A preliminary study on the composition of the early expressive lexicon in Egyptian infants and toddlers

Aisha F. Abdel Hady, Heba M. Farag, Ayatallah R. Sheikhany

Phoniatric Unit, ENT Department, Faculty of Medicine, Cairo University, Cairo, Egypt

Correspondence to Aisha F.A. Hady, MD, Phoniatric Unit, ENT Department, Faculty of Medicine, Cairo University, Cairo 11562, Egypt. Tel: +20 112 493 2611; fax: +20 26 28 884; e-mail: aishafawzy1@yahoo.com

Objective
The aim of this study was to gain insight into early vocabulary size in Egyptian children aged between 12 and 30 months and to study the children's lexicon composition at that young age while studying some demographic factors that might affect early vocabulary development.

Participants and methods
Parents of 150 children joining day care nurseries in Cairo were asked to fill in the Arabic vocabulary checklist designed in this study. The children were divided according to their age into three groups and then further subdivided according to their vocabulary size into low-vocabulary and high-vocabulary groups.

Results
Range and median and vocabulary percentages were estimated. Vocabulary size in older children is larger than younger children. More word types developed in the high-vocabulary group in younger age than in low-vocabulary group. All types of vocabulary continued to grow in older children in both low-vocabulary and high-vocabulary groups. There are differences between high-vocabulary and low-vocabulary groups in some demographic factors.

Conclusion
In this study, the vocabulary size for Arabic infants and toddlers is less than some languages. There are delays and differences between low-vocabulary and high-vocabulary groups regarding the linguistic composition and some demographic factors.

Keywords:
Arabic vocabulary checklist, composition of children's, lexicon, demographic factors, low-vocabulary and high-vocabulary groups, vocabulary production

Introduction

The style and rate at which children acquire a language have universal characteristics although there are individual variations in the process of language development. These differences in language development can often be seen between 1- and 3-year-old children, concerning, for instance, when the first words are uttered, and what kind of words the children are learning.

The study of early language acquisition was carried out in different methods. The difference might be owing to various data collection methods: either observational or controlled experiment, length of study: either cross-sectional or longitudinal study, collected data: either through parental-diary or transcription of language sample by professionals, sample size (number and distribution of sampling sessions), and degree of standardization: either informal diary data or checklist measures.

Various studies have testified the effectiveness of parent reports in measuring lexical development in the early years [1]. Although parents' notes can be helpful to record utterances over a longer period of time than the home visits, the observations of a parent would not be sufficient to draw reliable conclusions. Using questionnaires and checklists can be useful as the inventories on which they are based include the most common words produced by children of the parent's age and they will therefore remind parents of words their child produces. Checklists do not need high educational level of the parents. In contrast, language sampling through structured testing have limitations when carried out with children under the age of 2.6 years. Situational and temperamental factors ranging from disinterest in the tasks to anxiety can introduce unwanted variation in performance or preclude testing altogether. Moreover, it requires trained examiners.

Arabic is a semitic language. The morphology of Arabic is rich. Suffixes are used to express grammatical relations. Inflections of singular and plural words differ, as well.

Verbs are inflected for subject–verb agreement, tense, etc. Rich morphology also means that fewer words are needed to express the same meaning than, for example, in English. Prior studies have shown that the amount of child-directed speech affects children's early lexicon growth [2,3]. Egyptian Arabic displays both the VSO (verb–subject–object) order and the SVO (subject–verb–object) order. However, the SVO order seems to be increasingly dominant in Egyptian Arabic [4]. Deviations from the SVO order are used in negative, interrogative, imperative, and exclamatory sentences. Thus, there is a variety in the structure of the sentences Arabic children are exposed to. In addition to the Egyptian parents' style in the speech directed to children, they use imperative language to direct the children's behavior that does not call for verbalization in an optimum way. Consistent with this, one might assume that children acquiring a language with rich morphology might have slower vocabulary development rates and smaller expressive lexicon compared with the children acquiring a language with simple morphology.

The early language development is a dynamic process that is affected by a complex array of transacting factors from multiple levels of influence, such as genetics, sex temperament, the child's own skills in other developmental domains (cognitive, motor, and socioemotional), and a host of biological and social risk and resilience factors, for example, premature birth, prenatal exposure to substances, parental education, caregiver interactive style, parent–child mutual regulation, bilingualism, and other cultural influences [5,6].

Up to our knowledge, there is little information about the development of Egyptian children's vocabulary in their early years because of deficiency of standardized assessment tools, and believing that cultural and linguistic differences reflect on talk addressed to children affecting their early language acquisition, we think it is valuable to study if there is any language-specific difference in early vocabulary repertoire of Egyptian children.

Thorough study of early lexical development in Egyptian infants and toddlers and its pattern is of great benefit in the clinical practice, as it helps picking up cases at risk of delayed language acquisition and helps setting intervention programs simulating the pattern of normal language acquisition for those at risk or those who have delayed language development. It also provides a useful evaluation of the progress of vocabulary development of children after therapy.

We also find that it is important to analyze the repertoire of vocabulary acquisition in infants and toddlers who have high-vocabulary and low-vocabulary levels so as to know the nature of the difference: is it related only to vocabulary size or to types of words in their expressive lexicon? Knowing the underlying factors that contribute to vocabulary acquisition in this early period of life is crucial to provide proper counseling for low-vocabulary group to help them catch up with their peers.

Objective
The objective of this study was to gain insight into the approximate early vocabulary size in Egyptian children aged between 12 and 30 months and to study the composition and pattern of children's lexicon at that young age while studying some demographic factors that might affect early vocabulary development.

Participants and methods
Participants
Parents of 150 children joining day care nurseries in Greater Cairo area, aged between 12 and 30 months, gave a consent, after explaining to them the purpose of the study, and they were asked to fill in the Arabic vocabulary checklist designed in this study. The children were of the same middle socioeconomic standard. The parents had to fill in a brief history about them and their children regarding the educational level of the parents, the prenatal, natal, developmental history, and childhood illness in addition to some information about the order of the child in the family, the number of hours spent by the child with his mother and father as well as the number of hours the child is exposed to television.

Participants who met the following criteria were included in this study: (i) Arabic was the only language spoken at home; (ii) subjective normal hearing of the children; (iii) no reported history of prenatal, natal, developmental, or postnatal illness; and (iv) their parents held diploma or university degree. Participants who were exposed to language(s) other than Arabic at home, whose hearing was suspected to be impaired, or who had a diagnosis of developmental delay or known disability (e.g. Down's syndrome, autism, cerebral palsy, and seizure disorder) were not included in this study. There was no serious medical illness of children under study.

The children under study were divided according to their age into three groups: group A included 30 children aged 12–18 months, group B included 42 children aged more than 18–24 months and group

C included 78 children aged more than 24–30 months. The study protocol has been reviewed and approved by the research committee of the department.

Each of the three age groups was divided statistically into low-vocabulary and high-vocabulary groups by the median variance after estimating the range of the vocabulary in each age group.

Low-vocabulary and high-vocabulary groups in group A included 15 children each, group B included 21 children each, and group C included 39 children each.

Description of the instrument
The checklist was designed in this study and applied on 150 children joining day care nurseries in Greater Cairo area. The children under study were selected between the ages of 12 and 30 months to assess early vocabulary acquisition at this young age.

Initially, the checklist contained 662 Arabic words selected from the most frequent words expected to be present in the early child vocabulary lexicon. The words were then grouped into six large categories. The checklist is for vocabulary, so proper names as well as gestures were excluded. The Egyptian cultural effect was put into consideration while designing the checklist and selecting words. The checklist has a short introduction explaining the purpose of the checklist and the instructions that the parents participating in this study should follow while filling in the checklist. As the checklist is for assessing the expressive vocabulary of Egyptian infants and toddlers, the instructions clarify that the parents should mark only what the child spontaneously says regardless of the correct pronunciation of the produced words. Words that the child understands but does not say should not be included. The same is for the words that the child can imitate but does not use spontaneously, which should not be included in the study.

The checklist also contains a separate column in front of the vocabulary list to be checked by parents if each word is produced by the child. In addition to that, it includes a section filled in by parents of words that the child produces and not mentioned in the checklist. The parents added a total number of 36 words into that section and the number of words in the checklist turned into 698. All the parents were contacted again through the phone to check if their children produce any of the added 36 words.

The grouping of vocabulary used into categories served two purposes: representing various communicative intensions in addition to reducing the size of collected data. The six large categories included the following: onomatopoeic items (sounds of animals and transportation), social interactional words, nominal words, action words, descriptive words, and grammatically related words (pronouns and question words). This categorization of the vocabulary was guided by a study done by Nelson [7].

Nominals are further classified into the following: general nominal, specific nominal, and nominal related to personal interest. General nominal is representing nouns labeling people, living things, and objects. Specific nominal is representing nouns labeling time, places, and occupations, whereas nominal related to personal interest is representing nouns labeling games and toys and cartoon characters.

The data were collected in the period from June 2016 to March 2017. Afterward, the data were statistically analyzed and tabulated.

Statistical methods
Data management and statistical analysis were performed using the statistical package for social sciences, version 21 SPSS 21 (Armonk, NY, USA: IBM Corp.).

Numerical data were summarized using means and standard deviations or medians and ranges. Categorical data were summarized as percentages. Comparisons between the three age groups with respect to normally distributed numeric variables were done using the univariate analysis of variance and Bonferroni's post-hoc test. Non-normally distributed numeric variables were compared by Kruskal–Wallis test (nonparametric analogue for analysis of variance) followed by Mann–Whitney test for pairwise comparisons. For categorical variables, differences were analyzed with χ^2-test and Fisher's exact test when appropriate. Adjustment of P value was done by Bonferroni methods to correct for multiple comparisons.
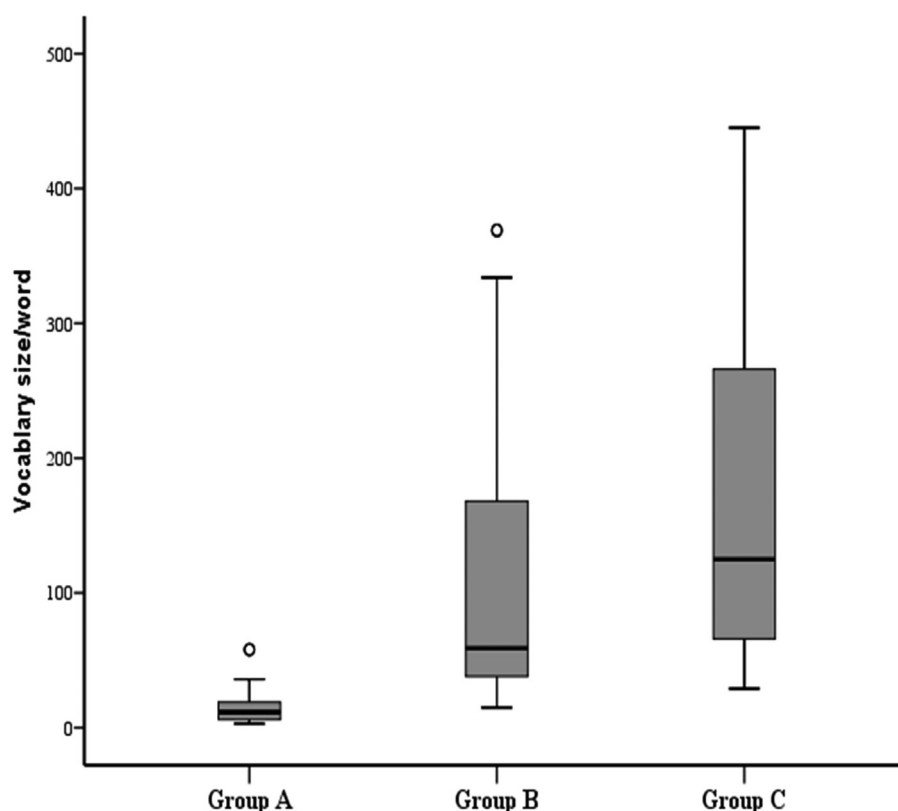
All P values are two sided. P values less than 0.05 were considered significant.

Results
Figure 1 shows the vocabulary size/word in the three age groups.

Comparison between median and range (minimum and maximum) of vocabulary and their percentages in the three age groups showed that there was a significant difference among the age groups

Figure 1

Vocabulary size/word in the three age groups.

Table 1 Comparison among median and range (minimum and maximum) of vocabulary and their percentages in the three age groups

	Group A			Group B			Group C			P value
	Median	Minimum	Maximum	Median	Minimum	Maximum	Median	Minimum	Maximum	
Total	11.5	3	58	59	15	369	125	29	445	<0.001
Total percentage	1.7	0.4	8.3	8.5	2.1	52.9	17.9	4.2	63.8	<0.001

regarding the total vocabulary size and their percentages. Pair-wise comparison shows the following P values: between groups A and B=0.001, groups A and C=0, and groups B and C=0.051 (Table 1).

Comparison regarding the percentage of vocabulary categories in all the three age groups showed that there was a significant difference among the three age groups regarding all the vocabulary categories (Table 2).

Pair-wise comparison regarding the percentage of vocabulary categories in all the three age groups showed that the order of the percentages of acquisition of vocabulary categories in group A was as follows: onomatopoeic words, social interactional words, and then nominals (nominals related to personal interest and general nominals). In group B, vocabulary acquisition was as follows: nominals showed the highest percentage, and in further analysis of the distribution of different nominal groups, the distribution was as follows: onomatopoeic words, grammar-related words, social interaction words, and general nominals as well as nominals related to personal interest, descriptive, action words, and then specific nominal, whereas in group C, vocabulary acquisition was as follows: the whole nominal group showed the highest percentage. In further analysis, the distribution was as follows: onomatopoeic words, social interactional words, grammar related, general nominals, action words, nominals related to personal interest, descriptive, and then specific nominals (Table 3).

Vocabulary composition in low-vocabulary and high-vocabulary children of group A showed that the same pattern was acquired in both low-vocabulary and high-vocabulary acquisition groups in the following order: general nominals, social interactional, and onomatopoeic words. In addition, in high-vocabulary group, other types emerged in the following order:

Table 2 Comparison regarding the percentage of vocabulary categories in all the three age groups

	Group A			Group B			Group C			P value
	Median	Minimum	Maximum	Median	Minimum	Maximum	Median	Minimum	Maximum	
Onomatopoeic items percentage	18.75	0	87.5	37.5	25	87.5	50	0	87.5	<0.001
Social interactional percentage	6.6	0	23.7	14.5	2.6	57.9	31.6	2.6	60.5	<0.001
Nominal all percentage	4	3	24	44.5	2	232	85.5	13	301	<0.001
General nominal percentage	1.06	0.71	7.42	11.84	0.71	72.08	25.62	4.59	89.05	<0.001
Specific nominal percentage	0	0	2.13	4.79	0	25.53	7.98	0	39.36	<0.001
Personal interest percentage	1.52	0	6.06	7.58	0	22.73	11.36	0	43.94	<0.001
Descriptive percentage	0	0	3	6	0	39.4	9.1	0	59.1	<0.001
Action words percentage	0	0	18.9	4.95	0	81.1	12.15	0	78.4	<0.001
Grammatical-related percentage	0	0	22.73	15.91	0	100	29.55	4.55	59.09	<0.001

Table 3 Pair-wise comparison regarding the percentage of vocabulary categories in all the three age groups

	Pair-wise comparison		P value
Onomatopoeic items percentage	Group A	Group B	**0.009**
		Group C	**<0.001**
	Group B	Group C	0.546
Social interactional percentage	Group A	Group B	**0.011**
		Group C	**<0.001**
	Group B	Group C	0.067
Nominal all percentage	Group A	Group B	**0.001**
		Group C	**<0.001**
	Group B	Group C	**0.021**
General nominal percentage	Group A	Group B	**0.001**
		Group C	**<0.001**
	Group B	Group C	**0.019**
Specific nominal percentage	Group A	Group B	**0.002**
		Group C	**<0.001**
	Group B	Group C	0.262
Personal interest percentage	Group A	Group B	**<0.001**
		Group C	**<0.001**
	Group B	Group C	**0.02**
Descriptive percentage	Group A	Group B	**0.004**
		Group C	**<0.001**
	Group B	Group C	0.091
Action words percentage	Group A	Group B	**0.017**
		Group C	**<0.001**
	Group B	Group C	0.504
Grammatical-related percentage	Group A	Group B	**0.001**
		Group C	**<0.001**
	Group B	Group C	0.073

nominals related to personal interest, grammar-related words, and action words), with no emergence of specific nominals or descriptive words (Table 4).

Vocabulary composition in low-vocabulary and high-vocabulary children of group B showed that all types of vocabulary emerged and continued to grow in group B in both low-vocabulary and high-vocabulary groups. More content words (action and descriptive words) proceeded with relative decline of social interactional and onomatopoeic words especially in the higher vocabulary acquisition group. The following are the patterns earned: in group B low vocabulary: general nominals, followed by social interactional, then action words, onomatopoeic words, nominals related to personal interest, specific nominals, grammar related, and descriptive words, whereas in group B high vocabulary, general nominals, followed by action words, then descriptive words, nominals related to personal interest, specific nominals, social interactional, grammar-related words, and onomatopoeic words (Table 5).

Vocabulary composition in low-vocabulary and high-vocabulary children of group C showed that nearly in group C, in both low-vocabulary and high-vocabulary groups, the pattern continued like that in the preceding group B, but with increasing on the level of the vocabulary size as in the following: in group C low vocabulary: general nominals, followed by action words, then social interactional words, nominals related to personal interest, specific nominals, grammar related, onomatopoeic words, and then descriptive words, whereas in group C high vocabulary: general nominals followed by action words, social interactional, descriptive words, nominals related to personal interest, specific nominals, grammar related, and onomatopoeic words (Table 6).

Factors affecting vocabulary acquisition in low-vocabulary and high-vocabulary children of the three age groups showed that regarding maternal age, there was a significant difference between low-vocabulary and high-vocabulary groups in only the older age group C regarding the maternal age. Regarding maternal and paternal contribution and time spent with children, the results showed there was a significant difference between low-vocabulary and high-vocabulary groups

Table 4 Vocabulary composition in low-vocabulary and high-vocabulary children of group A

	Group A						P value
	Low-vocabulary			High-vocabulary			
	Median	Minimum	Maximum	Median	Minimum	Maximum	
Onomatopoeic items	1	0	3	2	1	7	0.083
Onomatopoeic items percentage	12.5	0	37.5	25	12.5	87.5	0.083
Social interactional	2	0	4	3	1	9	0.058
Social interactional percentage	5.3	0	10.5	7.9	2.6	23.7	0.058
General nominal	3	2	3	10	3	21	<**0.001**
General nominal percentage	1.06	0.71	1.06	3.53	1.06	7.42	<**0.001**
Descriptive	0	0	0	0	0	2	0.073
Descriptive percentage	0	0	0	0	0	3	0.073
Specific nominal	0	0	0	0	0	2	**0.007**
Specific nominal percentage	0	0	0	0	0	2.13	**0.007**
Action words	0	0	0	1	0	21	<**0.001**
Action words percentage	0	0	0	0.9	0	18.9	<**0.001**
Grammatical related	0	0	0	2	0	5	<**0.001**
Grammatical-related percentage	0	0	0	9.09	0	22.73	<**0.001**
Personal interest	0	0	1	2	1	4	<**0.001**
Personal interest percentage	0	0	1.52	3.03	1.52	6.06	<**0.001**
Nominal all	3	3	3	14	5	24	<**0.001**
Nominal all percentage	0.68	0.68	0.68	3.16	1.13	5.42	<**0.001**

Table 5 Vocabulary composition in low-vocabulary and high-vocabulary children of group B

	Group B						P value
	Low-vocabulary			High-vocabulary			
	Median	Minimum	Maximum	Median	Minimum	Maximum	
Onomatopoeic items	3	2	5	3	2	7	0.281
Onomatopoeic items percentage	37.5	25	62.5	37.5	25	87.5	0.281
Social interactional	6	1	12	5	2	22	0.202
Social interactional percentage	15.8	2.6	31.6	13.2	5.3	57.9	0.202
General nominal	16	2	35	91	32	204	<**0.001**
General nominal percentage	5.65	0.71	12.37	32.16	11.31	72.08	<**0.001**
Descriptive	1	0	4	12	4	32	<**0.001**
Descriptive percentage	1.5	0	6.1	18.1	6	39.4	<**0.001**
Specific nominal	2	0	6	7	4	24	<**0.001**
Specific nominal percentage	2.13	0	6.38	7.45	4.26	25.53	<**0.001**
Action words	4	0	6	21	5	90	<**0.001**
Action words percentage	3.6	0	5.4	18.9	4.5	81.1	<**0.001**
Grammatical related	2	0	3	5	4	22	<**0.001**
Grammatical-related percentage	9.09	0	13.64	22.73	18.18	100	<**0.001**
Personal interest	3	0	7	10	4	15	<**0.001**
Personal interest percentage	4.55	0	10.61	15.15	6.06	22.73	<**0.001**
Nominal all	21	2	46	116	43	232	<**0.001**
Nominal all percentage	4.74	0.45	10.38	26.19	9.71	52.37	<**0.001**

regarding hours spent by father with children in both age groups A and B, whereas there was a significant difference between low-vocabulary and high-vocabulary groups in age group C only regarding hours spent by mother with children. In other words, the number of hours spent by father with his children was more in high vocabulary children in both age groups A and B than hours spent by father with children who have low vocabulary, whereas the number of hours spent by the mother with children who have high vocabulary was more than number of hours spent by the mother with children who have low vocabulary in age group C. Regarding exposure to television, the results showed significant differences between the low-vocabulary and high-vocabulary groups in all the three age groups. In other words, higher vocabulary children in all age groups were exposed less to television (Table 7).

Birth order as a factor affecting vocabulary acquisition in low-vocabulary and high-vocabulary children of all

Table 6 Vocabulary composition in low-vocabulary and high-vocabulary children of group C

	Group C						
	Low-vocabulary			High-vocabulary			
	Median	Minimum	Maximum	Median	Minimum	Maximum	*P* value
Onomatopoeic items	3	1	7	4	0	7	0.016
Onomatopoeic items percentage	37.5	12.5	87.5	50	0	87.5	0.016
Social interactional	5	1	16	17	3	23	<0.001
Social interactional percentage	13.2	2.6	42.1	44.7	7.9	60.5	<0.001
General nominal	35	13	72	150	73	252	<0.001
General nominal percentage	12.37	4.59	25.44	53	25.8	89.05	<0.001
Descriptive	2	0	14	15	6	39	<0.001
Descriptive percentage	3	0	21.2	22.7	9.1	59.1	<0.001
Specific nominal	4	0	12	10	2	37	<0.001
Specific nominal percentage	4.26	0	12.77	10.64	2.13	39.36	<0.001
Action words	6	0	16	29	5	87	<0.001
Action words percentage	5.4	0	14.4	26.1	4.5	78.4	<0.001
Grammatical related	4	1	11	8	2	13	<0.001
Grammatical-related percentage	18.18	4.55	50	36.36	9.09	59.09	<0.001
Personal interest	5	0	11	12	5	29	<0.001
Personal interest percentage	7.58	0	16.67	18.18	7.58	43.94	<0.001
Nominal all	49	13	86	174	85	301	<0.001
Nominal all percentage	11.06	2.93	19.41	39.28	19.19	67.95	<0.001

Table 7 Factors affecting vocabulary acquisition in low-vocabulary and high-vocabulary children of the three age groups

	Low-vocabulary		High-vocabulary		*P* value
	Mean	SD	Mean	SD	
	Group A				
Maternal age	28	6	27	4	0.786
Mother with children (h)	10.5	2.7	8.8	2.8	0.125
Father with children (h)	3	1.3	4.6	1.8	**0.016**
TV	3	1.9	1	1	**0.002**
	Group B				
Maternal age	28	5	27	4	0.463
Mother with children (h)	9.6	3.6	9.6	3.6	1.000
Father with children (h)	5.6	4.6	5.6	2.4	1.000
TV	4.6	1.9	2.3	2.6	**0.002**
	Group C				
Maternal age	26	1	29	3	<**0.001**
Mother with children (h)	7.5	2.3	9.3	4	**0.018**
Father with children (h)	3.6	1.7	4.1	3	0.402
TV	5.9	3.2	3.2	3.2	<**0.001**

Significant values bold.

the three age groups showed that there was a significant difference between the low-vocabulary and high-vocabulary groups only in the older age group (group C) for children who were born as the firstborn (Table 8).

Discussion
Estimation of vocabulary size in this study
The range of the vocabulary shown in this study (Fig. 1 and Table 1) indicates that there is a wide variation in number of vocabulary acquisition that goes with data collected from literature that pointed out that the variation in early lexical development is much larger than the variations that are usually observed in other maturational milestones like walking [8–10]. Individual differences among infants can be illustrated by the findings of Bates *et al.* [8] showing that infants aged 1 year spoke 0–52 words, those aged 1.4 spoke 0–347 words, and those aged 2.6 spoke 208–675 words. After 1.6 years, toddlers' vocabulary size increases very quickly [11,12].

When compared with other languages, at the age of 1.4 years, Estonian children's mean expressive lexicon was 43 words and median expressive lexicon was 29 words. Caselli *et al.* [13] reported that at the same age, Italian children had a mean lexicon of 27 words and

Table 8 Birth order as a factor affecting vocabulary acquisition in low-vocabulary and high-vocabulary children of all the three age groups

	Low vocabulary		High vocabulary		P value
	Count	%	Count	%	
Birth order			Group A		
1	3	50.0	3	50.0	NA
2	6	66.7	3	33.3	
3	6	100.0	0	0.0	
Birth order			Group B		
1	6	100.0	0	0.0	NA
2	3	33.3	6	66.7	
3	3	50.0	3	50.0	
Birth order			Group C		
1	12	36.4	21	63.6	0.047
2	0	0.0	9	100.0	
3	0	0.0	3	100.0	

NA, not available.

American-English-speaking children had a mean lexicon of 64 words. Hamilton et al. [14] reported median scores for British and American children of the same age, and the scores were 10 and 37 words, respectively. Stolt et al. [15] found that mean size of Finnish children's expressive lexicon was 17 words at the age of 1.3 years and 69 words at the age of 1.6 years.

A greater difference in the number of spoken words was recorded among different age groups of Slovenian-speaking toddlers aged 1.6–2 years (the vocabulary scores of these age groups varied from 46 words at 1.6 and 6–238 words at 2) and those aged 2–2.6 (the vocabulary scores of these age groups varied from 255 words at 2–424 words at 2.6) [16].

According to this study, it seems that the vocabulary size for Arabic infants and toddlers is less than other languages. This can be attributed to the difference in using the statistical variances mean (used in most studies) versus median (used in this study), in addition to the previously mentioned expectation that Egyptian early lexicon may include small vocabulary size, as Arabic is considered a morphologically rich language.

Pattern of language acquisition in this study
For studying the pattern of vocabulary acquisition, six categories were included. The nominal category was further subdivided into three categories: general nominals, specific nominals, and nominals related to personal interest. It is important to analyze the pattern of acquisition of nouns as well putting into consideration the distinction among different nouns as not all nouns serve the same function. Nouns can function as simple as labeling people, living things, and objects in their nearby surroundings; labeling things they are interested in as they begin to be exposed to more stimuli around them; nouns triggering some concepts as time; nouns triggering many details to get a whole picture as names of places (trees, flowers, birds, and butterflies in a garden); and names of occupation (person who wears a white coat, works in a hospital, and gives sick people medications is a doctor). This view was supported by Nelson et al. [17] who stated that nouns in early vocabularies might refer to a variety of concepts, not all of which can be classified as representations of objects.

In Tables 2 and 3, the high average percentages of onomatopoeic words and social interactional words were expected, as young children tend to acquire expressions to socially interact with others in addition to the experience they gain from the parental input that focuses on providing the young children with easily producible syllables related to basic semantic groups such as animal and transportation sounds.

The children then proceeded into labeling common objects in the environment around them (general nominal) and labeling their favorite toys giving them easy names (nominal related to personal interest). Few percentage of nominals are seen in the vocabulary repertoire of this young age group and this is in line with other studies [15,18], which stated that at the earliest stage of vocabulary development, noun dominance has not yet set in and that the first words tend to be from the categories of sound effects, games and routines, and people.

So in the previously mentioned distribution in groups B and C (Tables 2 and 3), it was noticed that different types of vocabulary emerged in group B. It has been demonstrated that adjectives and verbs start to increase

proportionally especially when the total lexicon exceeds 50 words [13,18].

The highest noun percentage is shown in group B, and then continues from group B to group C, especially in the acquisition of the two categories of general nominals and nominals related to personal interest. In contrast, the 'highest verb/action percentage' is shown in group B, with no significant difference in progression between groups B and C. It can be assumed that the presence of nouns with considerable number provokes further language development. The child has first to label nouns and then to name the qualities of objects (increasing number of descriptive words) and the relationship between them (growing number of functional words) as in line with other studies [19,20]. The growing number of nouns leads to decreasing share of both onomatopoeic words and social interactional words, although the two word groups were still found in older age groups B and C as well because the checklists did not put the frequency of usage of words into consideration.

The pattern of vocabulary acquisition shown in this study is in line with that in English speaking homes. The largest part of the child's early vocabulary is noun-like words and other 'frequent individual words which are expressions for displeasure or rejection (such as no) and various types of social interaction such as 'give and bye-bye', with verb-like and adjective-like words being the next most frequent category types [21]. The majority of the early vocabulary contains names of objects and animals where these early words tend to refer to things within the child's immediate environment in which the child could actively interact with, for example, mommy, doggie, or teddy [22]. Verbs are not as popular as nouns in the early vocabulary.

In contrast to the pattern seen in this study, sociolinguistic and typological differences can result in different shares of individual word types in the vocabulary of children [20]. Thus, for example, the larger share of verbs in the early vocabulary of Korean-speaking toddlers reflects the predominance of activity-oriented utterances over naming-oriented utterances in Korean, the frequent use of utterances that include only a verb, and sentence-final placement of verbs. The established differences in the vocabulary of toddlers speaking different languages (e.g. Korean and Mandarin Chinese) can also result from the methodological approach to language assessment, because assessment of Korean toddlers' spontaneous speech shows a high percentage of verbs in their

vocabulary, whereas the parental assessments show a higher percentage of nouns [20].

Using data from six very different languages, Gentler [23] argues that nouns are universally more accessible than verbs. The differences in the distribution of vocabulary repertoire in different languages could be attributed to the following: Cognitive ease of nouns representing objects: the view that nouns of countable and bounded objects are easily learnt than parts, features, and description of objects, ill-formed masses of material substances (plastic) [24]; the position of the word being salient (verbs are in the last position in sentences in Korean and Italian but not so in Arabic or English); the morphology of the words being simple (verbs in Korean but not in Arabic), as well as language structure which has an effect on the word acquisition speed. Korean-speaking and Chinese-speaking children acquire verbs earlier. In Korean language, subject–object–verb (hence SOV) is the common order of words. In addition, the subject and object can be omitted from sentences [25,26]. In Arabic, verbs are inflected more often than nouns, which might contribute to Arabic-speaking children having fewer verbs than nouns in their earlier expressive lexicon.

One of the objectives of this study was to have a more differentiated picture regarding characteristics of lexical development in low versus high vocabulary acquisition in the early ages. Gathered information on this issue might help us understand whether the low-vocabulary children followed the same pattern of vocabulary growth as high-vocabulary children, even though it would begin at a later age. If they do not show this learning trajectory, then this might be an evidence of atypical form of development, rather than delay.

It is important to know the nature of various word types of children with low-vocabulary and high-vocabulary so as to follow up with the low acquisition group, providing proper counseling to enhance their vocabulary acquisition that will reflect later on their grammatical development, as it is well known from literature and clinical practice that both grammatical development and lexical level are tied together [27].

Pattern of vocabulary in low-vocabulary and high-vocabulary groups in this study

In Tables 4–6, the results of this study pointed out that the vocabulary pattern of low-vocabulary group children may simply reflect a delay. However, differences were noted as well. The observed differences in vocabulary might be attributed to that

low-vocabulary group tends to produce words such as animal sounds and sound effects (onomatopoeic words) as this class of words is easier to produce than words classified as animals or transportations, and these words have easier syllables to be produced than words that have both closed and open syllables [e.g. meow vs. 'ʔotta' (cat), bib bib vs. ¶arabeya (car)].

Taken together, this study found that the highest increase in the vocabulary size and emergence of all types of different vocabulary categories were seen in the age group B in both low-vocabulary and high-vocabulary groups. High-vocabulary group preceeds low-vocabulary group in acquiring action and grammar-related words, so it seems thus reasonable to consider early higher lexical abilities as an indicator for early grammatical development and complexity. This finding supports the results of other studies, Rescorla *et al.* [28] in which lexical development is considered a valid predictor of further language acquisition.

Comparing factors that might affect early language development in low-vocabulary and high-vocabulary groups in this study (Table 7)

Language acquisition is an intricate process that involves auditory, linguistic, cognitive, and environmental factors [29]. Children's early language development is a dynamic process that is affected by a complex array of transacting factors from multiple levels of influence, such as genetics, sex, temperament, the child's own skills in other developmental domains (cognitive, motor, socioemotional), and a host of biological and social risk and resilience factors, for example, premature birth, prenatal exposure to substances, parental education, caregiver interactive style, parent–child mutual regulation, bilingualism, and other cultural influences [5]. Therefore, it was necessary to assess some factors that might affect the development of language during early childhood.

In this study, the following factors were fixed: parent's education, socioeconomic standard, and all children were full term healthy children with normal prenatal, natal, and postnatal histories with no serious medical illnesses.

Regarding maternal age

Children with high-vocabulary group had older mothers compared with mothers of low-vocabulary groups. This result supports other researches that found evidence that children of older mothers have higher outcomes [30].

In theory, there are various channels through which maternal age might affect children's lexical development outcome including the possibility that younger mothers may be emotionally unprepared for motherhood, that they may have less parenting skills than if they had become parents at an older age, that they may be less sensitive in responding to their infants, or that they may make more poorly informed choices about daycare, preschool, and kindergarten.

Regarding maternal and paternal contribution and time spent with children

The results shown in Table 7 could be explained by the fact that both the father and the mother have contribution to their children's language development in different ways. Mothers spent time in feeding and taking care of their children in their early age whereas fathers' role during this early time is more toward interacting with talking, vocalizing, and playing with their children and they are more affectionate. The quality and quantity of both fathers and mothers' talk and interaction affect early language development of their children. So these very early childhood years are a time in which parents might exert unique influence as their children rapidly acquire important foundational social, cognitive, and language skills.

These results are parallel to what was found in the literature that fathers provide material and social resources to young children and may contribute to the children's competency in acquiring communication and language skills [31,32] and that the father's education has been linked to greater income, more stable hours of work, and more benefits [33].

In particular, the language used by fathers in interactions with young children may positively affect early language development. On maternal language, input has indicated that the amount of talk mothers direct to their children is positively associated with their children's gains in linguistic abilities [2]. In particular, the diversity of maternal vocabulary has predicted children's later language development and literacy [34,35].

Regarding exposure to television

This study showed that higher vocabulary children in all age groups were exposed less to television. This is expected as exposure to television reduces interaction time between children and their parents in addition to the effect of content of the programs children are exposed to. This result supports some views from previous researches that report associations between screen time and cognitive development outcomes, such as short-term memory skills, academic achievement in reading and math, and language development [36–38].

This led the American Academy of Pediatrics to recommend that children younger than 24 months of age should not be exposed to television. It also suggests that children 24 months and older only be exposed to 2 h of screen time per day [29].

Parent–child interaction remains the most effective learning tool for young children. Some studies related to video and television such as by Anderson and Pempek [39] which showed that even though learning does occur when young children are taught through video and television programs, the learning experience is superior when it involves live displays and social interaction.

Regarding birth order (Table 8)

The result in Table 8 means that number of children who were firstborn and have high vocabulary acquisition is more than that those who were firstborn and have low vocabulary. Theoretically speaking, this is logical as first born children mainly get benefit from more attention of the parents and language model brought down to their level than late born children (second or third).

The studies in the literature showed controversial results. Some showed that early and late born children share the same general language development, as the late born children can benefit from witnessing conversation with parents and siblings, but they might be affected by the immature language of their older siblings while the early born that get benefit from more care and attention from parents. This is parallel to what was found in literature that child birth order has been linked with children's early vocabulary competence, with the first-born children demonstrating some advantages over later born children in language development [40,41].

Conclusion

This study gives a provisional idea about the vocabulary size of Egyptian infants and toddlers and provides additional information about the composition of their expressive vocabulary. There is a difference in linguistic composition of vocabulary in low-vocabulary and high-vocabulary groups. High-vocabulary group earned more types of words than low-vocabulary group in younger age. The main constituents at that young age were general nominals, onomatopoeic words, and social interactional words with emergence of other types of nominals (nominals related to personal interest), grammatical related, and action words in higher vocabulary group. Their higher linguistic skills

contributed to earlier emergence of grammatical structure. Both low-vocabulary and high-vocabulary groups showed the highest vocabulary size at 18–24 months age. So the vocabulary development in low-vocabulary group showed delay in addition to some characteristic differences in the linguistic composition. Some demographic factors played a role in early and high-vocabulary acquisition. The results of this study should be interpreted with caution owing to small size of the sample under study.

Conflicts of interest
There are no conflicts of interest.

References

1 Thal DJ, O'Hanlon L, Clemmons M, Fralin L. Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment. J Speech Lang Hear Res 1991; 42:482–496.

2 Huttenlocher J, Haight W, Bryk A, Seltzer M, Lyons T. Early vocabulary growth: relation to language input and gender. Dev Psychol 1991; 27:236–248.

3 Rowe ML, Raudenbush SW, Goldin-Meadow S. The pace of vocabulary growth helps predict later vocabulary skill. Child Dev 2012 83;508–525.

4 Khamis-Dakwar R, Al-Askary H, Benmamoun A, Ouali H, Green H, Leung T, Al-Asbahi K. Cultural and linguistic guidelines for language evaluation of Arab-American children using the Clinical Evaluation of Language Fundamentals (CELF); 2012.

5 Bates E. Explaining and interpreting deficits in language development across clinical groups: where do we go from here? Brain Lang 2003; 88:248–253.

6 Thorpe K, Rutter M, Greenwood R. Twins as a natural experiment to study causes of mild language delay. II. Family interaction risk factors. J Child Psychol Psychiatry 2003; 44:342–355.

7 Nelson K. Structure and strategy in learning to talk: monographs of the society for research in child development. Chicago, IL: University Press 1973. pp. 1-2.

8 Bates E, Marchman V, Thal D, Fenson L, Dale PS, Reznick JS, Harung JP. Developmental and stylistic variation in the composition of early vocabulary. J Child Lang 1994; 21:85–123.

9 Bleses D, Vach W, Slott M, Wehberg S, Thomsen P, Madsen T, Basboll H. The Danish Communicative development inventories: validity and main developmental trends. J Child Lang 2008; 35:651–669.

10 Devescovi A, Caselli MC, Marchione D, Pasqualetti P, Reilly J, Bates E. A cross linguistic study of the relationship between grammar and lexical development. J Child Lang 2005; 32:759–786.

11 Bates E, Goodman JC. On the inseparability of grammar and the lexicon: evidence from acquisition In: Tomasello M, Bates E, editors. Language development: the essential readings. Oxford, UK: Blackwell 2001. pp. 134-162.

12 Fernald A, Pinto JP, Swingley D, Weinberg A, McRoberts GW. Rapid gains of verbal processing by infants in the 2nd year In: Tomasello M, Bates E, editors. Language development: the essential readings. Oxford, UK: Blackwell 2001. pp. 49-56.

13 Caselli MC, Bates E, Casadio P, Fenson J, Fenson L, Sanderl I, Weir J. A cross linguistic study of early lexical development. Cogn Dev 1995; 10:159–199.

14 Hamilton A, Plunkett K, Schafer G. Infant vocabulary development assessed with a British communicative development inventory. J Child Lang 2000; 27:689–705.

15 Stolt S, Haataja L, Lapinleimu H, Lehtonen L. Early lexical development of Finnish children: a longitudinal study. First Lang 2008 28:259–279.

16 Marjanovic-Umke L, Fekonja-Peklaj U, Podlesek A. Characteristics of early vocabulary and grammar development in Slovenian-speaking infants and toddlers: a CDI-adaptation study. J Child Lang 2013; 40:779–798.

17 Nelson K, Hampson J, Kessler Shaw L. Nouns in early lexicons: evidence, explanations and implications. J Child Lang 1993; 20:61–84.

18 Schults A, Tulviste T, Konstabel K. Early vocabulary and gestures in Estonian children. J Child Lang 2012; 39:664–686.

19 Caselli MC, Casadio P, Bates E. Lexical development in English and Italian. In: Tomasello M, editor. Language development: essential readings in developmental psychology. London, UK: Blackwell 2001. pp. 78-110.

20 Conboy BT, Thal DJ. Ties between lexicon and grammar: cross-sectional and longitudinal studies of bilingual toddlers. Child Dev 2006 77:712–735.

21 O'Grady W, Dobrovolsky M, Aronoff M. Contemporary linguistics: an introduction. 2nd ed. London, UK: Pearson 2011.

22 Gleason B, Ratner NB: Psycholinguistics. 2nd ed. USA: Harcourt Brace; 1993.

23 Gentler D. Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In: Kuczaj S, editor. Language development: language, thought, and culture. Hillsdale, NJ: Erlbaum 1982. 2. pp. 301–333.

24 Hollich GJ, Hirsh-Pasek K, Golinkoff RM, Brand RJ, Brown E, Chung HL. Breaking the language barrier: an emergent's coalition model for the origins of word learning. Monogr Soc Res Child Dev. 2000; 65:123.

25 Gopnik A, Choi S. Do linguistic differences lead to cognitive differences? A cross linguistic study of semantic and cognitive development. First Lang 1990; 10 199–215.

26 Gopnik A, Choi S. Names, relational words, and cognitive development in English and Korean-speakers: Nouns are not always learned before verbs. In: Tomasello M, Merriman W, editors. beyond names for things: young children's acquisition of verbs. Hillsdale, NJ: Lawrence Erlbaum 1995. pp. 63–80.

27 Bates E, Goodman JC. On the emergence of grammar from the lexicon. In: MacWhinney B, editor. The emergence of language. Mahwah, NJ: Lawrence Erlbaum 1999. pp. 29–80.

28 Rescorla L, Roberts J, Dahlsgaard K. Late talkers at 2: outcome at age 3. J Speech Hear Res 1997; 40:556–566.

29 Chonchaiya W, Pruksananonda C. Television viewing associates with delayed language development. Acta Paediatr 2008; 79:977–982.

30 Leigh A, Gong X. Does maternal age affect children's test scores?. Aust Econ Rev 2010; 43:12–27.

31 Pleck JH, Masciadrelli BP. Paternal involvement by U.S. residential fathers: Levels, sources, and consequences. In: Lamb ME, editor. The role of the father in child development. Hoboken, NJ: John Wiley & Sons; 2004 pp. 222–271.

32 Pancsofar N, Vernon-Feagans L. Mother and father language input to young children: Contributions to later language development. J Appl Dev Psychol 2006; 27:571–587.

33 Booth A, Crouter A. Men in families. Mahway, NJ: Erlbaum 1998.

34 Pan BA, Rowe ML, Singer JD, Snow CE. Maternal correlates of growth in toddler vocabulary production in low-income families. J Child Dev 2005; 67:763–782.

35 Vernon-Feagans L, Pancsofar N, Willoughby M, Odom E, Quade A, Cox M. The Family Life Project Key Investigators. Predictors of maternal language input to infants during a picture book task in the home: family, SES, child characteristics, and the parenting environment. J Appl Dev Psychol 2008; 29:213–226.

36 Barr R, Lauricella A, Zack E, Calvert SL. Infant and early childhood exposure to adult-directed and child-directed television programming: relations with cognitive skills at age four. Merrill-Palmer Q 2010; 56:21–48.

37 Tomopoulos S, Dreyer BP, Berkule S, Fierman AH, Brockmeyer C, Mendelsohn AL. Infant media exposure and toddler development. Arch Pediatr Adolesc Med 2010; 164:1105–1111.

38 Zimmerman FJ, Christakis DA. Children's television viewing and cognitive outcomes: a longitudinal analysis of national data. Arch Pediatr Adolesc Med 2005; 159:619–625.

39 Anderson D, Pempek T. Television and very young children. Am Behav Sci 2005; 48:505–522.

40 Bornstein MH, Leach DB, Haynes OM. Vocabulary competence in first- and second born siblings of the same chronological age. J Child Lang 2004; 31:855–873.

41 Hoff-Ginsberg E. The relation of birth order and socioeconomic status to children's language experience and language development. Appl Psycholinguist 1998; 19:603–629.