Mean Length of Utterance in Cypriot Greek-speaking Children

Louiza Voniati
European University, Cyprus
L.Voniati@euc.ac.cy

Abstract

While research undertaken worldwide indicates that mean length of utterance (MLU) is a valuable index in investigations of child language development, to date there have been no studies exploring MLU in pre-primary Cypriot Greek (CYG)-speaking children. The participants in this study were 36 monolingual CYG-speaking children at ages 36, 40, 44 and 48 months, with a typical course of language development. The findings demonstrated that MLU counted in words (MLU-w) of typically developing CYG-speaking children had a positive correlation with age (from 36 to 48 months) and a non-significant difference in MLU-w by gender at each age point, and that typically developing CYG-speaking children, for the age range studied, tended to produce more multiword utterance types. An outcome of this study is an MLU-w database which could be used, with some caution, in the language assessment of a similar population or as the basis for future studies. Areas for further research are identified.

Keywords

mean length of utterance – Cypriot Greek – typical language development

1 Introduction

Non-standardized or informal language assessment methods offer alternative approaches to describe and analyse communication in a variety of naturalistic contexts with numerous interactive partners (Thomas-Tate et al., 2006; Heilmann, 2010a). The most popular among these is language sampling (Spaulding
et al., 2006), for samples can be collected as often as necessary, without performance bias, and children’s language production can be documented across a wide range of linguistic levels (Hewitt et al., 2005; Schuele, 2010).

Furthermore, there is evidence that language sample analysis (LSA hereafter) legitimizes the ordinary talk of every child as a clinical resource and may be less vulnerable to dialect and cultural variations than traditional formal tests (Stockman, 1996). LSA is considered an alternative language assessment procedure that brings cultural sensitivity, validity, accessibility and flexibility to the screening process (Gutierrez-Clellen et al., 2000). The mean length of utterance (MLU) measurement, which samples a natural behavior of children and which may also be less susceptible to the impact of dialect in comparison with formal testing, forms the backbone of LSA (Owens, 2003; Klee et al., 2004; Parker and Brorson, 2005).

Current clinical practice in Cyprus employs a variety of informal language assessment procedures in order to distinguish ‘typical’ and ‘impaired’ language development in young children. Informal assessing is the only available procedure because there are no standardized tests regarding Cypriot Greek (CYG), the linguistic variety of interest in the study on which the current paper reports. In addition, while the MLU has been used for almost a century to quantify children’s language production in certain languages and linguistic varieties, there are no reference databases on MLU scores in pre-primary children in the context of CYG. Therefore the aim of the study was to contribute to knowledge on the relationship between age, gender and MLU among CYG-speaking children (for the age range 36 to 48 months).

2 The Context of the Study

The linguistic situation in the Republic of Cyprus forms the backdrop to the present study. In the non-occupied parts of the country, the majority of the population consists of native speakers of CYG (Terkourafi, 2007). CYG is a southeastern dialect of Greek and it is not a codified system (Newton, 1972).

CYG is an under-studied variety, at least in terms of the morphosyntactic grammatical system and developmental stages in young learners (Grohmann, 2011). CYG exhibits a range of several syntactic and morphosyntactic patterns that are not found in Standard Modern Greek (SMG) and have not yet been studied (Yiakoumetti et al., 2005; Tsiplakou, 2009). The most important differences between SMG and CYG are in the domains of phonology and the lexicon (Grohmann, 2011), but with the development of mass media, CYG is becoming strongly influenced by SMG (Arvaniti, 2007).
The CYG phonological inventory consists of 29 consonants defined according to features of place and manner of articulation and like most Modern Greek varieties exhibits a five-vowel system (Petinou and Okalidou, 2006). Another topic that has been barely touched upon by researchers is the prosody of CYG, and in fact, little is known about its rhythm and prosodic structure (Terkourafi, 2007). Moreover, it does not have a generally accepted orthography and there is only one extant description that partially accounts for its grammar (Tsiplakou, 2006). Further, there are few studies on the acquisition of CYG: for example, there is some limited research on the phonetic/phonological development of CYG (Petinou and Okalidou 2006) and only a single study on clitic production (Petinou and Terzi, 2002).

Recently, Petinou et al. (2011) examined longitudinally the emergence of various linguistic measures, such as receptive and expressive vocabulary, phonetic production and MLU-w in 18 CYG-speaking children classified as late-talkers and 18 age-matched CYG-speaking children with typical course of language development. Participants were assessed at 28, 32 and 36 months. The two groups exhibited parallel developmental profiles, with a language lag favoring the late-talker group as compared to their typical counterparts. The linguistic lag persisted throughout the course of the investigation. Further studies following children across longer time periods are necessary in order to accurately describe the trends and the times at which critical markers (MLU-w included) appear in the language development of CYG-speaking children.

In general, it seems that CYG is an under-described dialect in a highly intriguing environment where bidialectism, bilingualism, and even trilingualism co-exist (Okalidou et al., 2010). Suffice it to say that language development of CYG-speaking children is quite limited, and many of its aspects remain little understood. More specifically, to date, the language development of CYG-speaking children with typical or impaired language development has not been comprehensively studied. One aim of this paper is to make a contribution in relation to the specific issues described above in order to extend our knowledge regarding the language development of monolingual CYG-speaking pre-primary children.

3 MLU

3.1 MLU-w as Opposed to MLU-m
Worldwide, researchers and clinicians use several different ways to compute utterance length in different languages, with the most common being MLU counted in words (MLU-w) and MLU counted in morphemes (MLU-m). Several scholars have recommended the use of word counts instead of morpheme
counts for scoring MLU (Hickey, 1991; Parker and Brorson, 2005) and there is still considerable debate over this point and with regard to the inherent reliability of MLU.

While Brown (1973) outlined some rules for calculating the MLU-m, these have been criticized for their lack of explicitness as well as their need for continual *ad hoc* decisions (Dromi and Berman, 1982; Parker and Brorson, 2005). Research into highly inflected languages, such as Finnish (Bowerman, 1973), Swedish (Hansson and Nettelbladt, 1995) and Spanish (Linares, 1983), revealed that Brown's (1973) rules were problematic when applied to them. In smg, where gender, number, and case are encoded in a single morpheme in the nominal system, in the early stages of language acquisition children may identify only a subpart of the feature encoded in one morpheme (Stephany, 1997). Such cases show that those affixes used by a child do not accurately reflect his/her performance, as they lack full feature identification and specification. This creates reliability problems in the use of MLU-m.

Most researchers (Neuman, 2003; Thordardottir and Namazi, 2007) concur that MLU-w is simpler to implement, more reliable, and less arbitrary than MLU-m. It is considered a neutral procedure that can be used reliably across various languages and dialects (Le Normand et al., 2008), maintains cross-linguistic consistency and comparability and can be applied successfully in bilingual and dialectal research (Gutierrez-Clellen et al., 2009). Moreover, MLU-w has been shown to be strongly correlated with MLU-m in several languages other than English, including Dutch (Arlman-Rupp, 1976), Irish (Hickey, 1991), and Icelandic (Thordardottir and Weismer, 1998), which points to it being as effective a language measure as the MLU-m for these languages.

For the purposes of this study, MLU-w was deemed the most appropriate form because of the rich inflectional morphology of CYG (Tsiplakou, 2006). The calculation of MLU-w eliminates, for the most part, the inconsistencies of inflectional morphology pattern use (Thordardottir and Namazi, 2007). The necessity of making *ad hoc* decisions, combined with the time needed for counting morphemes, may reduce the reliability of the results where the MLU-m is employed. As such, the reliability of the MLU-w does not present a problem in terms of counting words.

3.2 *Utility of MLU*

MLU is considered an essential procedure in the assessment and monitoring of language production across a range of communicative contexts, and in time, may be considered the gold standard for assessment as it is a measure that is free from dialect or cultural bias (Miller and Deevy, 2003). Further, the MLU-w approach has been recommended as an unbiased measure.
for use with children who speak dialects and with those who are learning multiple languages (Craig et al., 2003).

Overall, this is a valuable index that is widely relied upon in investigations of children’s language development. Some suggested uses for MLU-w include: determining the stage of language development (Laing and Kamhi, 2003); identifying language impairment (Petinou et al., 2011); guiding further language assessment, and, measuring change in language production (Paul, 2007). Numerous researchers have suggested MLU-w as an alternative and supportive language tool which can be used alongside standardized language ones (Craig and Washington, 2000; Costanza-Smith, 2010; Yoder et al., 2011). In addition, the MLU-w is considered a very useful way of benchmarking general language acquisition (Laing and Kamhi, 2003; Leonard, 2004).

Despite the widespread use and acceptance of MLU, it has been criticized on a number of grounds. For instance, the specific sampling context may affect the child’s score (Dethorne et al., 2005). Moreover, in order to establish the validity and utility of measures of language performance derived from language samples, it is necessary to address a range of problems which may arise concerning the reliability of sampling and the replicability of measurement (Johnston, 2001). More specifically, with regards to the use of MLU-m, Eisenberg et al. (2001) argued that a number of problems impede robust research, including the interpretation and application of available reference datasets, the language sample size, and utterance identification and selection, all of which can pose threats to validity.

3.3 MLU in Accordance with Age and Gender

Although the developmental course of MLU is not universally accepted, it has been the focus of several studies carried out cross-linguistically (Eisenberg et al., 2001). An interesting insight regarding the use of MLU is the fact that MLU normative scores in one language cannot be compared to that same MLU in another language (Thordardottir and Weismer, 1998). The developmental norms for the various dimensions and aspects of one language cannot be transferred to another; MLU scores must make reference to the specific language (Hickey, 1991).

A number of studies targeting MLU-w in several languages including Dutch (Arlman-Rup, 1976), Irish (Hickey, 1991), Icelandic (Thordardottir and Weismer, 1998), and Malaysian (Kok, 2011) have shown that MLU-w and age are highly correlated in young children with typical language abilities. Many such studies are not consistent with regard to their findings in terms of measuring the means and age levels, but nonetheless the MLU-w is found to be developmentally sensitive.
As such, age-MLU correspondence can be employed to identify children whose language development requires further attention (Eisenberg et al., 2001) and can also be used to identify children at a particular MLU stage and to group them accordingly. Additionally, in educational and research settings, children may be selected first by age and then by MLU. Hence, reference databases regarding the MLU of children can be important sources of developmental age-referenced data. It should be noted that there are several caveats when using these databases. For example, robust comparisons can only be made when the children to be assessed are from populations similar to the one sampled in the research, and MLU-w scores are obtained in the same way.

Regarding the effect of gender on MLU development, Bornstein et al. (2000) found MLU-m to be higher for two-year-old English-speaking girls than for their male counterparts. Apart from data reported for speakers of English, significant gender differences were reported regarding the MLU-w of 180 Cantonese Chinese-speaking pre-schoolers ranging in age from 3 to 5 years, with girls producing significantly and consistently longer utterances than boys (Tse et al., 2002). This finding was consistent with comparable data from a large-scale study on Mandarin Chinese-speaking children in Taiwan studied by Yang and Zhang (1974), where significant effects for gender were reported, in favor of girls’ MLU counted in characters (MLUc) as compared to that of boys’.

By contrast, in Indian Konkani-speaking children in the age range of 3 to 5 years, the MLU-w was found to be higher in boys than in girls (Boland Arif, 2008). Moreover, regarding the MLU-w of Parisian French-speaking children in the age range of 2 to 4 years, a small advantage for girls over boys until 36 months of age was revealed (Le Normand et al., 2008). Further to these studies, some researchers have reported no clear outcomes linking age with MLU. For instance, Kok’s (2011) study indicated that for 130 children in the age range from 1:0 to 6:11, gender did not correlate with MLU-w and MLU counted in syllables (MLUs). Similarly, the findings of Kazemi et al. (2011) revealed that for 171 Iranian Persian-speaking children aged between 30 and 66 months, there was no statistical difference between boys and girls.

From the extant research in the field, there is no widespread agreement regarding whether the above findings with respect to MLU are due to factors such as socioeconomic groupings, numbers of children being studied, the procedures and methods used to measure the MLU or, in fact, a complex combination of these. Notwithstanding, at present for CYG, there are no reference databases on MLU scores in pre-primaries (36 to 48 months of age), nor are there research findings about the relationship between age, gender and MLU.
3.4 *MLU as Opposed to Other Measures of Language Development*

Many researchers have provided evidence that *MLU* may be neither as holistic a measure as Brown (1973) and Nice (1925) suggested, nor well correlated with other predictors of language complexity and development (Dethorne et al., 2005; Eisenberg et al., 2001; Klee and Fitzgerald, 1985; Leadholm and Miller, 1992; Rollins et al., 2000). According to Rollins et al. (2000), a combination of different language measures, including the Index of Productive Syntax, Number of Different Words, the MacArthur Communicative Development Inventories, vocabulary size, Developmental Sentence Scoring, Type-Token Ratio, Total Number of Words and others, may be more accurate indicators of child language acquisition than a single general indicator such as the *MLU*. For, although it can provide valuable information about certain aspects of language development, no one measure can provide all of the information necessary to understand the language development of an individual.

Some studies on *MLU* refer to other indicators of language competence from a more qualitative point of view. One such additional measure is qualitative analysis of the utterances with respect to lexical diversity (Silverman and Ratner, 2002; Thordardottir, 2005). As children's single, two- and multiword utterances are produced to express content or meaning (Bloom and Lahey, 1978) their ability to produce lengthier and therefore more meaningful utterances emerges (Miller, 1981).

The combination of more than one word per utterance marks the child's realization that words can be combined in systematic ways to express meanings that cannot be expressed through words used in isolation (Bloom and Lahey, 1978). Typically, children start to combine words when they are between 18 and 24 months of age. At around 30 months of age, children's utterances become more complex and it is expected that, by this time, almost all children are producing multiword utterances (Miller and Chapman, 1981). Although children may develop at very different rates, the transition from two-word to multiple-word utterances marks the next stage of language development where children add grammatical features that hold together adult sentences. These include features such as tense and number markers, possessive markers, helping verbs, and certain prepositions (Lieven, 2006).

From a qualitative point of view, a relative development index is constructed for describing the type of utterances children tend to produce in terms of single, two-word, or multiword ones. This is based on the assumption that the production of an increasing number of words and elements in an utterance reflects a greater lexical and grammar complexity (Rice et al., 2006). As for CYG-speaking children at ages 36, 40, 44 and 48 months, to date there has not been any published evidence regarding the type of utterances that they produce for
this age range. This study attempted to remedy this situation by providing data from the language samples of children with typical language development for the age range studied.

4 Goals of the Study

The objectives of the study were: to find the MLU-w of typically developing CYG-speaking children at four age points, 36, 40, 44 and 48 months; to examine the data for a possible correlation between age and MLU-w for these children; to examine the data to ascertain whether MLU-w for the children in this study varies according to gender; and last, to find the type of utterances (single, two- or multiword utterances) typically developing CYG-speaking children at ages 36, 40, 44 and 48 months tend to produce.

5 Methodology

5.1 Participants

The sample for this study consisted of 36 monolingual CYG-speaking children (18 girls and 18 boys). In terms of subject recruitment, the parents of the participants responded to an advertisement notice issued from the Early Childhood Research Laboratory (the ECRL hereafter), the Department of Education, of the University of Cyprus that was sent to nursery schools and paediatricians. At the beginning of the initial data collection phase, all children were 36 months old.

The criteria for selecting the child-participants were as follows: they were typically developing, had an unremarkable developmental and medical history as reported in their child case history files and came from a monolingual Greek Cypriot environment. The sample cohort was largely homogeneous with respect to ethnicity and socioeconomic status (SES). Their SES was defined broadly in terms of family income, parental education level, parental occupation, and place of residence (Cyprus Republic, 2011). In brief, the majority came from middle or upper SES homes and their parents had incomes higher than average, with the annual average national income being US$12,692 in 2005 (WHO, 2005). Most parents had completed at least their high school education, many had acquired their Bachelor’s and Master’s degrees, and they were all employed at the time of the research.
5.2 **Data Gathering**

Two pilot sessions were conducted prior to the main study. For this, as with the main study, written consent forms were completed by the parents agreeing to their child's participation in the research.

As part of a language assessment or intervention, in clinical practice the convention is to calculate **MLU** from a language sample of a minimum of 50 to 100 contiguous intelligible utterances (Eisenberg et al., 2001; Costanza-Smith, 2010) and 50 are generally considered sufficient for the purposes of analysis (Heilmann et al., 2010b). Accordingly, language samples of at least 50 utterances were used and any samples under this size were not included in the data analysis.

Furthermore, it is recommended that speech and language therapists (SLTs) select an elicitation method that suits the goal of the evaluation, the child's personality and his/her communication style (Owens, 2003). Activities for preschool children typically include forms of free play, which give the child an opportunity to produce spontaneous language (Leadholm and Miller, 1992; Costanza-Smith, 2010). Moreover, the range of topics covered can be varied and thus the language samples are likely to be diverse (Miller, 1981). For the purposes of this investigation, child-directed free play between the parent and child was selected as the appropriate elicitation method for data collection.

To date there has been little agreement regarding the materials that are optimal for obtaining a language sample from children (Miller, 1981; Lee, 1974), but tailoring them to the child's developmental level is advised (Nelson, 1998; Eisenberg et al., 2001; Owens, 2003). Therefore, a variety of materials considered appropriate for the children's developmental level were used (plastic food items, cups and plates, dolls, plastic cars and tractors, books, puzzles and pictures). These materials, the interactant (the parent), the place for data collection (ECRL), and the elicitation method (child-directed free play) were held constant across all children and across all data collection sessions. The language sampling procedures were uniform in order to eliminate any potential variability and/or distraction.

Each child in the author's presence played in the dedicated session room with the selected set of toys while conversing with his/her parent. There was a mix of fathers and mothers involved and their attendance depended on who was available to accompany the child to the ECRL. Four sessions in total were held for each child. That is, one session at 36, 40, 44 and 48 months of age and these sessions lasted approximately 30 minutes each. The children's language samples were recorded using an Olympus DS-2500 digital voice recorder with a compact microphone, under the supervision of the parent. The recorder was placed on the table in front of the child.
5.3 Transcription

The International Phonetic Alphabet (IPA, 1999) was used to transcribe the collected language samples. According to Bloom and Lahey (1978) the first few utterances of a language sample should be omitted since this could be considered a ‘warming up’ period. However, during transcription, the identification of initial words and utterances is a key issue with the potential of corrupting the findings which are going to be used for analyzing the study (Leadholm and Miller, 1992). In this research, the first few utterances were not omitted in an attempt to retain valuable information (Leadholm and Miller, 1992). Moreover, although only 50 utterances of each language sample recorded for each participant were included in the data analysis, the entire language sample was transcribed.

In Smiley and Goldstein’s (1998) work, an utterance is considered a complete thought that is divided from other utterances by sentence boundaries, pauses and/or a drop in the voice. Miller and Chapman (1981) segmented utterances ‘primarily by apparent terminal intonation contour’ (155). However, this rule proved to be insufficient in the work of Miller and Chapman (1981), as there was inter-examiner disagreement for 10–15% of the utterances. Leadholm and Miller (1992) made some suggestions for utterance segmentation, considering pauses of greater than two seconds to determine utterance boundaries. They noted that segmentation is relatively easy ‘when the child is producing only one utterance per speaking turn’ (27).

Accordingly, for the purposes of this research, a combination of the pause and intonation rules (the use of changing pitch to convey information) was chosen to work consistently for identifying utterances throughout the language samples. The end of an utterance was indicated by a child’s definite pause (greater than two seconds) preceded by a drop or rise in pitch.

For determining utterance exclusion in a language sample, most textbooks suggest using Brown’s (1973) rules, according to which only fully transcribed utterances are included and all totally and partially unintelligible utterances are excluded (Miller and Chapman, 1981). In fact, deciding what constitutes words and utterances in a language sample is not a straightforward task and is controversial (Smiley and Goldstein, 1998). Some researchers recommend that imitations of the immediately prior adult utterance or exact self-repetitions should be excluded (Paul, 2007), while Lund and Duchan (1993) suggest excluding identical utterances, elliptical responses to questions, counting and other sequences of enumeration. They assert that only real, communicative information must come from the words and sentences.

In this study, the utterances that were memorized (e.g. in the case of rote passages) or imitated (immediately following a model utterance), were excluded...
for counting MLU-w, in an attempt to include the most representative of children’s language performance. Subsequently, MLU-w was counted by totaling up the words in each of the 50 selected utterances and dividing the sum by the number 50.

5.4 Scoring Reliability
Reliability in LSA procedures is concerned with the consistency of administration and scoring (Eisenberg et al., 2001). The accuracy of the transcription process rests on the presumption that the individual transcripts are accurate representations of the original oral production, thus providing clinically meaningful data. For the present study, in order to ensure accuracy of the transcription, one postgraduate student and one newly qualified SLT transcribed approximately 15% of randomly selected language samples. The accuracy of this independent transcription was checked by comparing their transcripts. The transcribers were blind to each other’s transcriptions and their decisions could not be influenced by the initial transcript. The accuracy estimate was calculated in terms of percentage agreement using the formula: number-of-agreements / (number-of-agreements + number-of-disagreements) (Eisenberg et al., 2001).

Overall, the differences observed between the two transcribers were minor, as reflected by high agreement percentages, including words identification rated at 97% (range = 93%–100%) and utterance segmentation at 93% (range = 90%–95%). To ensure consistency of scoring, the two transcribers’ calculated the total MLU-w in their transcripts and the level of agreement related to scoring issues for MLU-w was high at 98% (range = 97%–99%) (Cohen et al., 2009).

6 Results
6.1 The Mean MLU-w and Differences in MLU-w by Age and Gender
As explained above, the participant sample for this study consisted of 36 monolingual CYG-speaking children (18 girls and 18 boys). The two groups were considered separately to ascertain whether there is a gender effect on MLU-w in CYG for the above mentioned ages and the data for both were combined to obtain the normative MLU-w scores across the age range. The statistical analysis of the data was performed with the Statistical Package for the Social Sciences (SPSS). Prior to the statistical analyses, descriptive statistics were computed. The mean MLU-w of typically developing CYG-speaking children for the ages 36, 40, 44 and 48 months was 2.579 (SD=0.365), 3.032 (SD=0.515), 3.428 (SD=0.573),
TABLE 1 Means and standard deviations for $MLU_w$

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age in months</th>
<th>$M$</th>
<th>$SD$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls</td>
<td>36</td>
<td>2.590</td>
<td>0.332</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>3.064</td>
<td>0.458</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>3.506</td>
<td>0.574</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>4.020</td>
<td>0.693</td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>36</td>
<td>2.569</td>
<td>0.404</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>3.000</td>
<td>0.578</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>3.350</td>
<td>0.578</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>3.797</td>
<td>0.611</td>
<td></td>
</tr>
<tr>
<td>Combined—boys and girls</td>
<td>36</td>
<td>2.579</td>
<td>0.365</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>3.032</td>
<td>0.515</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>3.428</td>
<td>0.573</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>3.908</td>
<td>0.654</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 1](image)

**Figure 1** $MLU_w$ scores as a function of age and gender

and $3.908$ ($SD=0.654$), respectively. The descriptive statistics of $MLU_w$ as a function of gender and age are presented in Table 1.

These results are graphically represented in Figure 1, where $MLU_w$ scores appear to increase with age; also, the positive difference between the $MLU_w$ scores of girls compared with boys increased with age.
To investigate the effect of age and gender on MLU-w, a two-way (2×4) ANOVA analysis was run with MLU-w as the dependent variable and age and gender as the independent variables. A Mauchly’s test was run and indicated that the assumption of sphericity had been violated for the main effect of age, $\chi^2(5) = 29.767 (p < .0005)$. Therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.622$).

The effect of age was reported to be significant, $F (1.866, 63.456) = 233.833 (p < .0005)$, an effect of large size, $\eta^2 = .873$ (Cohen et al., 2009). Pairwise comparisons of the four age points using the Sidak adjustment indicated that all four age points were different from one another ($p < .0005$ in all cases). In particular, it was confirmed that at each age point, the MLU-w score was significantly higher compared with any previous age point. This can be seen in Table 2, where the standardized mean differences (Cohen’s d) increase with age: the greater the age difference, the greater the d value. The standardized mean difference is a measure of effect size showing how significant a difference between two samples means is: a d value of 0.5 is considered a medium effect size, while a d value of 0.8 and above is considered a large effect size (Cohen, 1977). This finding was further supported by a 1-tailed Pearson correlation analysis, which indicated a strong positive correlation between age and MLU-w ($r = .679, p < .0005$).

The effect of gender was not significant, $F (1, 34) = 0.476 (p = .495)$, partial $\eta^2 = .014$. The standardized mean difference between girls and boys was $d = 0.161$, with girls exhibiting higher MLU-w scores than boys (across all ages).

There was no significant interaction between age and gender, $F (3, 102) = 1.505, p = .218$, partial $\eta^2 = .042$. This finding was verified by simple main effects analysis: the difference in MLU-w between girls and boys at each age point was always positive, but none was significant, as can be seen in Tables 3 and 4.
Table 3: MLU-w difference (d) and significance between girls and boys as a function of age

<table>
<thead>
<tr>
<th>Age</th>
<th>Girls—Boys MLU-w difference (d)</th>
<th>Statistical significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>36 months</td>
<td>$d = 0.021$</td>
<td>$t(34) = 0.171, p = .865$</td>
</tr>
<tr>
<td>40 months</td>
<td>$d = 0.064$</td>
<td>$t(34) = 0.371, p = .713$</td>
</tr>
<tr>
<td>44 months</td>
<td>$d = 0.156$</td>
<td>$t(34) = 0.810, p = .423$</td>
</tr>
<tr>
<td>48 months</td>
<td>$d = 0.223$</td>
<td>$t(34) = 1.026, p = .312$</td>
</tr>
</tbody>
</table>

Table 4: Standardized mean MLU-w differences (d) between girls and boys as a function of age

<table>
<thead>
<tr>
<th>Age</th>
<th>Girls—Boys MLU-w difference (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>36 months</td>
<td>$d = 0.059$</td>
</tr>
<tr>
<td>40 months</td>
<td>$d = 0.127$</td>
</tr>
<tr>
<td>44 months</td>
<td>$d = 0.278$</td>
</tr>
<tr>
<td>48 months</td>
<td>$d = 0.352$</td>
</tr>
</tbody>
</table>

6.2 Type of Utterances in the Language Samples

The total number of single-word utterances that children 36 months of age produced in language samples of 50 utterances, ranged between four and 22 utterances. Accordingly, the total number of two-word utterances ranged between eight and 21 utterances and the total of multiword utterances ranged between 16 and 32. Most children at 36 months of age (28 out of 36 participants) produced more multiword utterances in a language sample of 50 utterances. Some of the children (six out of 36 participants) produced more two-word utterances, while others (two participants) produced more single-word utterances.

At 44 months of age, the total number of single-word utterances that children produced in language samples of 50 utterances, ranged between one and 18 utterances while the total number of two-word utterances ranged between four and 17 utterances and the total for multiword utterances ranged between 19 and 41. Almost all children at 40 months of age (35 out of 36 participants) produced more multiword utterances in a language sample of 50 utterances with only one producing more single-word utterances.

At 44 months of age, the total number of single-word utterances that children produced in language samples of 50 utterances, ranged between one and...
The total number of the produced two-word utterances ranged between four and 19 while the total of multiword utterances ranged between 21 and 45. All the children at 44 months of age produced more multiword utterances in the collected samples of 50 utterances.

At 48 months of age, the total number of single-word utterances that children produced in samples of 50 utterances ranged between one and 13 utterances. The total number of the produced two-word utterances ranged between two and 21 and the total of multiword ones ranged between 21 and 46. All the children at age 48 months produced more multiword utterances.

7 Discussion

One of the objectives of this study was to find the MLU-w of typically developing CYG-speaking children for the ages of 36, 40, 44, and 48 months and subsequently, to look at whether age and gender influenced MLU-w. To date, there are few published works pertaining to younger monolingual CYG-speaking children at 28 to 36 months of age (Petinou et al., 2011) or to monolingual Greek-speaking children for the age range 1:7 to 2:8 years (Marinis, 2003), thus it is difficult to compare the outcomes of the study with those from extant research results.

In various investigations with respect to other languages and their variants, the absolute mean values of MLU have been reported as being different, with either higher or lower values for the targeted age points (Thordardottir and Weismer, 1998; Kok, 2011). These differences may be explained in part by the differences in methodologies adopted by the researchers and the different linguistic features of the studied language. Differences among the samples may also have played a vital role. Additionally, the degree of competency of the transcribers is likely to account for some differences being reported in MLU values (Klee et al., 2004).

It is noteworthy that an extensive literature review revealed that there have been surprisingly few longitudinal studies that have directly examined the relationship of MLU with age. It can be asserted that longitudinal data provide the strongest evidence to date in terms of assuring robust reliability and the validity of MLU-w as an index of language acquisition and it is sustained development over time (Rice et al., 2010).

The results of the current study confirmed that at all four age points, the MLU-w increased with age. In particular, it was shown that at each age point, the MLU-w score was significantly higher compared with any previous age point. This was expected as language in most young children is developing rapidly.
in this particular age range. This finding confirmed that in the CYG context, this measure could be considered a useful developmental index that changes predictably for children within this age range.

These results are consistent with those of most studies targeting MLU-W in English, demonstrating a significant relationship between age and MLU-W among pre-schoolers (Parker and Brorson, 2005). It is also encouraging to compare the results of the present study with those of Rice et al. (2010), revealing that MLU-W appears to be a reliable and valid index of general language development and an appropriate grouping variable for English-speaking children from three to ten years of age. These findings are consistent with the findings of most studies carried out cross-linguistically regarding MLU-W, such as Dutch (Arlman-Rupp et al., 1976), Irish (Hickey, 1991) and Malaysian (Kok, 2011).

However, it should be noted that some prior studies concerning other languages and their variants (e.g. dialectical issues), suggest that the relationship between age and MLI is not a simple one (Owens, 2003). MLU-W may be a developmental index that increases with age but may also be experiential. More precisely, the more children are exposed to a rich vocabulary, the richer their vocabulary may become, and this is indicated in their utterances; this notion has been the focus of prior research studies (Owen and Leonard, 2001). Moreover, it is possible that the cited inconsistencies in relevant studies are due to the fact that in typically developing children the increase in MLU with age is expected, but the magnitude of the increase is likely to vary and may well be context-dependent (Klee and Fitzgerald, 1985; Conant, 1987). Accordingly, it is possible to hypothesize that individual differences in rates of development and the complexity of the language may also account for discrepancies in findings.

While a number of studies, the current one included, demonstrate a developmental course of MLU-W, this should not be interpreted as a pattern free of caveats. The correlation of MLU with age needs to be interpreted with caution. Firstly, the magnitude of the increase in MLU-W with age is likely to vary. Most importantly, a key potential issue is that age, per se, may not be the factor influencing children's language but, instead, it is simply a proxy for many developmental processes that we may or may not yet understand (Goffman and Leonard, 2000).

In respect of whether MLU-W varies according to the gender of the participant children, the study’s findings revealed that the difference in MLU-W between girls and boys at each age point was always positive in favor of girls but not statistically significant. This finding supports previous research into this area, as it has revealed that the effect of gender on MLU-W is not significant regarding a number of languages (Schachter, 1979; Bol and Arif, 2008; Le Normand et al., 2008; Kok, 2011). On the other hand, these results differ from
some published studies for the English language (Chapman, 2007) that have found that MLU-m varies according to gender, with girls having higher MLU-m values than boys. For instance, Bornstein et al. (2000) reported that MLU-m was higher for 2-year-old girls than for boys of the same age. However, it should be noted that there appears to be a gap in the extant literature regarding whether this finding holds true for older children.

In contrast to the present findings, Tse et al. (2002) reported significant gender differences regarding MLU-w, with girls producing significantly and consistently longer utterances than boys and the findings of Bol and Arif (2008) indicated that MLU-w was higher in boys as compared to girls. These inconsistencies might be due to a number of linguistic, cognitive, and sociocultural influences (Wray and Bloomer, 2006) on children's language performance. With respect to differences in MLU by gender, further studies are required to be carried out for other languages and for other age groups in order to determine whether girls or boys produce more words and lengthier utterances.

Turning to consider CYG, there are no data reporting on MLU by gender, thus it was impossible to make any relevant comparisons with extant studies. However, this factor is of great importance, because using combined reference data for boys and girls could lead to an over- or under-identification of boys’ or girls’ MLU-w. In the current study, data for the boys and girls was combined for the creation of the MLU-w normative scores database which was justified on the grounds that the analysis revealed no statistically significant difference between them at any age point.

The lattermost data presented above addresses the nature of the utterances in language samples of typically developing CYG-speaking children in the studied age range. In brief, with respect to the first two ages (36 and 40 months), most of the children produced more multiword utterance types and only a few produced more single- or two-word utterance types. At the two older ages (44 and 48 months), all the children produced more multiword utterance types.

In CYG there is no existing database pertaining to the types of produced utterances in children's language samples, thus it was impossible to make any relevant comparisons. However, the present results are consistent with the findings of most studies in a number of languages targeting types of utterances in children for the age range studied (Thordardottir, 2005; Lieven, 2006). Although there is no recognizable stage that marks the transition from two-word to multiple-word utterances, it is expected that almost all children after 30 months of age are producing multiword utterances (Paul, 2007). The present findings support previous research into this area for English-speaking children which suggested that during year three, most of transitional utterance forms...
are replaced by multiword utterance types thanks to the gradual increase in the number of words present in the child's lexical repertoire (Le Normand et al., 2008).

Although from some relevant studies (Smiley and Goldstein, 1998; Silverman and Ratner, 2002; Klee et al., 2007) it appears that there is consensus that during early pre-school years children produce more multiword utterances, the type of utterance is likely to vary in a language sample and may well be context-dependent. In addition it may be possible that any inconsistencies are due to individual differences in rates of development or to the complexity of the language (Owens, 2003). One of the issues emerging is that age, per se, may not be the key factor influencing the children's language but may be simply a proxy for developmental processes that we may or may not yet understand (Goffman and Leonard, 2000). Thus, these results regarding the type of utterances produced in early pre-school years need to be interpreted with caution.

8 Conclusion

The outcomes of the present study have contributed to a better understanding of MLU in CYG-speaking children and have helped to identify a number of avenues for future inquiry. The mean MLU-w scores for typically developing CYG-speaking children aged 36, 40, 44 and 48 months were successfully ascertained. This study first indicated that MLU-w scores for the above-mentioned children significantly increased at each subsequent age point, while gender played no role in the MLU-w scores. Based on the MLU-w analysis performed for the study, a performance database has been developed for typically developing CYG-speaking children aged from 36 to 48 months, in the context of standard speaking conditions.

Moreover, this study provides, for the first time, information on the type of utterances (single-, two-, or multiword) that typically developing CYG-speaking children at ages 36, 40, 44 and 48 months tend to produce. For the first time, researchers and clinicians who analyze and assess the language of CYG-speaking children will be able to employ performance standards for MLU-w and the type of utterances they tend to produce. Such language measures have been used in other counties and linguistic varieties, but to date none are being employed for CYG-speaking children for the age range studied. The implications of the current study can have a positive impact on several areas, including research of language and language assessment of CYG-speaking children aged from 36 to 48 months.
However, there are some points that should be considered in terms of the generalizations that can be made from this study data. An important limitation of the research lies in the small sample, which limits the trustworthiness of any generalizations. Furthermore, the investigation was not specifically designed to evaluate all factors related to MLU-w, and the participants were largely homogeneous with respect to their ethnicity and SES. As the findings of this research pertain to average scores, individual differences and outliers are not accounted for in the discussion. Therefore, an individual CYG-speaking child's trajectory may or may not follow the patterns described in this study, and the reference database for MLU-w produced provides only a comparison sample but not defined norms.

Continued investigation of potential applications of MLU-w and its measurement properties is warranted, as are the development of more extensive normative data and the qualitative analysis of the produced utterances (for instance in terms of content and function words). More specifically, a longitudinal investigation of MLU-w growth across different ages in CYG-speaking individuals with typical or impaired language development and further analysis of the children's language samples from a qualitative point of view is a pressing need. In addition, in the specific context of CYG speakers, other factors such as language-processing constraints and language measures related to MLU-w would benefit from research. This study can also serve as a strong base for future studies in LSA for other populations and age ranges.

As globalization continues, there is a pressing need for unbiased authentic language assessment of children’s oral language skills, and LSA measures prove to be an important part of researcher’s and clinicians’ work with bilingual and multilingual children. Given the change in the cultural composition of Cypriot society, future research of a similar nature could be conducted targeting populations other than children from a monolingual Greek Cypriot environment.

References


Grohmann, K. (2011) ‘Some directions for the systematic investigation of the acquisition


Kok, L.T. (2011) A Comparison of Mean Length of Utterances (MLU) on Mandarin Child Language Data of Chinese Children within the Age Ranges of 1; 0–6; 11 Years Old: MLUs (syllable) and MLU-w (word), unpublished PhD Thesis, University of Kebangsaan, Malaysia.


