OPEN ACCESS Freely available online

ADVANCES IN PEDIATRIC RESEARCH

The very first words of Estonian children: A comparison of two parental report types

Astra Schults [1*], **Tiia Tulviste** [1]

1 Institute of Psychology, University of Tartu, Estonia

Abstract

Background: Parental recognition has been shown to produce more accurate information about child's vocabulary compared to remembering. The differences in reporting vocabulary are of interest. Our aims were to determine the extent of vocabulary overlap in the two report types, to identify the word report frequency and the proportions of word categories according to the report type.

Methods: The subjects were 219 children (125 boys and 94 girls) at the age of 0;8 to 1;4 (M = 10.41, SD = 1.96) who had one to three words reported in the ECDI Infant Form. In this study parents reported children's current vocabulary in two ways: Estonian version of MacArthur-Bates Communicative Development Inventory: Language and Gestures (henceforth ECDI Infant Form) and free recall.

Results: On 40% of occasions a word was reported in both ways. Three most frequent words in both report types were aitäh 'thanks', nämm-nämm 'yum-yum', and emme 'mommy'. Words belonging to the categories of sound effects and animal sounds as well as social terms had higher average frequency of reports than common nouns regardless of report type.

Conclusions: These results may reflect differences in what parents consider to be a word and raise a question of social desirability in parental reports.

Citation: Schults A, Tulviste T (2015) The very first words of Estonian children: A comparison of two parental report types. Adv Pediatr Res 2:15. doi:10.12715/apr.2015.2.15

Received: February 6, 2015; Accepted: May 12, 2015; Published: June 4, 2015

Competing interests: The authors have declared that no competing interests exist.

Sources of funding: Research for this article was supported by the Estonian Research Competency Council (grant No. SF0180025s08) and the Estonian Science Foundation (grant No. ETF9033).

* Email: astra.schults@ut.ee

Introduction

Child's very first words are eagerly looked forward to by the parents and used as developmental milestone by pediatricians. Child's first words are oftentimes a subject of discussion and it is not uncommon that parents note them down as they appear in child's productive vocabulary. Parent reports about their child's current stage of language development have been shown to be a good means to study early language development as the parents have observed their child in various situations and have extensive knowledge about their child's vocabulary [1]. This kind of knowledge might be difficult to achieve otherwise (e.g., young children may prefer not to speak to strangers in laboratory setting).

To research early vocabulary open ended questions and various checklists for parents have been used. Of those methods checklists for parents are oftentimes used in large-scale research studies as demanding fewer resources compared to other methods and providing data to address important theoretical and clinical questions at the same time [2]. Also, relying on recognition instead of remembering has been proved to produce more accurate information about current status of child's vocabulary [1]. Information

provided by parents via checklists has been shown to be valid in several languages [1, 3-8].

Some problems that might arise from using checklists for parents include a possibility of biasing parents to provide more information about nouns production as compared to the production of words from other categories [9, 10]. In many languages the largest proportion of vocabulary sized 50 words or more has been shown to be formed of common nouns (e.g., words to name animals, vehicles, toys, clothing) [1, 2, 11-16]. Another possible problem is that as the vocabulary checklists contain only a certain number of items some words produced by a particular child might not be present in the list thus limiting the variability and size of vocabulary reported. In addition, parents with low-education levels or from low-income households might not be very accurate while providing information using parental checklists as they could fail to report some of the vocabulary their child is using [17]. As the parents are asked to fill in a checklist they might conform to the perceived expectancy (an example of social desirability [19] that their child should say at least one word. Thus, the answers to the open ended questions might be more accurate [20].

Open ended questions give to the parents a possibility to report all of the vocabulary currently used by their child. An advantage of open ended questions is that these do not bias parents towards any particular word type. Thus if the noun bias is an artifact created by using checklists it should not appear in the answers parents give to the open ended question. But if the noun bias is stemming from attention or memory related processes or from the fact that children use more nouns than other types of words it should appear in the parental free report. If parents are biased to name nouns when asked to report their children's first words there should be more common nouns in free recall compared to the Estonian version of MacArthur-Bates Communicative Development Inventory: Language and Gestures (henceforth ECDI Infant Form).

Using open ended questions relies heavily on parents' memory and some of the words might not be reported even if these are produced by the child. In this study, we tried to overcome this by including only those children whose productive vocabulary is just beginning to emerge.

To our knowledge, the very first words (one to three) have not been described in large scale studies. More is known about the first 10 or 50 words. As the size of vocabulary is about 10 words the largest proportion has been shown to consist of social terms (e.g., the words that are used as a part of daily activities and names for favorite people) [11, 12, 15, 21]. In this study we separated the category of sound effects and animal sounds from social terms. The reason for keeping them separate is twofold. First, the sound effects and animal sounds are more similar to babbling (with repeating the same syllable) than to word production thus making it more difficult for the parents to recognize them as words. Second, the category includes items that could be categorized as social terms (e.g., sound effects that are a part of daily routines) as well as items that could be categorized as common nouns (e.g., animal sounds can be used to name the animals). We assume that the very first words are similar to the first 10 words in terms of categories with mainly sound effects and animal sounds, social terms and some common nouns represented.

In addition social terms and common nouns two categories can be present to a smaller extent in early vocabulary: predicates (that is verbs and adjectives) and function words (e.g., pronouns, question words, prepositions and locations, and quantifiers) [10-12].

In this study, early vocabulary will be analyzed to reveal the extent of overlap as well as possible differences in two report types. We assume that more words are reported via ECDI Infant Form compared to free recall [1]. We will describe the frequencies of occurrences of different word categories reported via two different report types. We assume that most of the reported words belong to the categories of sound effects and animal sounds and social terms [11, 12, 15, 21]. We aim to clarify if there are more common nouns reported via ECDI Infant Form (indicating that the noun bias is stemming from the parents' attention or memory related processes) or reported via free recall (indicating that the noun bias stemming from actual noun use by the children).

Methods

Participants

In this study we used the data of 219 children (125 boys and 94 girls). The age of the children varied from 0;8 to 1;4 (M = 10.41, SD = 1.96) (Table 1).

All the children whose data was selected for this study were acquiring Estonian as their first language. According to parental report all of the children were healthy (not born prematurely, without significant birth trauma, or serious medical conditions). Most of the participants (94%) were from middle or higher social economic status (henceforth SES) homes (that is mother's education level at least high school), some children (2%) were from lower SES families (mothers having not more than primary school education). For 4 % of children mother's education level was not known.

Materials

ECDI Infant Form is an adaptation of the English MacArthur-Bates Communicative Development Inventory: Language and Gestures [22]. The structure of the ECDI Infant Form is parallel to the original. One of the main parts of it is a list of 383 words, organized into semantic categories (e.g., words for food and drink, words for games and routines) [21]. The list of words can be categorized according to the word types [11-13, 15].

Social terms category was originally combined of sound effects and animal sounds, people, and games and routines in earlier research. In this study we separated the category of *sound effects and animal sounds* from *social terms* (people and games and routines).

Common nouns included animals' names, vehicles, toys, food and drink, clothing, body parts, furniture and rooms, small household items.

Other words in our study included predicates (verbs and adjectives), and function words (pronouns, question words, prepositions and locations, and quantifiers). We combined the latter two to other words category as a few of these are expected to be present in the early vocabulary.

The subject-information sheet contained questions about the child's developmental history, language(s) spoken at home, and parents' education level. On the subject information sheet parents were asked to recall the first three words their child produced.

Design

The data of children who had either one, two or three words reported in the ECDI Infant Form was selected for this study. We assumed that it would be relatively easy for the parents to freely recall the first words of their child if the child had just started to produce the very first words. The parents who completed the ECDI Infant Form filled in the subject-information sheet at the same time. Thus parents reported the words produced by their child via two report types - recognition and free recall.

Table 1. Age and gender distribution in the sample

	Children's age									
	0;8	0;9	0;10	0;11	1;0	1;1	1;2	1;3	1;4	Row total
Boys	22	24	21	17	14	12	11	2	2	125
Girls	16	25	19	16	10	4	2	0	2	94
All groups	38	49	40	33	24	16	13	2	4	219

Procedure

Participants were recruited by the researchers, pediatricians, play-group teachers, internet forums, and by the parents who already had filled in the inventory by themselves. Parents, who agreed to participate, were handed out or sent copies of the ECDI Infant Form and a subject-information sheet. The parents were asked to complete the ECDI Infant Form by checking the words that their child produced. The subject-information sheet was asked to be filled in at the same time as the parents completed the ECDI Infant Form.

The work has been approved by the ethical committee of the University of Tartu (Nr. 170/T-12, 28.04.2008).

Results

Word frequency in two types of parental reports

For 77 % of the children in our sample the parents had written at least one word as an example of the first words produced by their child. For the remaining 23% of the children no information about the first words was given or parents had written letter combinations that could not be identified as a word (8 times). All together parents reported a word as produced by their child for 561 times. This includes all the times a word was checked in the ECDI Infant Form or was freely recalled or was reported in both ways. The list composed of all the words reported by parents included 64 different items.

The parents had reported one and the same word in ECDI Infant Form as well as in free recall for 222 (40%) times. The list composed of the words that parents had reported in both ways included 22 different words. Five of these words were reported more than ten times including *aitäh* 'thanks' (71 times), *nämm-nämm* 'yum-yum' (47 times), and *emme* 'mommy' (44 times), *aidaa* 'bye-bye' (13 times), *anna* 'give to me' (11 times) (Table 2).

The parents had checked a word on ECDI Infant Form but not reported the same word in free recall for 189 (34%) times. The list of these words included 50 different items. Six of these words had been checked at least ten times. These words were *nämm-nämm* 'yum-yum' (39 times), *aitäh* 'thanks' (30 times), *põrr-põrr* 'vroom' (19 times), *ai-ai* 'ouch' (13 times), *aidaa* 'bye-bye' (12 times), and *emme* 'mommy' (10 times) (Table 2).

The parents had reported a word via free recall that they had not checked in ECDI Infant Form for 147 (26%) times. There were 30 different words that the parents had only reported in free recall. These included eight different words that were not present in the ECDI Infant Form (12 times of report). Three of these words were reported in free recall more than ten times, namely *emme* 'mommy' (58 times), *nämm-nämm* 'yum-yum' (23 times), and *aitäh* 'thanks' (12 times) (Table 2).

Differences in word frequencies according to report types

First, correlation analyses were conducted with seven of the most frequently mentioned words (Table 2) to determine if the same words get reported via ECDI Infant Form, free recall and both ways (Fig. 1). None of the three correlations yielded significance indicating that different ways of reporting result in different word frequencies.

Second, analyses were conducted with seven most frequently reported words to see if there are differences between report types. Chi-square analyses showed that there are significant differences between word reporting frequencies according to report types (Table 3). *Põrr-põrr* 'vroom' ($(O-E)^2 / E = 324.0$) and *ai-ai* 'ouch' ($(O-E)^2 / E = 144.0$) were reported more frequently in ECDI Infant Form than in free recall ($\chi^2 = 555.9$, $p = 0.0000$). *Emme* 'mommy' ($(O-E)^2 / E = 26.3$) and *aitäh* 'thanks' ($(O-E)^2 / E = 23.7$) were more frequently reported in both ways compared to the sole reports in the ECDI Infant Form ($\chi^2 = 66.5$, $p = 0.0000$). *Aitäh* 'thanks' ($(O-E)^2 / E = 49.0$) and *nämm-nämm* 'yum-yum' ($(O-E)^2 / E = 12.3$) were reported more frequently in both ways compared to the sole reports via free recall ($\chi^2 = 77.1$, $p = 0.0000$).

Table 2. Comparisons of numbers of produced words in different types of parental reports

	Number of times the word was			ECDI Infant Form *vs* free recall		ECDI Infant Form *vs* both ways		Free recall *vs* both ways	
	Checked in ECDI Infant Form	Reported in free recall	Reported in both ways	O - E	$(O-E)^2/E$	O - E	$(O-E)^2/E$	O - E	$(O-E)^2/E$
nämm-nämm 'yum-yum'	39	23	47	16	11.1	-8	1.4	-24	12.3
aitäh 'thanks'	30	12	71	18	27.0	-41	23.7	-59	49.0
põrr-põrr 'vroom'	19	1	0	18	324.0	19	0.0	1	0.0
ai-ai 'ouch'	13	1	5	12	144.0	8	12.8	-4	3.2
aidaa 'bye-bye'	12	5	13	7	9.8	-1	0.1	-8	4.9
emme 'mommy'	10	58	44	-48	39.7	-34	26.3	14	4.5
anna 'give to me'	6	5	11	1	0.2	-5	2.3	-6	3.3
Sum	129	105	191	24	555.9	-62	66.5	-86	77.1

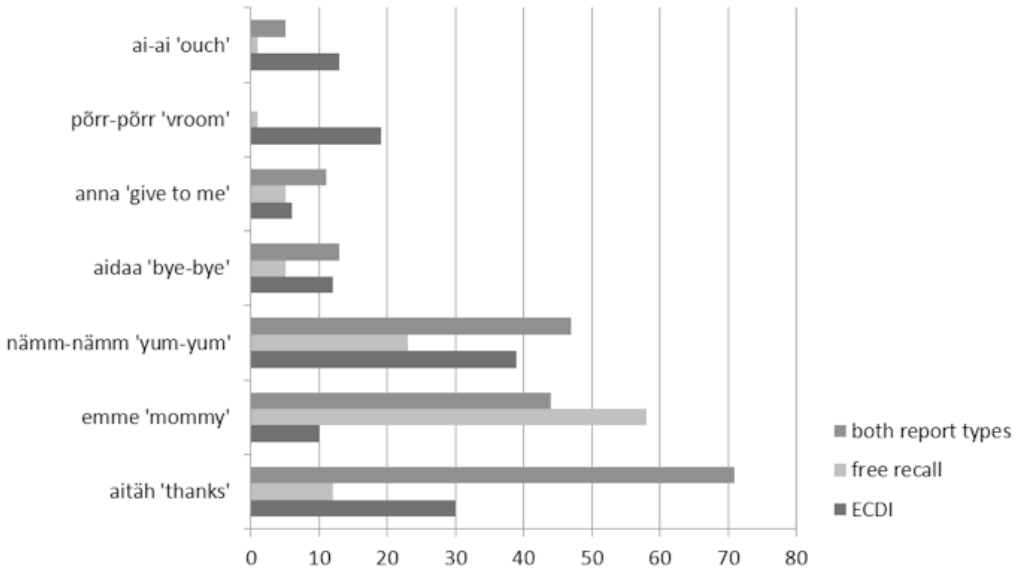

Figure 1. Comparison of word frequencies reported in different ways

Table 3. Mean frequencies of reporting words from different categories in different ways

	Report types combined			Reported in ECDI Infant Form			Reported in free recall			Reported in both ways		
	Words	Reports	M	Words	Reports	M	Words	Reports	M	Words	Reports	M
Sound effects and animal sounds	12	177	14.8	12	90	7.5	4	26	6.5	5	58	11.6
Social terms	15	323	21.5	11	68	6.2	12	100	8.3	11	155	14.1
Common nouns	24	45	1.9	19	23	1.2	8	14	1.8	5	8	1.6
Predicates and other words	13	16	1.2	8	8	1.0	6	7	1.2	1	1	1.0
All words	64	561	8.8	50	189	3.8	30	147	4.9	22	222	10.1

Note: M - Mean probability of being reported per word

Proportions of different word categories

We calculated the proportion of each word category in the whole data as well as in each of the report types (Table 3). There were all together 64 different words that parents had reported in one way or the other or in both ways. Twelve (19%) of the words belonged to sound effects and animal sounds, 15 (23%) to social terms, 24 (38%) to common nouns, and 13 (20%) to the category of other words.

There were 50 different words that were reported in ECDI Infant Form. Twelve (24%) of the words belonged to sound effects and animal sounds, 11 (22%) to social terms, 19 (38%) to common nouns, and 8 (16%) to the category of other words.

There were 30 different words that were reported in free recall. Four (13%) of the words belonged to sound effects and animal sounds, 12 (40%) to social terms, 8 (27%) to common nouns, and 6 (20%) to the category of other words.

There were 22 different words that were reported in both ways. Five (23%) of the words belonged to sound effects and animal sounds, 11 (50%) to social terms, 5 (23%) to common nouns, and 1 (5%) to the category of other words.

As the report frequencies of words varied according to the word as well as to the report type we performed the following analyses to see the proportional differences in between the word categories.

Mean report frequency of words from different categories

We calculated mean report frequencies per word for each word category and each report type (Table 3). Of the words that had been checked in ECDI Infant Form a sound effect or animal sound was reported 7.5 times average. A social term had 6.2 and a common noun had 1.2 reports on average. A word from other categories was reported just once on average.

Of those words that had been recalled free a sound effect or animal sound was reported on 6.5 occasions and a social term 8.3 occasions on average. A common noun had 1.8 reports on average. A word belonging to other categories was reported 1.2 times on average.

Of the words that had been reported in both ways a sound effect or animal sound was reported 11.6 times and a social term 14.1 times on average. A common noun had 1.6 reports and a word from other categories 1.0 report on average.

Correlation analyses showed that mean word report frequencies in different categories are similar even if these are reported in different ways. Correlation between word report frequencies in different categories obtained by both types of report and by ECDI Infant Form was significant ($r = 0.95$, $p < 0.05$) as well as correlation between word report frequencies in different categories obtained by both types of report and by free recall ($r = 1.00$, $p < 0.05$). The correlation between word report frequency in

different categories reported in ECDI Infant Form and reported via free recall did not reach significance ($r = 0.93$, $p < 0.08$). Chi-square did not reach significance in any of the three comparisons.

Discussion

Our first aim was to determine the extent of the overlap of words reported in different ways as well as to see what words get reported more often, and if the frequency depends on report type. 39.6% of reports had the same word reported in ECDI Infant Form as well as in free recall. This finding indicates that the overlap of the words reported in two different ways at the same time is far from perfect even as the vocabulary to be reported consists of one to three words. This result is in part due to the fact that some of the words reported via free recall were not present in ECDI Infant Form. At the same time, there were occasions when parents had checked a word in ECDI Infant Form but had not put it down in free recall. Thus we were interested to see what the words that contributed the most to the difference were. As we compared the frequency of word reports the words *emme* 'mommy' and *aitäh* 'thanks' the parents reported those words more often via both report types than just via recognition. Thus as a child has started to say either of those words these catch parent attention and are memorable so that it is easy for the parents to report them via free recall as well as via vocabulary checklist. The word *nämm-nämm* 'yum-yum' was reported more often via both ways than just via free recall. A possible explanation for this is that as a parent recognizes this sound effect as a word they are able to report it in both ways easy enough but as some of the parents do not consider it to be a real word there are only a few cases when it is reported via free recall and not checked in the ECDI.

In our study common nouns formed more than a third of the list of words reported by the parents. Social terms accounted for about a quarter of the word list. If we combine the last with the sound effects and animal sounds (that have been included in social terms category in previous research) that accounted for almost 20% of the different words reported by the parents we can conclude, that this study has confirmed the finding from previous studies that the very first words tend to be either social terms or common nouns [1, 2, 11-16]. In previous research predicates and other words have been shown to be present very rarely within very first words. But if we look at the predicates and other words in the list of different words reported by the parents in this study we can see that these account for about 20%. We were interested to determine if there were differences in the proportions of word categories reported in different ways. We found that all the report methods elicited more sound effects and animal sounds as well as more social terms than common nouns or predicates and other words if the frequency of word reports was taken into account. Thus we can conclude that even as some of the children have common nouns or predicates and other words represented within their very first words most of the children's very first words belong to either sound effects and animal sounds category or to social terms category.

It has been suggested by the previous research that vocabulary checklists might bias parents to provide more information about nouns production than about production of words belonging to other categories [9, 10]. We found that compared to the free recall parents had checked somewhat more common nouns in ECDI Infant Form. At the same time, the list of words combined of the parental reports via ECDI Infant Form was longer than the list of words combined of the ones reported via free recall. This result suggests that it might be easier for the parents to report the words their child has produced in general and common nouns in particular if they are provided with a vocabulary checklist. Vocabulary checklist provides a memory aid as it lists a variety of words and the parents need just to recognize the words they have heard in their child's vocabulary instead of remembering them. Thus as expected, relying on recognition instead of remembering produced more information about child's vocabulary [1, 5, 17]. As we look at the words that have been reported more frequently via ECDI Infant Form compared to the words reported via free recall it is noteworthy that two of the words contributing the most to the difference are from sound effects and animal sounds category (*põrr-põrr* 'vroom' and *ai-ai* 'ouch'). Thus, we can interpret this finding that for some parents it is difficult to take verbalizations belonging to the category of sound effects and animal sounds into account as real words and so they do not report them

via free recall as they have been asked to provide the words used by their child.

For three quarters of the children in our sample parents had reported at least one word in free recall. Thus, there were several cases when parents had not reported any word being used by their child via free recall even as they had checked words in the ECDI Infant Form at the same time. This finding can be explained with the tendency of MacArthur-Bates CDI to over represent vocabulary in the early age [18]. This tendency may be due to either social desirability [19] with parents checking the words in the list to conform with the perceived expectancy that their child should say at least one word. Thus, the answers to the open ended questions might be more accurate [20]. Still, the question of social desirability might arise for the free recall as well. *Emme* 'mommy' was the word that had the most occasions of free recall without it being checked in the vocabulary checklist at the same time. This might be due to parents' expecting it to be one of the first words their child should say.

A limitation of this study comes from the fact that in addition to the two parent report instruments no direct measure of the child's first words was used. As we used a part of a data from a large scale study the direct measurements were not included in the design. It would be important in future studies to include some direct measure of child's vocabulary.

Conclusions

Very first words tend to be either from the categories of sound effects and animal sounds or social terms. Using parental checklists provides more information about child's current vocabulary compared to free recall as the children are very young and their vocabulary small indeed.

Acknowledgments

The authors would like to thank the parents who participated in the study.

References

1. Fenson L, Dale P, Reznick JS, Bates E, Thal DJ, Pethick S. Variability in early communicative development. Monogr Soc Res Child. 1994;59:1-185.
2. Bornstein MH, Cote LR, Maital S, Painter K, Park SY, Pascual L, et al. Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. Child Dev. 2004;75:1115-39.
3. Berglund E, Eriksson M. Reliability and content validity of a new instrument for assessment of communicative skills and language abilities in young Swedish children. Logop Phonatr Vocol. 2000;25:176-85.
4. Bleses D, Vach W, Slott M, Wehberg S, Thomsen P, Madsen TO, et al. The Danish communicative developmental inventories: validity and main developmental trends. J Child Lang. 2008;35:651-69.
5. Dale PS, Bates E, Reznick JS, Morisset C. The validity of a parent report instrument of child language at twenty months. J Child Lang. 1989;16:239-49.
6. Eriksson M, Marschik PB, Tulviste T, Almgren M, Pérez Pereira M, Wehberg S, et al. Differences between girls and boys in emerging language skills: Evidence from 10 language communities. Brit J Dev Psychol. 2012;30:326-43.
7. Marchman VA, Martínez-Sussmann C. Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. J Speech Lang Hear R. 2002;45:983-97.
8. Rescorla L, Alley A. Validation of the language development survey (LDS): a parent report tool for identifying language delay in toddlers. J Speech Lang Hear R. 2001;40:556-66.
9. Pine JM. How referential are "referential" children? Relationships between maternal-report and observational measures of vocabulary composition and usage. J Child Lang. 1992;19:75-86.
10. Tardif T. Nouns are not always learned before verbs: evidence from Mandarine speakers' early vocabulary. Dev Psychol. 1996;32:492-504.
11. Bates E, Marchman V, Thal D, Fenson L, Dale P, Reznick JS, et al. Developmental and stylistic variation in the composition of early vocabulary. J Child Lang. 1994;21:85-123.
12. Caselli C, Cassadio P, Bates E. A comparison of the transition from first words to grammar in English and Italian. J Child Lang. 1999;26:69-111.
13. Caselli MC, Bates E, Casadio P, Fenson J, Fenson L, Sanderl L, et al. A cross-linguistic study of early lexical development. Cognitive Dev. 1995;10:159-99.
14. Nelson K. Structure and strategy in learning to talk. Monographs of the Society for Research in Child Dev. 1973;38:1-135.

15. Stolt S, Haataja L, Lapinleimu H, Lehtonen L. Early lexical development of Finnish children: a longitudinal study. First Language. 2008;28:259-79.

16. Wehberg S, Vach W, Bleses D, Thomsen P, Madsen TO, Basbøll H. Danish children's first words: analyzing longitudinal data based on monthly CDI parental reports. First Language. 2007;27:361-83.

17. Roberts JE, Burchinal M, Durham M. Parents' report of vocabulary and grammatical development of African American preschoolers: child and environmental associations. Child Dev. 1999;70:92-106.

18. Fenson L, Bates E, Dale P, Goodman J, Reznick JS, Thal D. Measuring variability in early child language: don't shoot the messenger: [reply]. Child Dev. 2000;71:323-8.

19. Lehman DR, Krosnick JA, West RL, Li F. The focus on judgment effect: a question wording effect due to hypothesis confirmation bias. Pers Soc Psychol B. 1992;18:690-9.

20. Oller DK, Eilers RE, Neal AR, Schwartz HK. Precursors to speech in infancy: the prediction of speech and language disorders. J Commun Disord. 1999;32:223-45.

21. Schults A, Tulviste T, Konstabel K. Early vocabulary and gestures in Estonian children. J Child Lang. 2012;39:664-86.

22. Fenson L, Dale PS, Reznick JS, Thal D, Bates E, Hartung JP, et al. MacArthur communicative development inventories: user's guide and technical manual. San Diego, London: Singular Publishing Group; 1993.