

Chondrogianni, V & Marinis, T. (2011). Asynchronous development of vocabulary, morphology and complex syntax in successive bilingual children: differential effects of internal and external factors. *Linguistic Approaches to Bilingualism* 1 (3), 318-345.

Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children

Vasiliki Chondrogianni^{1,2} & Theodoros Marinis¹

1. School of Psychology & Clinical Language Sciences, University of Reading

2. Aalborg University, Denmark

Running title: Differential effects on the development of language domains in L2 children

Corresponding author:

Vasiliki Chondrogianni

E-mail: v.chondrogianni@reading.ac.uk

Address: School of Psychology and Clinical Language Sciences, University of Reading,

Harry Pitt Building, Reading RG6 6AL

ABSTRACT

The present study investigates the effects of child internal (age/time) and child external/environmental factors on the development of a wide range of language domains in successive bilingual (L2) Turkish-English children of homogeneously low SES. Forty-three L2 children were tested on standardized assessments examining the acquisition of vocabulary and morpho-syntax. The L2 children exhibited a differential acquisition of the various domains: they were better on the general comprehension of grammar and tense morphology and less accurate on the acquisition of vocabulary and (complex) morpho-syntax. Profile effects were confirmed by the differential effects of internal and external factors on the language domains. The development of vocabulary and complex syntax were affected by internal and external factors, whereas external factors had no contribution to the development of tense morphology. These results are discussed in light of previous studies on the impact of internal and external factors in child L2 acquisition.

Keywords: child L2 acquisition, internal and external factors, profile effects

Introduction

A considerable amount of research has revealed that sequential bilingual (L2) children are less accurate in standardised assessments compared to monolingual (L1) children (Marinis & Chondrogianni, 2010; Paradis, 2005). Their performance has been shown to resemble the performance of children with Specific Language Impairment (SLI) in terms of their overall accuracy (Hakansson & Nettelbladt, 1996; Paradis, 2008), but some recent studies have also revealed qualitative differences between typically developing (TD) L2 children and children with SLI (Chondrogianni & Marinis, under review; Gutierrez-Clellen, Simon-Cereijido, & Wagner, 2007; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005; Marinis & Chondrogianni, 2010, 2011; Paradis, 2005; Paradis, Rice, Crago, & Marquis, 2008).

Research on L2 children has focused mainly on their language abilities comparing them to TD L1 children and children with SLI with limited attention to the impact of age/time related variables internal to the child, such as age of onset (AoO), length of exposure (LoE), and age at the time of testing (AToT) and environmental or external factors, such as, parents' level of education, use of English in the home, amount of language input, and L2 proficiency of the parents. Moreover, many existing studies addressing internal and external factors have investigated primarily heterogeneous populations with varied L1 backgrounds, a wide range of age of onset and years of exposure, and different socio-economic status (SES). The present study addresses systematically the extent to which internal and external factors affect L2 children's performance on a range of standardised assessments widely used in clinical settings tapping morpho-syntax and vocabulary in a

relatively homogeneous group of L2 children who have Turkish as their L1 and English as their L2.

Effects of internal and external factors on the acquisition of morpho-syntax vs. vocabulary

Child L2 learners, i.e. learners exposed to the L2 between the ages of 3-4 and 7 years old (Schwartz, 2003), constitute an interesting population to study the interplay between internal and external factors and how they affect language development, since L2 children differ from their L1 counterparts in a number of important ways. Firstly, child L2 learners are biologically and cognitively more mature than L1 learners when they are first exposed to the L2 (Schwartz, 2003; Unsworth, 2005). This cognitive maturity has been shown to affect the rate of acquisition of lexical material with L2 children with a higher age of onset (AoO) exhibiting an accelerated acquisition of vocabulary (Golberg, Paradis, & Crago, 2008).

Apart from cognitive maturity, child L2 learners upon exposure to the L2 already have a grammatical system in place with well-developed processing routines (Schwartz, 2003). This established grammatical system may exert an influence in the manner and rate of L2 acquisition and most studies in the field of child L2 acquisition have primarily focused on the effects of L1 transfer as a child internal factor (Chondrogianni, 2008; Haznedar, 1997; Unsworth, 2005; Whong-Barr & Schwartz, 2002). This L1 transfer can be attested at the level of underlying syntactic representations without any effect at the level of the morpho-phonological exponents of the lexical items, suggesting that the acquisition of grammar and

morphology can be asynchronous and the two domains can be dissociated (Haznedar & Schwartz, 1997).

A third parameter distinguishing L2 children from their L1 counterparts is the amount of input that the former receive in each language. Children exposed to the L2 in a setting outside the home receive less input than L1 children exposed to a single language in both the home and social settings. This difference is both quantitative and qualitative in terms of the contexts that each language may be used in and is influenced by factors, such as the educational system, the status and the power relations between the two languages, and the institutional support that the minority language receives (Oller & Eilers, 2002).

Most studies examining child L2 acquisition have primarily focused on age/time related or child internal factors such as AoO, LoE, and AToT by investigating their effects on language outcomes and/or rate of acquisition of primarily morpho-syntactic phenomena (Unsworth, 2005). External or environmental factors, such as parents' level of education, use of English in the home, amount of language input, and parents' L2 proficiency have predominantly been related to the acquisition of vocabulary or general academic achievement (Oller & Eilers, 2002). Very few studies have compared directly the acquisition of morpho-syntax with the acquisition of vocabulary in the same group of children looking at the extent to which internal and external factors differentially affect each one of those domains (Bohman, Bedore, Peña, Mendez-Perez, & Gillam, 2010).

Importantly, several of these studies have shown that internal and external factors differentially affect distinct domains of language, such as morpho-syntax, vocabulary, and reading, and give rise to profile effects, namely to a differential or asynchronous acquisition

of the various language domains (Bohman et al., 2010; Oller, Pearson, & Cobo-Lewis, 2007; Verhoeven, 2000). Profile effects may also be related to the degree and level of transfer and generalization that is allowed between the two languages depending on the domain, e.g. what can be transferred from the L1 and what can be generalised at the level of vocabulary and what at the syntactic or morphological level.

In terms of the acquisition of morpho-syntax, several studies have shown a negative association between AoO and language outcomes or rate of acquisition. For example, Jia & Aaronson's (2003) longitudinal study with ten Chinese children and adolescents who immigrated to the US when they were 5 to 16 years of age showed that earlier AoO affects participants' performance, but this effect was evident only after three years of exposure to the L2. However, this study also showed differences in the environmental factors between participants with early and late AoO. Children with early AoO had a higher preference for carrying out activities, such as reading or listening in the L2 compared to children with late AoO. So, the effect of AoO on the acquisition of morpho-syntax was confounded by environmental factors. Jia & Fuse (2007) also looked at the effect of age/time related (internal) factors, such as AoO and LoE on the acquisition of morpho-syntax, such as tense and non tense morphology. Children who started learning English at younger ages were more accurate after five years of exposure than children who started learning English at an older age for the most difficult grammatical forms. In addition, there was a positive correlation between the production of grammatical morphology and LoE to English.

Gutierrez-Clellen & Kreiter (2003) investigated how external factors affect L2 children's morpho-syntax in a group of 7-to-8-year old Spanish-English children in the US,

using parental and teacher questionnaires and assessing children's language through a narrative task. Results showed that for Spanish, the language exposure variables combined (years of exposure, language input in the home, hours of reading and other literacy activities, parents' and teachers' rating of language use, and language input at school) seemed to account for 35% of the variance in grammatical utterances in Spanish and only the percentage of exposure to Spanish in the home was a significant predictor of grammatical performance in Spanish, accounting for 26% of the variance. However, for English there was no relation between any of the exposure variables on the one hand and grammatical performance in English on the other. This discrepancy was taken to be caused by the parents' limited English, and therefore, a difficulty to determine the children's input to English.

Finally, Gathercole (2002a; 2002b) investigated the acquisition of mass/count nouns and *that*-trace phenomena in English using a grammaticality judgment task in a large-scale study with Hispanic children in Miami attending 2nd and 5th grade. In the acquisition of mass/count nouns, there was an effect of SES for the children in the 2nd grade. High SES children were more accurate than low SES children in the 2nd grade, but this difference disappeared at the 5th grade. In the task tapping *that*-trace phenomena, there was also an effect of SES, but this was more prominent in children in the 5th grade. These results show that external factors do not affect all morpho-syntactic phenomena in the same way.

In contrast to the results on the acquisition of morpho-syntax, studies looking at the acquisition of vocabulary have shown a positive association between AoO and rate of acquisition. Golberg, et al. (2008) investigated the comprehension of single-word

vocabulary in a group of L2 children from a diverse L1 background. After 34 months of exposure, children performed within the low-average range for L1 children on the vocabulary task. Furthermore, children with AoO after 5 years made faster gains than children with AoO below 5. In addition, Goldberg et al. (2008) found that some environmental factors, such as mother's education level affected vocabulary outcomes, but there was no association between use of English in the home and vocabulary acquisition. However, English in the home was negatively correlated with mother's SES, i.e. mothers with high SES tended to use their L1 at home. This could be related to the prestige of the L1 because the group of L2 children had a mixture of high and low prestige L1s.¹

Cobo-Lewis, Pearson, Eilers, & Umbel (2002) conducted a large-scale study on the acquisition of vocabulary and literacy in bilingual school-aged children in Miami. This study showed an association between languages spoken in the home, instructional programme, SES, and vocabulary acquisition; vocabulary was more affected by SES and language spoken at home than literacy. The highest performance in vocabulary was observed in high SES children who spoke both English and Spanish at home and attended English-only programmes.

Vocabulary and reading abilities were also investigated in Oller et al. (2007), which reanalysed data from Oller & Eilers (2002). L2 children showed a comparable performance with L1 children in terms of reading, but lower vocabulary than L1 children in both English and Spanish. This profile effect and in particular the lower accuracy in vocabulary was taken to reflect the fact that vocabulary is partially distributed across the two languages.

¹ We would like to thank an anonymous reviewer for bringing this issue to our attention.

This is based on the idea that the children's experience with words is distributed across different contexts in the two languages. For some vocabulary items, L2 children know the words in only one language, whereas for others they know the translational equivalents in both languages. In this sense, bilingual bootstrapping (Gawlitzek-Maiwald & Tracy, 1996) was not observed in the domain of vocabulary due to the difference in context of use.

Importantly, profile effects were consistent across SES status, language spoken at home (Spanish-only, English and Spanish), and school setting (Immersion, English-Spanish).

One of the few studies that directly compared the acquisition of morpho-syntax with the acquisition of vocabulary in the same group of children and their interplay between external and internal factors is the one by Bohman et al. (2010). This is a large scale study with 756 Hispanic children in the US attending pre-kindergarten and kindergarten and compared how these children acquire morpho-syntax and semantics using a screening test assessing mainly language production. Accuracy in morpho-syntax was positively associated with English language input, output, and AToT. Accuracy in semantics, on the other hand, was related to the school district they lived in and whether or not the children received free lunch at school, i.e. SES affected the acquisition of semantics, but not of morpho-syntax.

Present study

The present study focuses on profile effects in Turkish-English L2 children by investigating associations of age/time related, child internal factors and environmental or external factors in a large range of standardised assessments used in a clinical setting that provide an index of grammar and vocabulary, but which also differentiate between tense morphology and

(complex) morpho-syntax in phenomena that are vulnerable in English children with SLI, i.e. tense marking, articles, passives, and wh-questions. These assessments were conducted in the same groups of children. This enables us to make not only between-group, but also within-group comparisons and compare the performance of the children in the different tasks used. This gave rise to the following research questions:

1. How do L2 children perform on standardized assessments and how do they compare with their L1 peers?
2. Are there any profile effects in the L2 children's performance on standardised assessments? Namely are there any differences between L2 children's performance on measures of general grammatical vs. vocabulary ability, and between tense morphology and complex morpho-syntax?
3. What is the relationship between age/time related, child internal and environmental or external factors and performance on language assessments?
4. What is the relative contribution of each factor on influencing child L2 learners' performance on different components of grammar and vocabulary?
5. Which combinations of factors predict L2 children's performance on standardized tasks and their ability to reach age-appropriate norms?

Methodology

Participants

Forty three typically developing (TD) successive bilingual Turkish-speaking children and thirty-three age-matched TD monolingual English-speaking children participated in the

study. The L2 children had a mean age of 7;8 (range:6;2–9;8; SD:12 months), and the L1 children a mean age of 7;4 (range:6;1–8;11;SD:8 months) ($F(1,73)=1.108, p>.2$). Both groups attended schools whose percentage of free school meals was well above the national average indicating low SES.

All L2 children were from the Turkish community in London and were recruited from schools with a high density in Turkish-speaking children. The monolingual children were recruited from schools in Reading. The selection criteria for the L2 children were that the language spoken at home should be Turkish, and the children should not have any history of speech and/or language delay or impairment. Additional information about the L2 children's years of exposure, quantity, and quality of input was collected through a parental and child questionnaire. The L2 children had an age of onset to the L2 between 2;6 and 5;0 years (mean=3;3;SD=6 months). The children's years of exposure to English ranged from 1;10 to 6;8 (mean=4;3;SD=14 months) depending on the age at the time of testing.

Materials and procedures

Parental questionnaire

A parental questionnaire (see Appendix) was designed to examine age/time related (internal), and environmental (external) variables that may influence language development and predict language attainment. The questionnaires were administered orally by Turkish research assistants to ensure that there was no language barrier, since most of the parents had a low proficiency in English. In the present study, a subset of the information elicited through the questionnaire is analysed.

Age/Time related (Internal) factors

The age/time related (internal) factors targeted in the questionnaire were AoO, LoE, and AToT. All children who participated in the study were born in the UK apart from four, who were born in Turkey and only three children had not attended an English-speaking nursery. Nursery attendance begun at the ages between 2;5 and 4;0 and only three children in the sample had begun schooling at the age of 5;0. Nursery attendance varied from full to part-time and some of the children also attended Saturday Turkish schools in London.

Environmental (External) factors

The external factors assessed through the questionnaire involved environmental variables including mother's and father's self-rated proficiency (mSRP, fSRP), mother's and father's level of education (mLOE, fLOE), and the use of English at home (ENGH). SRP was measured on a five-point scale (0=not at all, 1=a bit, 2=some, 3=very well, 4=advanced). The mothers had a mean proficiency of 1.7 (SD=1.01) and the fathers a mean of 1.8 (SD=1.05) suggesting that both parents' proficiency in English was quite low. Mothers had a mean of 7.6 years of education (range=4-16 years;SD=2.8) and fathers 8.1 (range=3-16;SD=2.9). Since the years of primary and secondary education in Turkey are 12, and SES is measured on the basis of the mother's educational level, this sample belongs to a population largely representative of low SES. The majority of the mothers were either unemployed or homemakers. Frequency of speaking English in the home (ENGH) with the parents and the siblings was measured on a five-point scale, and was transformed into

percentages for further statistical analyses (never=0%, occasionally=25%, often=50%, very often=75%, always=100%). The mean proportion of ENGH was 28% (range=0%-50%;SD=16.3) indicating low use of English in the home.

Standardized assessments

The standardized assessments used were the Test for Reception of Grammar-2 (TROG-2) (Bishop, 2003), the British Picture Vocabulary Scale II (BPVSII) (Dunn, Dunn, Whetton, & Burley, 1997), the Test for Early Grammatical Impairment (TEGI) (Rice & Wexler, 2001), and the Diagnostic Evaluation of Language Variation (DELV) (Seymour, Roeper, de Villiers, & de Villiers, 2005) .

The TROG-2 is a standardised receptive grammar task, in which children are shown four pictures on a panel and are instructed to point to the one that goes with the sentence uttered by the experimenter. The BPVSII involves comprehension of single-word vocabulary with outline and instructions similar to that of the TROG. The TROG-2 and the BPVSII are normed with L1 British English children. The children's raw scores are converted into standard scores (mean:100;SD=15).

The screening test of the TEGI was used to assess L2 children's production of tense marking, i.e. third person singular *-s* and regular and irregular past tense. Third person singular *-s* is elicited through picture probes depicting professionals engaging in various activities; past tense is elicited through picture pairs depicting ongoing and completed activities. The accuracy rate on each structure is added up to give rise to a single score. TEGI scores are criterion referenced, namely the accuracy rate on each morpheme is

compared to a cut-off score, above which performance can be considered in the range of TD children.

The DELV comprises four domains: syntax, pragmatics, semantics, and phonology. In the present study only the syntax domain was administered, namely the comprehension of *wh*-questions (10 items) and passives (10 items) and the production of definite and indefinite articles (8 items). The individual scores for each structure give rise to a syntax domain total raw score (overall=28). For each individual domain, the DELV has norm-referenced scaled scores, that is, scores of the same-age norm group. The TEGI and the DELV have been normed for American English. Although for the structures examined in the present study we do not expect to find any differences between American and British English, comparison of the children's performance to the norms should be treated with caution due to possible variations in word frequency between British and American English². Raw scores were used in the present study to carry out all statistical analyses for the TEGI and the DELV apart from the logistic regression.

Results

To examine how the L2 children perform in assessments standardized with monolingual children, we first analysed the results from the TROG-2, BPVSII, TEGI, and DELV in comparison to age-matched L1 children, as shown in Table 1.

² American words in the TEGI and the DELV that are infrequent in British English, such as 'baseball player', 'firefighter', 'soccer ball', were replaced by British equivalent words, i.e. 'cricket player', 'fireman', and 'football' respectively.

INSERT TABLE 1 HERE

To examine differences between the two groups on each task as well as to address any profile effects between the comprehension of grammar and vocabulary, we entered the standard scores from the TROG-2 and the BPVSII into a repeated-measures ANOVA with Group (L1,L2) as a between-subjects and Task (TROG-2,BPVSII) as a within subjects factor. This revealed a main effect of Group ($F(1,73)=49.280;p<.001,\eta^2=.403$), a main effect of Task ($F(1,73)=9.263;p<.01,\eta^2=.113$), and a significant interaction between Group and Task ($F(1,73)=15.702;p<.001,\eta^2=.177$). L2 children were less accurate than the L1 children on both tasks (TROG-2: $F(1,73)=10.082;p<.01,\eta^2=.641$); BPVSII: $F(1,73)=132.049;p<.001,\eta^2=.116$), but pairwise comparisons using Bonferroni correction showed that L1 children were equally good on the TROG-2 and the BPVSII, whereas the L2 children were significantly less accurate on the BPVSII compared to the TROG-2 ($F(1,42)=23.956;p<.001,\eta^2=.363$).

Subsequently, we examined whether there is a difference between the TEGI and the DELV, the two tasks that tap into morpho-syntax. Note, that the TEGI examines only the production of bound tense morphology (third person *-s*, past tense *-ed*), whereas the DELV assesses comprehension of complex syntax (passives and *wh*-questions) and production of free morphology (articles). To examine differences between the two groups on each task as well as to address profile effects between tense morphology and (complex) morpho-syntax, we entered the composite raw scores from TEGI and DELV into a repeated-measures

ANOVA with Group (L1, L2) as a between-subjects and Task (TEGI, DELV) as a within subjects factor. This revealed a main effect of Group ($F(1,73)=32.994, p<.001, \eta^2=.320$), a main effect of Task ($F(1,73)=47.177, p<.001, \eta^2=.403$), and a significant interaction between Group and Task ($F(1,73)=11.993, p<.001, \eta^2=.146$). L2 children were less accurate than L1 children on both tasks (TEGI: $F(1,73)=7.505, p<.01, \eta^2=.148$; DELV: $F(1,73)=55.762, p<.001, \eta^2=.443$). Pairwise comparisons using Bonferroni correction showed that both groups were better on the TEGI in comparison to the DELV, but this pattern was more pronounced for the L2 children (L1 children: $p<.01$, L2 children: $p<.001$).

In the subsequent analyses we focused solely on the performance of the L2 children. The asymmetry between comprehension of grammar vs. vocabulary and between tense morphology vs. (complex) morpho-syntax was further confirmed by examining the proportion of L2 children who were able to reach age appropriate norms on each task.

INSERT FIGURE 1 HERE

As Figure 1 indicates, the majority of the L2 children were able to reach age-appropriate norms for the TROG-2 (33 out of 45), whereas the most problematic tasks were the BPVSI and the DELV with only one third of the children reaching age-appropriate norms on these tasks. The acquisition of tense morphology was less problematic than (complex) morpho-syntax (DELV), almost half of the children reached age-appropriate norms on third person singular *-s* ($N=21$) and more than half on past tense ($N=28$).

To examine the relationship between age/time related (internal) and environmental (external) variables for the L2 children, we conducted non parametric Spearman's correlations between the age/time and environmental variables and the raw and standard scores on each subcomponent of the standardized tasks separately, as shown in Table 2.

INSERT TABLE 2 HERE

Better performance on most assessments was associated with longer exposure. For all components of the DELV and the raw scores on the BPVS, this relationship was moderate to strong; the same relationship was weak for the TEGI and the raw scores on the TROG-2. The standard scores on the TROG-2 and the BPVSII were not affected by LoE. Older age at the time of testing was also associated with better performance (raw scores) on most assessments, whereas AoO was weakly associated with complex syntax on the DELV (passives, *wh*-questions) and the overall score.

With regards to the environmental factors, of the six environmental variables analysed, better performance on vocabulary and complex syntax (comprehension of passives, *wh*-questions, and overall score) was weakly associated with mother's self-rated proficiency (mSRP), and the latter was also weakly associated with the raw scores on the TROG and the standard scores on the BPVS. The frequency of the use of English in the home (ENGH) was moderately correlated with the use of articles and the overall score of the DELV and weakly correlated with the performance on *wh*-questions and the past tense score of the TEGI.

To examine which age/time and environmental variables can better predict L2 children's performance on grammar, vocabulary, tense morphology, and (complex) morpho-syntax and the relative contribution of each factor, we analysed the age/time and environmental variables in three steps. First, correlations were carried out within the age/time related variables and the environmental variables to examine the nature and the strength of the relationship within internal and external variables.

INSERT TABLES 3 & 4 HERE

As shown in Table 3, age/time related variables were highly correlated with each other, especially LoE and AToT. This is because the majority of the children had a similar AoO, and therefore, the higher the age of the children at the time of testing, the more years of exposure they had. Within the environmental variables, mSRP was moderately correlated with mLoE and ENGH. This suggests that mother's perception of their proficiency in the L2 was in accordance with their level of education in the L1, and this also related to the degree to which they spoke English in the home.

All variables yielding significant bivariate correlations with the performance variables were selected for a multiple regression analysis using backward-elimination regression analysis as the method of regression with the standard and raw scores on the standardised assessments as outcome variables. Backward-elimination regression analysis was chosen to address any multi-collinearity effects, since some of the predictor variables were correlated with each other, suggesting that they may be responsible for roughly the same portion of

performance variance, and thus, could be cancelling each other out when entered simultaneously into the regression analysis. The factors that emerged as significant from this regression analysis were entered into a further multiple regression with block entry as the data entry method only for the tasks whose outcome was influenced by a contribution of both internal and external factors. These were TROG-2 (raw scores), BPVSII (raw scores) and DELV. Block entry was used to enter first the age/time related variables and then the environmental variables, as the former were shown to correlate higher with the outcome variables, and thus, it was predicted that they will account for a higher portion of the performance variance than the environmental variables. In order to avoid complex presentation of the data due to the number of variables and possible analyses, Tables 5, 6, and 7 show the results from the multiple regressions after eliminating the non-significant predictors for each task.

INSERT TABLES 5, 6 & 7 HERE

As shown in Table 5, for the raw scores on the TROG-2 a combination of LoE and mSRP can predict children's performance on this task, although the actual variation predicted is quite low ($R^2=.219$ or 21.9% of the performance variance). The multiple regression analysis with block entry further revealed that both of these factors can equally account for around 20% of the variation in performance (LoE: $R^2=.119$, mSRP: $R^2=.100$). Conversely, no factor could predict the variation on the standard scores on the same task. For the BPVSII, a combination of LoE, AoO, and mSRP can predict almost half of the

performance variance on the raw scores ($R^2=.462$ or 46.2%), with the block entry regression analysis showing again that the age/time related variables had a greater contribution to the outcome variance ($R^2=.352$ or 35.2%). The standard scores on the BPVSII can be very weakly predicted by mSRP ($R^2=.090$ or 9% of the performance variance).

For the TEGI, only a small proportion of the performance on *-s* and the Composite Grammar Score (CGS) can be accounted for by a combination of LoE and AoO (third person: $R^2=.244$ or 24.4%; CGS: $R^2=.188$ or 18.8%), whereas ENGH could very poorly predict performance on past tense (ENGH: $R^2=.085$ or 8.5%).

Turning to the various structures examined on the DELV, these were mostly affected by a combination of internal and external factors. Performance variance on articles was accounted by LoE ($R^2=.246$ or 24.6%), whereas performance on *wh*-questions and overall performance on the DELV could more strongly be predicted by a combination of LoE and mSRP (*wh*-.: $R^2=.456$ or 45.6%; overall DEL: $R^2=.483$ or 48.3%) with again a higher contribution from age/time related variables (overall DELV: $R^2=.361$ or 36.1%) and a lower contribution by environmental variables ($R^2=.123$ or 12.3%). Performance on passives was predicted to a low degree by age/time related variables such as LoE and AoO (passives: $R^2=.268$ or 26.8%).

Logistic regression analysis

To examine which factors or combinations of factors can predict L2 children's performance on reaching age-appropriate norms on the standardised tasks, we conducted a logistic

regression analysis. A logistic regression examines the probability of an outcome variable that is categorical or dichotomous with predictor variables that are either continuous or categorical. In the present study, the length of exposure is the continuous predictor variable and the child's performance the dichotomous outcome variable (1=able to reach age-appropriate norms, 0=unable to reach age-appropriate norms).

The predictor variables that emerged as significant in the previous multiple regression analysis were also entered into the logistic regression with the backward conditional in order to address any multi-collinearity effects. These were LoE, AoO, mSRP, and ENGH. The z-scores on each task were converted into binary values with 1 indicating that the score is within 1SD of the mean, and thus, within age-appropriate norms, and 0 indicating that the score does not meet age-appropriate norms and falls below 1SD from the mean. These scores were then inserted into the logistic regression as outcome variables. Here the results from the best model are presented. Tables 8 and 9 report the overall evaluation of the logistic model for the TROG-2, BPVSII, TEGI (third person), and DELV, the statistical tests for individual predictors, and the goodness-of-fit statistics. No logistic models are reported for past tense and the overall scores on the TEGI, because none of the factors accounted for the L2 children's ability to reach age-appropriate norms on these two measures.

INSERT TABLES 8 & 9 HERE

According to Table 8, the variable that emerged as a significant predictor of children's ability to reach age-appropriate norms for the TROG-2, the BPVS and the DELV was mSRP. In all three tasks, the likelihood ratio test was significant indicating that the logistic model provides a better fit to the data than the null hypothesis. The goodness-of-fit test was insignificant, suggesting that the model was fit to the data as well. For the TEGI, the significant predictor variable was LoE, and this model was good fit to the data, as indicated by the significant effect in the likelihood ratio test and the insignificant result in the goodness-of-fit test.

Table 10 is the general classification table that documents the validity of the predicted probabilities for each task.

INSERT TABLE 10 HERE

According to Table 10, with the cut-off set at 0.5, the prediction for L2 children who manage to reach age-appropriate norms on the TROG-2 is more accurate than for those who do not reach age-appropriate norms. This observation was supported by the magnitude of sensitivity (93.9%) compared to specificity (16.7%). Sensitivity in the present analysis measures the proportion of children who were able to reach age appropriate norms out of the number of children who were predicted to reach age appropriate norms (31 out of 33 children), and specificity measures the proportion of children who were not able to reach age appropriate norms out of the number of children who were predicted not to meet age appropriate norms (2 out of 12 children) (Peng, Lee, & Ingersoll, 2002). This is also shown

by the false positive and the false negative rate (Table 10). The false positive rate shows the proportion of children misclassified as reaching age-appropriate norms over all children predicted to reach age-appropriate norms. The false positive rate for the TROG-2 was 24.4%, i.e. 10 children out of 41 children, who were predicted to reach age-appropriate norms on the TROG-2, did not. The false negative rate measures the proportion of children misclassified as not reaching age-appropriate norms over all children predicted not to reach age-appropriate norms. This was 50%, i.e. 2 out of the 4 children who were predicted as not able to reach age-appropriate norms, did reach age-appropriate norms. This suggests that the model was better at predicting the ability to reach age-appropriate norms than the opposite. Additionally, the overall correction prediction was 73.3% which is an improvement over the chance level.

In the BPVSII, mSRP accounted for only 10% of the performance variance. The model was more accurate at predicting the proportion of children who are not able (specificity:96.7%) rather than those who are able (sensitivity:14.3%) to reach age-appropriate norms and the overall correction prediction is again an improvement from chance level (70.5%). For the third person singular *-s* from the TEGI, the significant predictor factor was LoE which accounted between 15% and 20% of the performance variation (Cox and Snell $R^2=.146$;Nagelkerke $R^2=.194$ ³) and the overall prediction was different from chance (68.9%). LoE was also almost equally accurate at predicting children

³ Cox and Snell R^2 and Nagelkerke R^2 are descriptive measures of goodness-of-fit. These are variations of the R^2 in linear regression. In logistic regression none of these renders the meaning of variance explained (Peng, Lee, & Ingersoll, 2002). Therefore, these two R^2 indices should be treated with caution and as supplementary to each other.

who reach age-appropriate norms and who fail to reach these norms, and the proportion of false positives and false negatives was a little more than 30% in both cases.

For the DELV, the significant predictor was mSRP which had an overall correction prediction of 78.6%. This external factor was more accurate at predicting the proportion of children who were not able to reach age-appropriate norms (specificity:92.9%) and less accurate for those who were within age-appropriate norms (sensitivity:50%).

Discussion

The present study investigated how Turkish-English L2 children perform on various standardised language assessments examining general grammatical abilities (TROG-2), vocabulary (BPVSII), tense morphology (TEGI), and (complex) morpho-syntax (DELV) in comparison to their monolingual TD age-matched peers, and addressed how this performance is affected by internal and external factors, i.e., time and age related factors on the one hand, and environmental factors on the other. This is one of the few studies to examine diverse language domains at the same time in the same population of L2 children (Bohman et al., 2010; Paradis, this volume) and to investigate the factors that can predict development and affect their ability to reach age-appropriate norms on widely used assessments.

Firstly, results showed that L2 children differed from their monolingual age-matched peers on all assessments regardless of the task and the language domain examined. This is in line with previous findings showing that L2 children are less accurate in standardised assessments compared to monolingual children (Gutierrez-Clellen et al., 2007; Marinis &

Chondrogianni, 2010; Paradis et al., 2008), and has important implications for the diagnosis of language impairment in bilingual populations and the evaluation of their language abilities.

Profile effects across language domains

An important novel finding that emerged from our study is that the L2 children showed a differential performance depending on the task and an asynchronous acquisition of the different structures, thus, exhibiting profile effects (Oller et al., 2007). The L2 children performed better on tasks targeting general comprehension of grammar and production of tense marking morphology, whereas they had a lower accuracy on the comprehension of single word vocabulary and (complex) morpho-syntax (articles, passives, wh-questions). Profile effects were also confirmed by the number of children who were able to reach age-appropriate norms, with only one third of the children achieving this for vocabulary and (complex) morpho-syntax, around half for tense morphology, and nearly two thirds for the comprehension of grammar.

The question that arises at this point is why the various language domains had differential accuracy rates in the same sample of children and how these profile effects can be accounted for. Previous studies attesting profile effects across (Oller & Eilers, 2002; Oller, et al., 2007) or within language domains (Gathercole, 2002a, 2002b) have drawn a clear distinction between vocabulary and grammar and have proposed different explanations depending on the language domain examined.

With regards to vocabulary, Oller et al. (2007) explain the bilingual profile effects on the basis of the “distributed characteristic”, namely vocabulary is partially distributed across languages with some concepts encoded in the L2 but not the L1 depending on the context. Apart from factors, such as the context of language use, another feature of vocabulary acquisition is that root lexical items are learned on an item-by-item basis, and rule generalisation for the formation of new lexical items is only possible at the level of derivational morphology. Additionally, lexical items are learned independently in each language regardless of whether or not the relevant concept has been acquired or is common between languages. Therefore, unless there is some degree of overlap between the two vocabularies in the form of cognates, bilingual bootstrapping effects between the two languages remain limited (Meara, 1992; Singleton, 1999). Conversely, in the case of the acquisition of morphology and syntax, distributed characteristics are limited and each morphological and syntactic system develops independently in each language (Gathercole, 2002b; Paradis & Genesee, 1996). Bilingual bootstrapping can be facilitatory in terms of transfer effects. Work on transfer in L2 acquisition has indicated that the presence of an underlying syntactic category in the L1 may facilitate acquisition in the L2. In the present study we could not address transfer effects because all children had Turkish as their L1. However, we follow Schwartz & Sprouse (1994) in arguing that transfer affects underlying syntactic categories and that transfer at the level of the morpho-phonological exponents of lexical items is excluded. This account argues for asynchronous acquisition between syntax and morphology and a dissociation between the two in L2 acquisition (Haznedar & Schwartz, 1997). In this account, the various domains of language need not develop

simultaneously, but differential performance is allowed both across and within domains, e.g. between third person *-s* and past tense (Haznedar, 2001). This asymmetry between tense morphemes was attested in the present study as in previous studies (Chondrogianni & Marinis, under review; Marinis & Chondrogianni, 2010). Additionally, within individual language systems, rules can be generalised at the level of morphology and syntax to generate an infinite number of utterances. This level of generalisation is not always possible at the level of vocabulary.

These profile effects and differences between vocabulary acquisition and morpho-syntax also raise theoretical questions regarding domain-specific (Guasti, 2002) vs. domain-general accounts of language acquisition (Marchman, 1997; see also Paradis, this volume). According to domain-general accounts, there should be a strong association between vocabulary and grammar with the latter developing in synchrony with the former. If, on the other hand, syntax and morphology can develop independently of each other and of vocabulary, then the tight correlation assumed by associative-learning accounts is not supported.

Profile effects and internal and external factors

Profile effects were further highlighted by the different age/time related (internal) and environmental (external) factors that affected children's performance on the various tasks. Performance on vocabulary (raw scores) was predicted at a ratio of almost 50% by a combination of LoE, AoO, and mSRP, suggesting that L2 children's vocabulary develops along with lengthier exposure. AoO also predicted L2 children's performance; children

with later AoO also exhibited higher vocabulary score, suggesting that cognitive maturity may indeed facilitate faster vocabulary acquisition (Golberg, et al., 2008). However, this does not necessarily mean that lengthier exposure and later AoO correlates with L2 children's ability to reach-age appropriate norms. Only one third of the children were able to reach age-appropriate norms. Rather, their inability to reach these norms was predicted by the mother's low proficiency in English. Mother's self-rated proficiency in this study correlated with their level of education in the L1. Since most of the mothers in our study had less than 12 years of formal education, the sample is largely representative of low SES. Thus, our findings are consistent with previous studies showing that children from families with low SES have lower vocabulary abilities compared to children with higher SES (Golberg, et al., 2008; Oller, et al., 2007). Golberg, et al (2008) suggested that SES makes a difference in children's lexical development in a language learned mainly outside the home because higher order verbal interaction associated with SES is fundamental and may be quantitatively and qualitatively reduced in a low SES environment. In the present study, low SES was coupled with the mother's low proficiency in the L2, suggesting reduced use of L2 vocabulary both qualitatively and quantitatively. In accordance with Golberg, et al. (2008) and contrary to Oller, et al. (2007), the use of English in the home did not predict children's performance on vocabulary neither their ability to reach age-appropriate norms primarily because most of the parents reported low use of English in the home (mean = 28%) and only few were fluent in English.

Environmental factors and mainly mother's self-rated proficiency and the use of English in the home were good predictors also of the children's development of the (complex)

syntactic phenomena from the DELV (*wh*-questions, overall score) and affected their inability to reach age-appropriate norms at a rate of 48.3%. This finding evokes again the relationship between mother's proficiency, educational level, and SES. As mentioned previously, in the present study low SES is intertwined with mother's low proficiency in the L2. It is very likely that the mothers in the sample do not use complex forms in the L2, e.g. multiple embedded *wh*-questions, due to their low L2 proficiency. Another factor may be the low SES of the mothers in the sample, as it has been suggested that higher educated mothers talk more to their children and tend to use longer and more complex sentences (Hoff, 2003). However, the extent to which SES affects the children's ability to acquire complex structures remains controversial (Noble, Norman, & Farah, 2005; Vasilyeva, Waterfall, & Huttenlocher, 2008) and in the particular sample it is difficult to disentangle SES from educational level and mothers' self-rated proficiency.

The development of general grammatical abilities as assessed by the TROG-2 was predicted by length of exposure and mother's proficiency. This suggests that L2 children's accuracy on this task improved with lengthier exposure and was related to higher mother's L2 proficiency. However, reaching age-appropriate norms was marginally predicted by mother's L2 proficiency. This suggests that, whereas L2 children's general grammatical abilities develop with time, what determines their ability to reach age-appropriate norms on grammar is their mother's proficiency in the L2. In our study, this external factor was better at predicting the amount of children who were able to reach age-appropriate norms rather than the amount of children who did not, and the rate of misclassifying an L2 child as not having reached age-appropriate norms was high.

Turning to the acquisition of tense marking morphology, L2 children showed an asymmetrical performance on the two tense marking morphemes with better performance on past tense compared to third person *-s*. This is consistent with previous findings on tense marking (Chondrogianni & Marinis, under review; Ionin & Wexler, 2002; Marinis & Chondrogianni, 2010; Paradis, et al., 2008). The children's performance on the third person singular *-s* was predicted by length of exposure, but this was not the case for past tense. This may well be because the sample of L2 children in this study performed high on past tense and there was not enough variation to give rise to any statistically significant relationship between the predictor and the outcome variables. Thus, the asymmetry between past tense and third person *-s* in our study is further highlighted by their differential susceptibility to age/time related variables.

A further novel finding of our study is the lack of any effect of environmental (external) variables, defined as mother's proficiency and English use in the home, on the acquisition of tense marking morphology. This was evidenced in the factors which affected L2 children's ability to reach age-appropriate norms for third person *-s* and past tense.⁴ Contrary to the findings regarding general grammatical abilities (TROG-2) and complex syntax (wh-questions from the DELV), the acquisition of tense marking morphology was not related to any external variables, such as the mother's proficiency or to the use of English in the home. This finding does not necessarily contradict previous results reported

⁴ Note also that L2 children's performance on the DELV component targeting article production was not predicted by mother's self-rated proficiency or use of English in the home, but rather by lengthier exposure, as shown by the multiple regression analysis. This result further corroborates the finding that the acquisition of morphology seems to be less susceptible to environmental factors.

by Jia & Fuse (2007) who found that the language environment across the years was a stronger predictor of performance variance than age of arrival for tense morphemes. Input effects defined as the accumulation of language experience across the years of exposure to English cannot generally be excluded from the acquisition of morphology. However, such effects were not uniformly attested in the present study for third person *-s* and past tense. This finding indicates that the acquisition of morphology seems to be less susceptible to input factors compared to the acquisition of vocabulary and complex syntax (wh-questions). This finding taken together with the differential performance on vocabulary and complex syntax, reveals an asynchrony in the development of these three language domains and can be regarded to support the domain specificity of the acquisition of tense marking (Haznedar & Schwartz, 1997; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005; Paradis, et al., 2008; Rice & Wexler, 1996; Rice, Wexler, & Hershberger, 1998).

References

- Bishop, D. V. M. (2003). *Test for Reception of Grammar, Version 2 (TROG-2)*. London: The Psychological Corporation.
- Bohman, T. M., Bedore, L. M., Peña, E. D., Mendez-Perez, A., & Gillam, R. B. (2010). What you hear and what you say: language performance in Spanish–English bilinguals. *International Journal of Bilingual Education and Bilingualism*, *13*, 325 - 344.
- Chondrogianni, V. (2008). The Acquisition of Determiners and Clitic Pronouns by Child and Adult L2 learners of Greek. Unpublished PhD dissertation, University of Cambridge.
- Chondrogianni, V., & Marinis, T. (under review). Production and processing asymmetries in the acquisition of tense morphology by sequential bilingual children *Bilingualism: Language and Cognition*.
- Cobo-Lewis, A. B., Pearson, B., Eilers, R. E., & Umbel, V. C. (2002). Effects of bilingualism and bilingual education on oral and written Spanish skills: A multifactor study of standardized test outcomes. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 64-97). Clevedon: Multilingual matters.
- Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale* (2nd ed.). Windsor: NFER-Nelson.
- Gathercole, V. C. M. (2002a). Grammatical gender in bilingual and monolingual children: a Spanish morphosyntactic distinctio. In L. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 207-219). Clevedon: Multilingual Matters.
- Gathercole, V. C. M. (2002b). Monolingual and bilingual acquisition: Learning different treatments of that-trace phenomenon in English and Spanish. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 43-63). Clevedon: Multilingual Matters.
- Gawlitzeck-Maiwald, I., & Tracy, R. (1996). Bilingual Bootstrapping. In N. Müller (Ed.), *Two languages. Studies in bilingual first and second language development* (pp. 901-926). Amsterdam: John Benjamins.
- Golberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied Psycholinguistics*, *29*, 41-65.
- Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.
- Gutierrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, *24*, 267-288.
- Gutierrez-Clellen, V. F., Simon-Cerejido, G., & Wagner, C. (2007). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics*, *29*, 3-19.
- Hakansson, G., & Nettelbladt, U. (1996). Similarities between SLI and L2 children: Evidence from the acquisition of Swedish word order. In I. Gilbert & Johnson (Eds.), *Children's language* (pp. 135-151). Mahwah, NJ: Erlbaum.

- Haznedar, B. (1997). L2 acquisition by a Turkish-speaking child: Evidence for L1 influence. In E. Hughes, M. Hughes & A. Greenhill (Eds.), *Proceedings of the Boston University Conference on Language Development 21* (pp. 245-256). Somerville, MA: Cascadilla Press.
- Haznedar, B. (2001). The acquisition of the IP system in child L2 English. *Studies in Second Language Acquisition*, 23, 1-39.
- Haznedar, B., & Schwartz, B. D. (1997). *Are there optional infinitives in child L2 acquisition?* Paper presented at the Boston University Conference on Language Development, Boston.
- Hoff, E. (2003). Causes and consequences of SES-related differences in parent-to-child speech. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting, and child development* (pp. 147–160). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ionin, T., & Wexler, K. (2002). Why is 'is' easier than '-s'? acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research*, 18, 95-136.
- Jacobson, P., & Livert, D. (2010). English past tense use as a clinical marker in older bilingual children with language impairment. *Clinical Linguistics & Phonetics*, 24, 101-121.
- Jacobson, P., & Schwartz, R. (2005). Elicited production of English past tense by bilingual children with language impairment. *American Journal of Speech-Language Pathology*, 4, 313-323.
- Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*, 24, 131–161.
- Jia, G., & Fuse, A. (2007). Acquisition of English Grammatical Morphology by Native Mandarin-Speaking Children and Adolescents: Age-Related Differences. *Journal of Speech, Language and Hearing Research*, 50, 1280-1299.
- Marchman, V. A. (1997). Models of language development: An “emergentist” perspective. *Mental Retardation & Developmental Disabilities Research Reviews*, 3, 293 – 299.
- Marinis, T., & Chondrogianni, V. (2010). Production of tense marking in successive bilingual children: When do they converge with their monolingual peers? *International Journal of Speech-Language Pathology*, 12, 19-28.
- Marinis, T., & Chondrogianni, V. (2011). Comprehension of reflexives and pronouns in sequential bilingual children: Do they pattern similarly to L1 children, L2 adults, or children with specific language impairment? *Journal of Neurolinguistics*, 24, 202-212.
- Meara, P. M. (1992). Vocabulary in a second language. *Reading in a Second Language*, 9, 761–837.
- Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science*, 8, 74-87.
- Oller, D. K., & Eilers, R. E. (2002). *Language and Literacy in Bilingual Children*. London: Multilingual Matters.

- Oller, D. K., Pearson, B., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28, 191-230.
- Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with Specific Language Impairment. *Language, Speech and Hearing Services in the Schools*, 36, 172-187.
- Paradis, J. (2008). Tense as a clinical marker in English L2 acquisition with language delay/impairment. In E. Gavrusseva & B. Haznedar (Eds.), *Current trends in child second language acquisition: a generative perspective* (pp. 337-356). Amsterdam: John Benjamins.
- Paradis, J., Crago, M., Genesee, F., & Rice, M. (2003). Bilingual children with specific language impairment: How do they compare with their monolingual peers? *Journal of Speech Language and Research*, 46, 1-15.
- Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*, 18, 1-25.
- Paradis, J., Rice, M. L., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics*, 29, 689-722.
- Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An Introduction to Logistic Regression Analysis and Reporting. *Journal of Educational Research*, 96, 3-14.
- Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of Specific Language Impairment in English-speaking children. *Journal of Speech and Hearing Research*, 39, 1239-1257.
- Rice, M. L., & Wexler, K. (2001). *Rice/Wexler Test of Early Grammatical Impairment*. New York: The Psychological Corporation, Harcourt Assessment Company.
- Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech and Hearing Research*, 38, 850 - 863.
- Schwartz, B. D. (2003). Child L2 acquisition: Paving the way. In B. Beachley, A. Brown & C. F. (Eds.), *Proceedings of the 27th Boston University Conference on Language Development* (pp. 26-50). Somerville, MA: Cascadilla Press.
- Schwartz, B. D., & Sprouse, R. (1994). Word order and Nominative Case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German Interlanguage. In T. Hoekstra & B. D. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar* (pp. 317-368). Amsterdam: John Benjamins.
- Seymour, H., Roeper, T., de Villiers, J., & de Villiers, P. (2005). *Diagnostic Evaluation of Language Variation (DELV™)—Norm-Referenced*. San Antonio: Harcourt Assessment.
- Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.
- Unsworth, S. (2005). *Comparing Child L1, Child L2 and Adult L2*. Unpublished PhD dissertation, Utrecht University, Utrecht.
- Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: commonalities and differences across children. *Developmental Science*, 11, 84-97.

- Verhoeven, L. T. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading, 4*, 313–330.
- Whong-Barr, M., & Schwartz, B. D. (2002). Morphological and syntactic transfer in child L2 acquisition of the English dative alternation. *Studies in Second Language Acquisition, 24*, 579-616.

Table 1. Scores on standardized assessments

		L1 children (N=33)			L2 children (N= 43)		
		Mean	SD	Range	Mean	SD	Range
TROG	Raw score	13.4	(2.0)	10-18	11.4	(3.8)	1-18
	Standard score	100	(9.3)	85-120	90.3	(15.9)	55-113
	z-score	0	(0.6)	-1.0–1.3	-.7	(1.1)	-3-.90
BPVS	Raw score	75.6	(11.2)	46-91	53.9	(11.7)	31-83
	Standard score	101	(7.6)	87-115	81	(7.4)	61-97
	z-score	0	(0.5)	-0.9–1.0	-1.3	(0.5)	-2.60--.20
TEGI	Third person(%)	92.3	(16.7)	20-100	78.5	(25.6)	0-100
	Past tense(%)	96	(6.4)	77.8-100	86.8	(22.4)	0-100
	CGS(%)	94	(9.9)	60-100	83	(21.4)	5.5-100
DELV	Articles(%)	93	(9.5)	75-100	51.2	(25.3)	0- 87.5
	Passives(%)	76.6	(19.4)	0-100	51.4	(22.8)	0-90
	Wh-(%)	91	(13.5)	50-100	73.7	(24.9)	0-100
	Overall	86.8	(10.4)	65-100	58.8	(19.2)	0-89.2
	z-score	0	(0.9)	-1.3–2.3	-1.5	(0.6)	-3.0--.30

CGS:Composite Grammar Score

Table 2. Correlations between age/time related (internal) and environmental (external) variables and language tasks for L2 children

Participant variables		TROG-2		BPVSII		TEGI			DELV			
		raw	stand	raw	stand	-s	-ed	overall	articles	Wh-	passives	overall
Internal factors	AoO	-.216	-.151	-.194	-.229	-.144	-.282	-.261	-.161	-.346*	-.379*	-.365*
	LoE	.328*	-.051	.598**	-.082	.445**	.337*	.448**	.431**	.601**	.513**	.704**
	AToT	.326*	-.091	.605**	-.143	.439**	.173	.360*	.359*	.517**	.348*	.601**
External factors	mLOE	.118	.039	.228	.121	.087	.270	.115	.062	.224	.183	.181
	mSRP	.325*	.289	.243	.343*	.128	.161	.133	.134	.365*	.367*	.331*
	fLOE	.107	.002	.136	-.080	.045	.147	.058	.233	.186	-.012	.185
	fSRP	.015	-.069	.050	-.046	.118	-.014	.038	.052	-.238	-.183	-.148
	ENGH	.184	.083	.256	.068	.061	.344*	.155	.417**	.345*	.128	.422**

*p<.05, **p<.01

Table 3. Correlation matrix for age/time related variables

	LoE	AoO
AoO	-.559**	
AToT	.863***	-.130

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

Table 4. Correlation matrix for environmental variables

	mLoE	mSRP	fLoE	fSRP	sSRP
mSRP	.595**				
fLoE	.537**	.295*			
fSRP	.165	-.104	.312		
ENGH	.393**	.301*	.265	.092	.201

**p<.01, *p<.05

Table 5. Multiple regression analysis for the TROG-2 and the BPVSII

	TROG-2			BPVSII						
	raw			raw			standard			
	B	SE	β	B	SE	β	B	SE	β	
LoE	1.20	.52	.38*	7.12	1.34	.70***				
AoO				3.48	1.42	.298*				
mSRP	1.38	.58	.35*	8.67	3.02	.391**	2.22	1.10	.30*	

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

Table 6. Multiple regression analysis for the TEGI

	TEGI								
	-s			-ed			CGS		
	B	SE	β	B	SE	β	B	SE	β
LoE	12.20	3.35	.59***				8.78	2.93	.51**
AoO	13.25	7.10	.35*				9.73	6.20	.27
ENGH				4.11	.20	.30*			

**p<.01, *p<.05

Table 7. Multiple regression analysis for the DELV

	DELV											
	Articles			Wh-questions			Passives			Overall		
	B	SE	β	B	SE	β	B	SE	β	B	SE	β
LoE	6.04	3.10	.295*	12.62	2.48	.62***	8.10	3.00	.37*	10.02	1.88	.58***
AoO							-13	5.23	-.36*			
mSRP				6.65	2.83	.28*				4.89	2.16	.27*

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

Table 8. Logistic regression for the TROG-2 and the BPVSII

Predictor	TROG-2						BPVSII					
	B	S.E.	Wald	df	<i>p</i>	Exp(B)	B	S.E.	Wald	df	<i>p</i>	Exp(B)
Constant	-.97	1.34	.54	1	.46	.37	-1.99	.78	6.59	1	.01	.19
mSRP	.93	.49	3.58	1	.03	2.53	.65	.35	3.37	1.06	.06	1.92
Test				χ^2	df	<i>P</i>				χ^2	df	<i>p</i>
Overall model evaluation												
Likelihood ratio test			7.287	3	0.015		6.325	3	.054			
Goodness-of-fit test												
Hosmer & Lemeshow			11.166	7	.606		4.213	8	.837			
TROG:Cox and Snell R^2 =.150;Nagelkerke R^2 =.218, BPVS:Cox and Snell R^2 =.134;Nagelkerke R^2 =.188												

Table 9. Logistic regression for the TEGI (third person) and the DELV

Predictor	TEGI -s						DELV					
	B	S.E.	Wald	df	<i>p</i>	Exp(B)	B	S.E.	Wald	df	<i>P</i>	Exp(B)
Constant	-3.29	1.36	5.86	1	.02	.04	2.25	3.44	.43	1	.51	9.51
LoE	.73	.41	5.74	1	.02	2.08						
mSRP							.96	.42	5.14	1	.02	2.61
Test			χ^2	df	<i>p</i>				χ^2	df	<i>P</i>	
Overall model evaluation												
Likelihood ratio test			7.08	1	.01				11.67	2	.01	
Goodness-of-fit test												
Hosmer & Lemeshow			7.20	7	.41				4.54	7	.72	
TEGI: Cox and Snell R^2 = .146; Nagelkerke R^2 = .194, DELV: Cox and Snell R^2 = .243; Nagelkerke R^2 = .337												

Table 10. Observed and predicted frequencies for meeting age appropriate norms on the TROG-2, BPVSII, TEGI and DELV by Logistic Regression with Cut-off of 0.50

Observed	TROG-2			BPVSII			TEGI			DELV		
	Predicted		% correct	Predicted		% correct	Predicted		% correct	Predicted		% correct
	Yes	No		Yes	No		Yes	No		Yes	No	
Yes	31	2	93.9%	2	12	14.3%	13	8	61.9%	7	7	50%
No	10	2	16.7%	1	29	96.7%	6	18	75%	2	26	92.9%
Overall % correct			73.3%			70.5%			68.9%			78.6%

TROG-2: False positive=10/(10+31)%=24.4%, False negative=2/(2+2)=50%, BPVSII: False positive=1/(1+2)%=33.3%, False negative=12/(12+29)%=29.3%, THIRD PERSON: false positive=6/(6+13)%=31.6%, False negative=8/(8+18)%=30.8%, DELV: False positive=2/(2+7)%=22.2, False negative=7/(7+26)%=21.2%.

Figure 1:Percentage of L2 children reaching age-appropriate norms

