**

When interacting with children, adults frequently self-repeat their own words and utterances, extending their own discourse in different ways that may include the use of objects or gestures (Suanda et al., 2017). Wirèn et al (2016), for example, included adults' exact self-repetitions among the utterances considered as VS. They argued that there are prosodic and non-verbal variations among verbal repetitions that introduce variations that are important in CDS.

Adults' use of VS is, therefore, an interesting topic within the operationalization of CDS and its role on children's communicative development. Caregivers' production of sequences of self-repeated utterances (i.e. using VS in their speech) has demonstrated several benefits to linguistic development, given that their recurrent structure constitutes an input that facilitates the identification of relevant lexical, morphological and syntactic features (Goldstein, 2019; Onnis et al., 2008). For instance, Waterfall et al. (2008) showed that the frequency of use of VS in adults' speech predicts children's syntactic development.

Recent research has stressed that the use of VS is not homogeneous. Rather, it changes as a function of development, and it could be related to the specific activities underway. First, from a developmental point of view, we know that the use of VS decreases with age (Lester et al., 2022; Waterfall, 2006) especially after the age of two (Wirén et al., 2016). This could be because, from this point on, children's linguistic and pragmatic skills make them less in need of contingent repetition. Furthermore, the fact that children increase their participation during conversational interchanges generates an increase in adults' retakes of children's speech and a subsequent decrease in self-repeated utterances (Waterfall, 2006). Second, in a study with Spanish speaking Argentinian children at 14 months of age, Alam et al. (2021) analyzed the speech they received during child-centered social activities, such as book-reading or play interactions, and showed how these scenarios increased the probability of listening to utterances included in VS. In addition, the study found that the use of VS in CDS is rather frequent since, overall, 20% of the utterances that were directed to each child were included in VS. This proportion raised to 35% in middle-class mothers and during child-directed activities, similarly to what was observed in Hebrew and English-speaking children (Tal & Arnon, 2018).

While the use of VS is well documented in the literature related to CDS, there are no studies yet that specifically identify and analyze this form of self-repetition in nursery school settings, which could be particularly informative for characterizing young children's linguistic experiences. Research on the use of CDS in early childhood education settings has traditionally attempted to identify the communicative dimensions that are most strongly associated with children's linguistic growth over time (Justice et al., 2018). Several studies have analyzed the quantitative and qualitative characteristics of teachers' talk in nursery schools and kindergartens, including their use of a number of verbal strategies, such as repetitions, recasts, expansions, labeling, questions, or modeling of new vocabulary (Kidd & Rowland, 2018). Teacher's use of these strategies sophisticates their speech and is related to children's linguistic achievements that occur later in development (Dickinson & Porsche, 2011; Girolametto et al., 2003; Hindman et al., 2021; Wasik et al., 2006).

Considering that the characteristics of CDS change across children development (Rowe & Snow, 2020), some of them could have a stronger influence during the early stages of language and communication, especially those that highlight the linguistic structure –or some parts of it– and promote the participation of children (Huttenlocher et al., 2002).

In this sense, everyday group activities, such as shared book reading, may foster teachers' use of certain strategies aimed at maintaining the attention from the group, contributing to children development of key linguistic abilities. For instance, studies suggest that the interdependency among developing competences may explain how teachers' use of attention-getters increase children's probability of expanding their receptive vocabulary, which is key in the subsequent development of reading comprehension (Dickinson & Porsche, 2011).



Prior research has documented that some forms of repetition and self-repetition are typically used by early education teachers during their interactions with young children (Hindman et al., 2021; Girolametto et al., 2003). For example, they use rhymes, routines and songs as part of daily interactions in the classroom, which includes self-repetitions of words and utterances in contexts that are meaningful for the children. While routines and songs constitute only a small portion of teachers' talk (Kidd & Rowland, 2018), they are a way of emphasizing the form of the words, similar to how self-repeated utterances typically characterize CDS. However, even when VS has been shown to be an effective attention-getter during dyadic interactions (Küntay & Slobin, 2002), their specific use has not been explored in early childhood education settings. Their recurrent structure could equally contribute to get children's attention during group interactions, while repeating the same word in different syntactic frames increases the chances of children hearing that word and bootstrapping its meaning (Tal & Arnon, 2018).

Some recent research has also pointed out that the effectiveness of CDS depends to a large extent on children's linguistic experiences and their opportunities for participation (Justice et al., 2018). During group interactions, teachers frequently direct their speech to more than one child, leading to diverse experiences both in quantitative and qualitative terms (Chaparro-Moreno et al., 2018; Perry et al., 2019; Soderstrom & Witterbolle, 2013). These differences could also depend on the type of activity and other structural factors such as the teacher-child ratio. In this sense, structured playtime activities, including book reading instances, typically concentrate higher proportions of teachers' linguistic utterances (Soderstrom & Witterbolle (2013), pointing to the potential of their analysis and characterization as a supportive context for language development.

Furthermore, there is also a need to consider teacher-child ratios, as it may help explaining why some children experience fewer opportunities to interact with the teacher. Soderstrom et al. (2018) found an indirect relationship between the adult-child ratio in the classroom and the number of words that children heard. Interestingly, the total number of adults in the classroom also had an influence on increasing the opportunities for direct teacher-child interactions. Similarly, Degotardi et al. (2016) demonstrated that the number of words that children heard in the classroom was influenced by the teacher-child ratio, as well as it was mediated by the general quality of the interaction, that the authors measured through a standardized scale. The findings from these studies highlight the interest in studying the relationship between the number of adults and children present in the classroom and the specific characteristics of teachers' talk. This should account too for cultural variations in adult-child interactions (Kuchirko et al., 2020), especially in countries in which the maximum number of children enrolled per classroom is relatively high.

For decades, numerous researchers have claimed the need to additionally analyze the multisensory nature of the speech directed to children, as well as the coordination between verbal and non-verbal behaviors in CDS (Gogate et al., 2015; Rodrigo et al., 2006; Suanda et al., 2017). For example, Rodrigo et al. (2006) evidenced that mothers of 1- and 2-year-olds adapt their gesture production to their children's developmental level, exaggerating their actions as compared to adult-directed gestures, in a way that highlights and directs the child's attention to particular words and meanings (Iverson, 1999). Similar functions have been reported for verbally analyzed VS (Alam et al., 2021), but the extent to which non-verbal behaviors (e.g., gestures) are coordinated with repeated verbal structures as in a VS remains unanswered in relation to group interactions in early educational settings.

Kidd & Rowland (2018) found that practitioners that received specific training in language development, produced more co-speech gestures during their dyadic interactions with two- and three-year-old children than practitioners who did not receive such training, although these differences did not reach significance. To our knowledge, however, there are no studies that explicitly analyze the use of co-speech gestures during group interactions with two-year-old children at nursery schools, which could contribute to depicting and understanding how and why early education teachers spontaneously coordinate their verbal speech in a multimodal fashion, especially with gestures.

### ***Rationale and contextualization***

For all these reasons, the overall purpose of this study was to contribute a new approach for establishing links among three main topics on CDS research: the use of VS as a characteristic of CDS, the way in which CDS is used in early childhood education, and the multimodal components that could accompany teachers' speech.

The present study had four specific research goals, which involved the examination of:

1. The extent to which teachers' verbal utterances were included in VS.
2. The multimodal nature of teachers' use of VS.
3. Whether there is a relationship between the frequency of use of VS (both verbal and verbal-gestural) and the addressee of the interaction (dyadic or polyadic).
4. Whether there is a relationship between the frequency of use of VS and the teacher-child ratio.

Following Huttenlocher et al. (2002) we assumed that children under four years of age may only process the speech that is specifically directed to them. However, adults –

including teachers— are very good at identifying their conversational partners’ skills and adapting their speech to their level, especially with children (Rowe & Snow, 2020).

Therefore, we expected teachers’ speech to be different when they addressed a single child or the group. Consequently, to account for their particularities when depicting the interactions, we differentiated between dyadic (i.e. directed to a single child) and polyadic (i.e. directed to the whole group) communicative behaviors from teachers.

While we acknowledge that the speech directed to a single child during a group activity is not purely ‘dyadic’, we adopted this term to differentiate the children exposure to teachers’ direct and indirect speech within the same group.

According to the literature, the percentage of VS in the speech of mothers of two-year-old children is less than 20% (Alam et al., 2021; Wirén et al., 2016). However, CDS in classroom settings differs from that at home, especially due to the number of children that are present. Thus, we expected teachers to use higher proportions of VS in order to get children to do or to attend to something during the ongoing activity. Group interactions also require participants to use a number of strategies to obtain attention from their communicative counterparts (Dickinson & Porsche, 2011; Soderstrom & Wittebolle, 2013), where gestures may be particularly useful to capturing children’s attention (Schmidt, 1996; Shaffer et al., 1983). We thus expected higher frequencies of utterances accompanied by gestures within VS, than outside them, and teachers using more often VS to the group than to addressing a single child (i.e., that polyadic VS would be used more frequently than dyadic VS), both for verbal and verbal-gestural VS. Finally, we expected higher rates of VS and, specifically, verbal-gestural VS in classrooms with more children, as compared to classrooms with lower ratios (Degotardi et al., 2016).

**Method**

***Participants***

This study involved children and teachers in 16 classrooms of 7 nursery schools, located in middle income urban areas in Madrid (Spain). The participants were 162 children enrolled in the two-to-three-year-old groups of the nursery schools ( $M = 29.16$ ;  $SD = 3.92$ ). The current Spanish legislation allows enrolling up to 20 two-year-old children per classroom with a single lead teacher, although not every classroom reach this peak ratio. However, this means that the number of children per classroom varies from one school to another (Table 1). The mean number of children per classroom in our sample was  $n = 10.06$ .

**Table 1**

*Number of children enrolled in each classroom*

	Classroom ID															
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
n	8	6	7	15	9	6	14	12	12	7	7	12	14	10	15	8

***Procedures***

The nursery schools were selected from the database of schools that had participated in the previous studies of the research team, according to their availability for this study, the number of children in each classroom, and ensuring a representation of both private and public providers. Informed consents were collected from the teachers and children's primary caregivers. Children whose families declined their participation ( $n = 4$ ) joined another classroom during the filming sessions.

Teachers and children were video recorded during “circle time”, which is defined as a daily activity that promotes communication and group bonding. Although recent research has shown that there is a strong variation in the effects of the type of activity in the quality of classroom interactions with 3-year-olds (Kook & Greenfield, 2020), Chaparro-Moreno et al. (2019) evidenced that the interactions between three- to- four-year-old children and their teachers were more homogeneous and more frequent during circle time than in other activities during the school day. In addition, the structure of circle time is very similar among schools (usually including individual and group activities, rounds of presentations, singing, storytelling or book-reading, as well as several daily routines), facilitating the comparisons between the 16 classrooms in our sample.

The recording sessions took place after the adaptation period, between December 2019 and January 2020. In order to familiarize the children with the researchers and the recording materials, we initially conducted a participant observation session, during which we placed the tripods without cameras where they were supposed to be when we conduct the filming. A researcher observed the classroom dynamics from a few minutes before and after circle time and took field notes about the way in which circle time started, whether all the children were present or not during the whole activity, how were children arranged around the teacher, and other relevant environmental factors such as the noise level, the amount of light, or whether there was any obstructing furniture.

We set up two cameras during the recording session, one towards the teacher and the other towards the children. Teachers wore a wireless microphone as well that were connected to the camera focusing them. They were instructed to interact with the children as they would usually do during circle time. The filmed circle time activities led to more than 4 hours of footage. The duration of the circle time activities varied between

classrooms; with a mean of 18 min 8 sec ( $SD = 6$  min 37 sec; Minimum = 8 min 27 sec; Maximum = 34 min 0 sec).

### ***Coding***

All videos were transcribed and coded in ELAN (Lausberg y Sloetjes, 2009). Teachers' verbal speech was transcribed following the conventions of the CHAT program from the CHILDES project (MacWhinney, 2000), so we divided teachers' speech into utterances that coincided with the segments in ELAN. We used CLAN programs from the CHILDES project to automatically calculate teachers' Mean Length of Utterance (MLU) and the ratio of types/tokens using the index D for Lexical Diversity (see Hindman et al., 2021 for a similar analysis).

The variables used in this study are described in Table 2. We coded the frequency of use of utterances, gestures and VS observed in teachers' speech. Following Bernstein Ratner & Brundage (2015), we included all *utterances* in the same conversational turn that met the intonation syntactic completion and semantic unity criteria. Also, inter-word pauses contained in an utterance should be less than 2 seconds. *Gestures* were operationalized as all motor actions used to convey meaning to the conversational partners and may include the use of objects (Murillo & Belinchón, 2012; Rodrigo et al.; 2016). For the purposes of this investigation, we did not distinguish between the types of gestures (i.e., deictic, conventional, representational, etc.).

For the operationalization of *variation sets* (VS) we followed Alam et al. (2021) and Wirén et al. (2016). According to these authors, VS were defined as two or more adjacent utterances that repeated some part or its whole structure, and that conveyed the same meaning. Note that this definition is based on the transcribed verbal behavior, while we

also analyzed other non-verbal information from the videos. As in prior research, we considered an utterance as partially or fully self-repeated as long as there were less than 5 seconds between the first utterance and its repetition (Ko et al., 2016). Longer pauses may result in children not being able to connect the sequence of repeated utterances as a unity. Additionally, we considered that a VS could also include (a) short utterances that did not affect the meaning of the preceding and following utterances and (b) short child-interventions that were ignored by the adult. This was a particularly important criteria during polyadic interactions, since the bigger the group, the better probability for children's interventions. Because the songs, rhymes and routines posed by teachers were also followed by some children, we considered them too as part of VS.

**Table 2**

*Adults' behavioral codes used in this study*

Main codes	Secondary codes	Descriptions
<b>Modality</b>	<i>Verbal utterance</i>	An utterance that is verbally produced.
	<i>Verbal-gestural utterance</i>	An utterance that at least includes one gesture coordinated with the speech, lasting for more than 0.5 seconds.
<b>Addressee</b>	<i>Dyadic</i>	The utterance is addressed to a single child.
	<i>Polyadic</i>	The utterance is addressed to a group of children.
<b>Type of VS</b>	<i>Verbal VS</i>	A VS in which all the utterances included are verbal.
	<i>Verbal-gestural VS</i>	A VS in which at least one utterance included is verbal-gestural.
	<i>Dyadic VS</i>	A VS in which all the utterances included were addressed to the same child.
	<i>Polyadic VS</i>	A VS in which all the utterances included were addressed to the group.
	<i>Mixed VS</i>	A VS that combines utterances addressed to a single child and to the group.



An additional researcher coded 15% of randomly selected video clips. We calculated intercoder reliability for the identification of VS (93.7% agreement,  $K = .793$ ;  $p < .001$ ), gestures (88.5% agreement,  $K = .746$ ;  $p < .001$ ) and the addressee of the teachers' speech (81.0% agreement,  $K = .793$ ;  $p = .048$ ).

## **Results**

For simplicity, the results are reported following the four specific research goals of this study. We initially analyzed the extent to which teachers' verbal utterances were included in VS, comparing their proportions with the rest of the speech they produced, and characterizing the multimodal nature of the VS that we identified. We then explored the relationship between the frequency of use of VS and some structural variables, in particular considering the addressee of the interaction, and the teacher-child ratio.

### ***Proportions of teachers verbal utterances included in VS***

We calculated the proportion of utterances that were included in VS over the total number of utterances produced by a particular teacher and compared it with the proportion of utterances in the rest of the speech. Overall, the mean proportion of utterances included in VS was 0.59 (Minimum = 0.43, Maximum = 0.85), leaving a mean proportion of utterances in the rest of the speech of 0.41 (Minimum = 0.15, Maximum = 0.57). Chi-square tests revealed no significant differences.

### ***Multimodal characterization of teachers' use of VS***

We first calculated the proportion of verbal and verbal-gestural utterances produced by teachers during the whole session, including those utterances contained in VS. Table 3 shows the mean proportions, as well as minimum and maximum values. Verbal-gestural

utterances represent a 23% of the overall speech that teachers directed to children, evidencing a higher proportion of verbal utterances in our sample. We calculated a repeated measures ANOVA to compare mean proportions, while computing log transformations to ensure the normality of the distribution, and confirmed that this difference was significant ( $F(1,15) = 237.799$ ;  $p < .001$ ,  $\eta^2 = .941$ ).

**Table 3**

*Descriptive statistics of teachers' verbal and verbal gestural utterances*

	Mean	Minimum	Maximum
<b>Verbal</b>	.76	.64	.86
<b>Verbal-gestural</b>	.23	.14	.36

Then, we calculated the proportion of verbal and verbal-gestural utterances that were produced within VS, taking the total number of utterances contained in VS as the denominator. We found a similar trend that we observed for overall proportions, where a 25.2% of the utterances contained in VS were verbal-gestural. As shown in Table 4, the proportion of verbal VS was higher than that of verbal-gestural VS. We conducted a repeated measures ANOVA to compare these differences, also calculating log transformations to ensure the normality of the distribution, confirming they were significant ( $F(1,15) = 161.028$ ;  $p < .001$ ,  $\eta^2 = .915$ ).

**Table 4**

*Mean proportion of teachers' verbal and verbal-gestural utterances included in VS*

	Mean	Minimum	Maximum
<b>Verbal VS</b>	.75	.70	.79

Verbal-gestural VS	.25	.29	.21
--------------------	-----	-----	-----

372

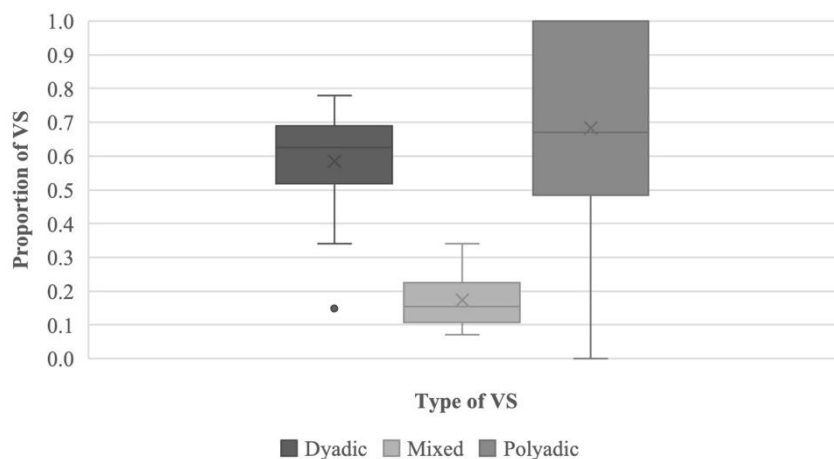
373 Additionally, for comparing whether VS were more or less multimodal than the rest of  
374 the speech, we initially calculated the proportion of verbal and verbal-gestural  
375 utterances in the speech directed to children, excluding those utterances contained in  
376 VS. We then compared the proportions of each type of utterance and tested whether the  
377 VS were more likely to include verbal-gestural utterances than the rest of the speech  
378 directed to children. We calculated a repeated measures ANOVA using the VS as the  
379 factor (i.e. whether it was included or not in VS), and the proportion of verbal-gestural  
380 utterances as the dependent variable. Results showed a higher proportion of verbal-  
381 gestural utterances included in VS ( $M = 0.25$ ) than its proportion in the rest of the  
382 speech ( $M = 0.24$ ), and that this difference was significant ( $F(1,15) = 5,658; p = .031, \eta^2$   
383  $= .274$ ).

#### 384 ***Relationship between the use of VS and the addressee of teachers' speech***

385 Considering VS as the unit of analysis, regardless the number of utterances that they  
386 included, we calculated the proportion of dyadic, polyadic and mixed VS that were  
387 identified over the total VS production. Figure 2 shows higher mean proportions of  
388 polyadic VS (0.58) than those of dyadic VS (0.24) and mixed VS (0.17).

#### 389 **Figure 2**

390 *Mean proportion of each type of VS according to the addressee*

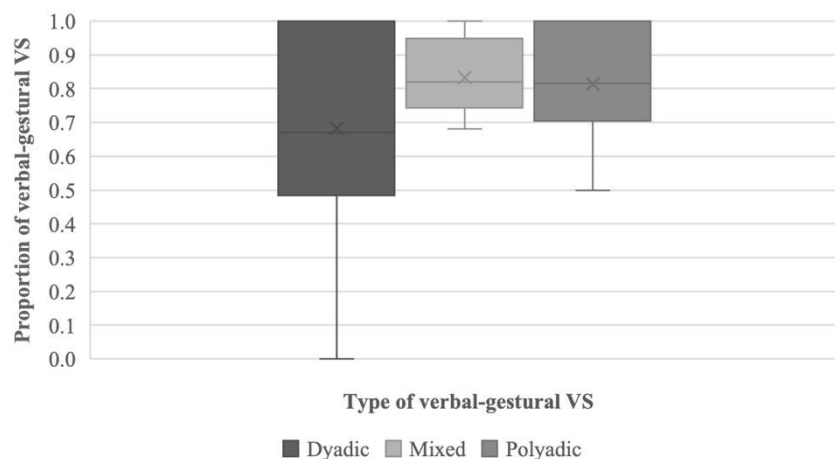


391

392 We conducted a repeated measures ANOVA to test whether the observed differences  
 393 were significant, using the type of VS (i.e., dyadic, polyadic and mixed) as the factor,  
 394 and the proportion of each category over the total number of VS as the dependent  
 395 variable. We calculated log transformations to ensure the normality of the distribution.  
 396 Results showed a main effect of the type of interaction ( $F(2,14) = 38,659; p < .001, \eta^2 =$   
 397  $.847$ ), so we proceeded to calculate lower level comparisons. Polyadic VS were  
 398 significantly more frequent than dyadic ( $p = .002$ ) and mixed VS ( $p < .001$ ). Although  
 399 dyadic VS were more frequent than mixed VS, this difference did not reach  
 400 significance. We also tested to what extent verbal-gestural VS were addressed to a  
 401 single child or to the group, by calculating the proportion of each type of verbal-gestural  
 402 VS (Figure 3). Although the particular number of verbal-gestural utterances included in  
 403 VS was significantly lower than solely verbal utterances, the percentage of VS that  
 404 contain at least one utterance coordinated with gestures ranged from 68% to 83% of the  
 405 whole VS sample.

### 406 **Figure 3**

407 *Mean proportion of verbal-gestural VS according to the addressee*



408

409 We tested the differences between dyadic, polyadic and mixed verbal-gestural VS  
 410 through a repeated measures ANOVA using as factors the proportion of each type of  
 411 verbal-gestural VS (i.e., dyadic, polyadic and mixed). Results showed an effect of the  
 412 type of VS ( $F(2,14) = 3,62; p = .039, \eta^2 = .194$ ), so we calculated lower level  
 413 comparisons that confirmed that the proportion of polyadic verbal-gestural VS were  
 414 significantly higher than that of dyadic verbal-gestural VS ( $p = .024$ ). The differences  
 415 between polyadic and mixed verbal-gestural VS did not reach significance, nor did the  
 416 differences between dyadic and mixed verbal-gestural VS.

#### 417 ***Relationship between the use of VS and the teacher-child ratio***

418 We tested the relationship between both the frequency of use of VS and each of their  
 419 types and the teacher-child ratio by conducting a series of bivariate correlations between  
 420 the variables related to VS: the overall proportion of VS; the proportion of verbal and  
 421 verbal-gestural VS; the proportion of dyadic, polyadic and mixed VS (both verbal and  
 422 verbal-gestural); and the number of children in the classroom. We found a positive and  
 423 significant correlation between the number of children and the overall proportion of VS  
 424 ( $r(16) = .538, p = .032$ ) and that of mixed VS ( $r(16) = .588, p = .017$ ). This finding

suggests that the more children that were present in the classroom, the better probability for teachers to use VS, and that this is both a general trend and one focused on combining utterances addressed single children and the whole group within the same self-repeated sequence of utterances. Furthermore, the number of children in the classroom significantly correlated with the use of mixed verbal-gestural VS, supporting the hypothesis that with larger groups of children teachers tend to produce more VS coordinated with gestures to address both a single child and the whole group within the same sequence.

## **Discussion**

The results of the present study showed that Spanish nursery school teachers frequently self-repeat their own utterances in forms of variation sets when they are interacting with two-year-old children. This finding is aligned with studies on dyadic interactions at home, and those showing that teachers self-repeat their speech when interacting with toddlers, by using rhymes, routines and songs (Chaparro-Moreno et al., 2019; Kidd & Rowland, 2018). Some of these studies also stressed that the use of VS at home only constitutes a 20% of the overall speech that middle-class mothers direct to their 18-month-old children (Alam et al., 2021; Tal & Arnon, 2018), and that it decreases as children approach their second birthday (Wirén et al., 2016). Interestingly, teachers' self-repetitions in forms of VS nearly reach a 60% of the utterances in our sample, suggesting that during circle time teachers tend to be more repetitive than what was initially thought.

This study showed that the use of VS is also a key characteristic of the speech directed to groups of toddlers in early childhood education classrooms. The high proportion of VS in teachers' speech reveals the need to consider them as a frequent communicative

strategy, at least during circle time activities with children around two years of age. Furthermore, the small proportion of rhymes and routines that previous studies found (Kidd & Rowland, 2016) may suggest that some of the functions of these strategies are carried out through the use of VS. In fact, previous studies analyzing the use of VS in dyadic interactions at home, showed that these repetitions allow the adult to direct the attention of the child to those linguistic forms that were repeated (Kuntay & Slobin, 2002), increasing the children exposure to those linguistic forms (Onnis et al., 2008). Teachers could frequently use VS as a way to ensure that children are listening to them, and to emphasize those parts of the message that require more attention, as occurs with other attention-getter strategies in interactions with older children (Dickinson & Porshe, 2011).

Regarding the multimodal nature of teachers' uses of VS, our results showed that they produced more utterances accompanied by gestures within VS than in the rest of their speech, aligned with previous studies with dyads evidencing that mothers direct more verbal than verbal-gestural utterances to one and two-year-old children (Casla et al., 2021; Iverson et al., 1999). Nonetheless, when we considered VS as the unit of analysis (i.e. instead of taking the number of utterances that constitute them), we found that most of the VS identified were accompanied by the use of gestures. In particular, more than 80% of the polyadic and mixed VS included at least one gesture. Both the use of VS and the use of gestures in CDS have been proven to be clearly effective strategies for children's linguistic development (Rowe & Goldin-Meadow, 2009; Tal & Arnon, 2018; Tamis-LeMonda et al., 2012), supporting the implications that our study could bring for nursery school teachers. As these authors suggested, the combination of verbal repetitions and gestures may serve the same functions of VS, where teachers accentuate a message that was already been stressed through verbal repetition (Casla et al., 2021).

Attention-getters may promote children's active participation as well during certain activities, such as storytelling (Dickinson & Porsche, 2011). Therefore, teachers explicit use of verbal-gestural VS during their interaction with very young children could serve to meet this goal, especially in those contexts in which participation gets hindered.

The findings of this study also showed that teachers did not address their repeated utterances to single children as often as they did to the group, since most of the VS identified were polyadic. Moreover, when teachers used VS combined with gestures, they are more frequently addressed to the group than to a single child, suggesting again that teachers could consider co-speech gestures as a way to stress their verbal messages.

Teachers also produced VS that combined utterances directed to a single child and to the group. When analyzing these sequences from a multimodal point of view, we found that mixed and multimodal VS were more frequent than dyadic VS, and as frequent as polyadic VS. This means that VS combined with gestures were preferably used by teachers when they addressed the group, highlighting certain differences in the speech directed to the group and to a single child even in the same classroom and during the same activity. These findings suggest that teachers may need to repeat more often and use more gestures when they are addressing the group, which is compelling to understand how the participation of children is organized in conversational sequences (Chaparro-Moreno et al., 2019; Perry et al., 2018; Piasta et al., 2012). Interestingly, very few related studies have differentiated between dyadic and polyadic speech.

While we acknowledge the difficulties involved in examining to what extent two- and three-year-olds may process this group-directed speech, further research should at least consider the differences between the amount of speech that they receive both in one-to-one and polyadic interactions. Teachers are more likely to use VS during polyadic than



during dyadic interactions, as well as verbal-gestural VS are more often addressed to the group than to a single child. Therefore, there could be an spontaneous adaptation of the teachers' speech not just to the linguistic level of the child, but also to other social characteristics of the interaction. This result has the potential to inform training programs for early education teachers revolving the development of language and communication.

Regarding the relationship between the teachers' use of VS and the teacher-child ratio, we observed differences between dyadic and polyadic VS that seem to be related to the number of children that were present in the classroom. We found weak but significant correlations between the number of children and the frequency of dyadic VS. The number of children also correlated with the frequency of mixed VS combined with gestures. Altogether, these findings suggest that in classrooms with larger groups of children teachers frequently self-repeat their utterances to each child, and accompanied them by gestures, especially when switching from addressing one child to the group, or from addressing the group to a single child. Results from the correlations analysis supported as well that teachers produce more VS in classrooms with larger groups of children. Note that the smallest ratio in our study was 1:6, and that there was a single teacher in all classrooms during the circle time observed. This means that every teacher in our study had to distribute their speech among several children. Interestingly, the use of gestures were less frequent when teachers use VS to address larger groups of 12 to 16 children, than when they were directed to single children or smaller groups. Multimodal VS may be useful with larger groups of children to gather the attention of specific children and to stress some parts of the linguistic message, even when some of the utterances are switched sometimes to the group as in mixed VS. These results are aligned with previous research conducted in nursery schools and in other countries,

suggesting that classroom ratios are not a unique influential factor on teachers' selecting specific communicative strategies (Dalgaard et al., 2022; Soderstrom et al., 2018). Notwithstanding, classroom ratios are an important variable to consider in further research of the characteristics of CDS and its particularities in nursery schools settings (Degotardi, 2021; Degotardi et al., 2018).

All in all, this study yields new data from Spanish nursery schools that evidenced VS as an important characteristic of the speech that teachers direct to two-year-old children. Furthermore, we provide new data showing that VS are frequently combined with gestures in early childhood education settings. These findings highlight the need to consider the coordination between verbal and non-verbal behaviors in teacher training programs. Additionally, we have shown that circle time is not just an activity that concentrates a considerable number of teachers' verbal utterances (Chaparro-Moreno et al., 2019; Soderstrom & Wittenbolle, 2013), but also that a large proportion of them are coordinated with gestures.

Our study also stressed the importance to analyze group interactions and to also consider children's behavior when the speech is not directly addressed to them. Children's experiences may differ as a function of the distribution of teachers' speech and as a function of the teacher-child ratio. It is thus important to recognize circle time as an activity that promotes linguistic socialization (Poveda, 2003), since children are exposed to different forms of speech, whether they are explicitly directed to them or not (i.e., group directed utterances). Besides, during the same activity children are also exposed to the speech directed to other participating children (i.e., teachers' talk to another child).

Together, the differences between dyadic and polyadic VS and the relationship between their production and some structural factors of the classroom provided relevant information on the potential effects of CDS in educational contexts. While previous research failed to find direct relationships between children's linguistic skills and teachers' specific linguistic and communicative strategies (Hindman et al., 2021; Justice et al., 2018), their impact in the children's linguistic growth should be considered in terms of the promotion of their active participation (Cabell et al., 2018; Dickinson & Porche, 2011; Perry et al., 2018; Piasta et al., 2012). Although analyzing children's specific responses to VS were beyond the scope of this particular study, further research should examine such responses and whether teachers' self-repetitions contribute to children's attention and participation. In addition, other related variables, such as the pragmatic functions of teachers' VS or the visual feedback from children, may be also of interest. This research was limited too by the fact that we could not include longitudinal data, which could have allowed us to track the changes in the use of VS over the school year and evaluate if, as it was found in home interactions, it tends to decrease as children grow older (Waterfall, 2006; Wirèn et al., 2016). The high rates of VS that we observed during polyadic interactions should impulse new research to understand how these behaviors evolve through different timescales, especially as children increase their communicative and linguistics abilities. This study could set a new starting point in the study of multimodality and CDS in nursery schools, including the investigation of the use of gestures during different activities. Exploring how teachers combine verbal and multimodal communicative resources during their everyday interactions with toddlers can ultimately contribute to our understanding of the role of CDS in shaping children's linguistic development beyond dyadic interactions.

## References

- 571 Alam, F., Rosemberg, C., Garber, L., & Stein, A. (2022). Variation sets in the speech  
572 directed to toddlers in Argentinian households. SES and type of activity effects.  
573 *Journal of Child Language*, 49(4), 799-823.  
574 <https://doi.org/10.1017/S030500092100043X>
- 575 Bernstein Ratner, N., & Brundage, S. (2015). *A clinician's complete guide to CLAN and*  
576 *PRAAT*. Talkbank Organization: <https://talkbank.org/manuals/Clin-CLAN.pdf>
- 577 Cabell, S.Q., Justice, L.M., McGinty, A.S., DeCoster, J., & Forston, L.D. (2015).  
578 Teacher-child conversations in preschool classrooms: Contributions to children's  
579 vocabulary development. *Early Childhood Research Quarterly*, 30, 80-92.
- 580 Chaparro-Moreno, L.J., Justice, L.M., Logan, J.A.R., Purtell, K.M., & Lin, T.J. (2019)  
581 The preschool classroom linguistic environment: Children's first person  
582 experiences. *PLoS ONE* 14(8): e0220227.  
583 <https://doi.org/10.1371/journal.pone.0220227>
- 584 Dalgaard, N. T., Bondebjerg, A., Klokke, R., Viinholt, B. C. A., & Dietrichson, J.  
585 (2022). Adult/child ratio and group size in early childhood education or care to  
586 promote the development of children aged 0–5 years: A systematic review.  
587 *Campbell Systematic Reviews*, 18, e1239. <https://doi.org/10.1002/cl2.1239>
- 588 Degotardi, S. (2021). The language environment of infant child care: Issues of quantity,  
589 quality, participation and context. In O. Saracho (2021, Ed.) *Contemporary*  
590 *perspectives on research on child care in Early Childhood Education* (pp. 85-  
591 107). Information Age Publishing.

- Degotardi S, Han F, Torr J. (2018). Infants experience with near and clear educator talk: individual variation and its relationship to indicators of quality. *Int J Early Years Educ.* 26(3):278–294. <https://doi.org/10.1080/09669760.2018.1479632>
- Degotardi, S., Torr, J., & Nguyen, N. (2016). Infant-toddler educators' language support practices during snacktime. *Australasian Journal of Early Childhood Education*, 41(1), 52-62.
- Dickinson DK, Porche MV. 2011. Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Dev.* 82(3):870–886. <https://doi.org/10.1111/j.1467-8624.2011.01576.x>
- Gogate L, Maganti M, Bahrick LE. 2015. Cross-cultural evidence for multimodal motherese: Asian Indian mothers' adaptive use of synchronous words and gestures. *J Exp Child Psychol.* 129:110–126. <https://doi.org/10.1016/j.jecp.2014.09.002>
- Girolametto, L., Weitzman, E., & Greenberg, J. (2003). Training day care staff to facilitate children's language. *American Journal of Speech-language Pathology*, 12(3), 299–311. [https://doi.org/10.1044/1058-0360\(2003/076\)](https://doi.org/10.1044/1058-0360(2003/076)).
- Hindman, A., Farrow, J., Anderson, K., Wasik, B., & Snyder, P. (2021). Understanding Child-Directed Speech Around Book Reading in Toddler Classrooms: Evidence From Early Head Start Programs. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.719783>

- 612 Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S.C. (2002). Language input  
613 and child syntax. *Cognitive Psychology*, 45, 337-374.  
614 [https://doi.org/10.1016/S0010-0285\(02\)00500-5](https://doi.org/10.1016/S0010-0285(02)00500-5)
- 615 Iverson, J. M., Capirci, O., Longobardi, E., & Caselli, M. C. (1999). Gesturing in  
616 mother-child interactions. *Cognitive Development*, 14(1), 57-75.
- 617 Justice, L. M., Jiang, H., & Strasser, K. (2018). Linguistic environment of preschool  
618 classrooms: What dimensions support children's language growth. *Early*  
619 *Childhood Research Quarterly*, 42(1), 79-92.  
620 <https://doi.org/10.1016/j.ecresq.2017.09.003>
- 621 Kidd, L., & Rowland, C. (2018). The effect of language-focused professional  
622 development on the knowledge and behaviour of preschool practitioners.  
623 *Journal of Early Childhood Literacy*, 21, 1-33.  
624 <https://doi.org/10.1177/1468798418803664>
- 625 Ko, E. S., Seidl, A., Cristia, A., Reimchen, M., & Soderstrom, M. (2016). Entrainment  
626 of prosody in the interaction of mothers with their young children. *Journal of*  
627 *Child Language*, 43(2), 284-309. <https://doi.org/10.1017/S0305000915000203>
- 628 Kook, J. F., & Greenfield, D. B. (2021). Examining variation in the quality of  
629 instructional interaction across teacher-directed activities in head start  
630 classrooms. *Journal of Early Childhood Research*, 19(2), 128–144.  
631 <https://doi.org/10.1177/1476718X20942956>

- 632 Kuchirko, Y., Schatz, J., Fletcher, K., & Tamis-LeMonda C. (2020). Do, say, learn: The  
 633 functions of mothers' speech to infants. *Journal of Child Language*, 47(1), 64-  
 634 84. <https://doi.org/10.1017/S0305000919000308>
- 635 Küntay, A., & Slobin, D. I. (1996). Listening to a Turkish mother: Some puzzles for  
 636 acquisition. In D.I. Slobin, J. Gerhardt, A. Kyratzis, & J. Guo (Eds.) *Social*  
 637 *interaction, social context, and language: essays in honor of Sudan Ervin Tripp*  
 638 (pp. 265-86). Lawrence Erlbaum Associates.
- 639 Küntay, A., & Slobin, D. I. (2002). Putting interaction back into child language:  
 640 Examples from Turkish. *Psychology of Language and Communication*, 6(1), 5-  
 641 14.
- 642 Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-  
 643 ELAN system. *Behavior Research Methods*, 41(3), 841–849. [https://doi.org/](https://doi.org/10.3758/BRM.41.3.841)  
 644 [10.3758/BRM.41.3.841](https://doi.org/10.3758/BRM.41.3.841).
- 645 Lester NA, Moran S, Küntay AC, Allen SE, Pfeiler B, Stoll S. 2022. Detecting  
 646 structured repetition in child-surrounding speech: evidence from maximally  
 647 diverse languages. *Cognition*. 221:104986.  
 648 <https://doi.org/10.1016/j.cognition.2021.104986>
- 649 MacGillion, M.L., Herbert, J.S., Pine, J.M., Keren-Portnoy, T., Vihman M.M., &  
 650 Matthews D.E. (2013). Supporting Early Vocabulary Development: What Sort  
 651 of Responsiveness Matters?. *IEEE Transactions on Autonomous Mental*  
 652 *Development*, 5(3), 240-248.

- 653 MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Lawrence  
654 Erlbaum Associates.
- 655 Murillo, E., & Belinchón, M. (2012). Gestural-vocal coordination: Longitudinal changes  
656 and predictive value on early lexical development. *Gesture*, 12(1), 16-39.  
657 <https://doi.org/10.1075/gest.12.1.02mur>
- 658 Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning  
659 language from variation set cues. *Cognition*, 109(3), 423-430.
- 660 Perry, L. K., Prince, E. B., Valtierra, A. M., Rivero-Fernandez, C., Ullery, M. A., Katz,  
661 L. F., ... Messinger, D.S. (2018). A year in words: The dynamics and  
662 consequences of language experiences in an intervention classroom. *PLoS One*,  
663 13(7), e0199893. <https://doi.org/10.1371/journal.pone.0199893>
- 664 Piasta, S.B., Justice, L.M., Cabell, S.Q., Wiggins, A.K., Turnbull; K.P., & Curenton,  
665 S.M. (2012) Impact of professional development on preschool teachers'  
666 conversational responsivity and children's linguistic productivity and  
667 complexity. *Early Childhood Research Quarterly*, 27(3), 387-400.  
668 <https://doi.org/10.1016/j.ecresq.2012.01.001>
- 669 Poveda, D. (2003). Literature Socialization in a Kindergarten Classroom. *Journal of*  
670 *Folklore Research*, 40(3), 233-272. <http://www.jstor.org/stable/3814910>
- 671 Rodrigo, M. J., González, A., Ato, M., Rodríguez, G., Vega, M., & Muñetón, M.  
672 (2006). Co-development of Child-Mother Gestures over the Second and the  
673 Third Years. *Infant and Child Development*, 15(1), 1-17.



- Rowe, M., & Goldin-Meadow, S. (2009). Early gesture selectively predicts later language learning. *Developmental Science*, 12, 182–187.  
<https://doi.org/10.1111/j.1467-7687.2008.00764.x>
- Rowe, M.L., & Snow, C.E. (2020). Analyzing input quality along three dimensions: interactive, linguistic, and conceptual. *Journal of Child Language*, 47(1), 5-21.  
<https://doi.org/10.1017/S0305000919000655>
- Schmidt, C. (1996). Scrutinizing reference: How gesture and speech are coordinated in mother-child interaction. *Journal of Child Language*, 23(2), 279-305.  
<https://doi.org/10.1017/S0305000900008801>
- Schaffer, H. R., Hepburn, A., & Collis, G. M. (1983). Verbal and nonverbal aspects of mothers' directives. *Journal of Child Language*, 10, 337-55.
- Schwab, J.F., Rowe, M.L., Cabrera, N., & Lew-Williams, C. (2018). Fathers' repetition of words is coupled with children's vocabularies, *Journal of Experimental Child Psychology*, 166, 437-450.
- Soderstrom, M., Grauer, E., Dufault, B., & McDivitt, K. (2018). Influences of number of adults and adult: Child ratios on the quantity of adult language input across childcare settings. *First Language*, 38(6), 563–581.  
<https://doi.org/10.1177/0142723718785013>
- Soderstrom, M., & Wittebolle, K. (2013). When do caregivers talk? The influence of activity and time of day on caregiver speech and child vocalizations in two childcare environments. *PLoS ONE*, 8(11). e80646.  
<https://doi.org/10.1371/journal.pone.0080646>

696 Soler MC, Murillo E, Nieva S, Rodríguez J, Mendez-Cabezas C, Rujas I. 2022. Verbal  
697 and more: multimodality in adults' and toddlers' spontaneous repetitions.  
698 Language Learn Dev. 19(1):16–33.  
699 <https://doi.org/10.1080/15475441.2021.2008939>

700 Suanda, S.S., Smith, L.B., & Yu, C. (2016). The Multisensory Nature of Verbal  
701 Discourse in Parent–Toddler Interactions. *Developmental Neuropsychology*,  
702 41(5-8), 324-341.

703 Tal S., & Arnon, I. (2018). SES effects on the use of variation sets in child-directed  
704 speech. *Journal of Child Language*, 45(6), 1423-1438.

705 Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why Is Infant Language  
706 Learning Facilitated by Parental Responsiveness? *Current Directions in*  
707 *Psychological Science*, 23(2), 121–126.

708 Tamis-LeMonda, C., Song, L., Smith Leavell, A., Kahana-Kalman, R., & Yoshikawa,  
709 H. (2012). Ethnic differences in mother–infant language and gestural  
710 communications are associated with specific skills in infants. *Developmental*  
711 *Science*, 15(3), 384-397. <http://doi.org/10.1111/j.1467-7687.2012.01136.x>

712 Wasik, B. A., Bond, M. A., & Hindman, A. H. (2006). The effects of a language and  
713 literacy intervention on Head Start children and teachers. *Journal of Educational*  
714 *Psychology*, 98(1), 63–74. <https://doi.org/10.1037/0022-0663.98.1.63>

715 Waterfall, H.R. (2006). A little change is a good thing: feature theory, language  
716 acquisition and variation sets [Doctoral dissertation]. University of Chicago.

717 Wirén, M., Nilsson Björkenstam, K., Grigonytė, G., & Cortes, E. E. (2016).  
718 Longitudinal studies of variation sets in child-directed speech. In *The 54th*  
719 *Annual Meeting of the Association for Computational Linguistics*, Berlin,  
720 Germany, August 11, 2016 (pp. 44-52). Association for Computational  
721 Linguistics.