Production of tense marking in successive bilingual children: when do they converge with their monolingual peers?

Article

Accepted Version

In: ***International Journal of Speech-Language Pathology 12 (1), 19-28***

Production of tense marking in successive bilingual children: when do they converge with their monolingual peers?

Theodoros Marinis & Vasiliki Chondrogianni

School of Psychology & Clinical Language Sciences

University of Reading

Running Head: Production of tense marking in L2 children

Address for correspondence:
Theodoros Marinis
School of Psychology & Clinical Language Sciences
University of Reading
Reading RG6 6AL
UK
Tel. +44-118-378 7465
Fax +44-118-3784693
e-mail: t.marinis@reading.ac.uk

Production of tense marking in successive bilingual children: when do they converge with their monolingual peers?

ABSTRACT

Children with English as a second language (L2) with exposure of 18 months or less exhibit similar difficulties to children with Specific Language Impairment in tense marking, a marker of language impairment for English. This paper examines whether L2 children with longer exposure converge with their monolingual peers in the production of tense marking.

38 Turkish-English L2 children with a mean age of 7;8 and 33 monolingual age-matched controls completed the screening test of the Test of Early Grammatical Impairment (TEGI). The L2 children as a group were as accurate as the controls in the production of *-ed*, but performed significantly lower than the controls in the production of third person *–s*. Age and YoE affected the children's performance. The highest age-expected performance on the TEGI was attested in eight and nine year-old children who had 4-6 YoE. L1 and L2 children performed better in regular compared to irregular verbs, but L2 children overregularized more than L1 children and were less sensitive to the phonological properties of verbs. The results show that tense marking and the screening test of the TEGI may be promising for differential diagnosis in eight and nine year-old L2 children with at least four YoE.

INTRODUCTION

Research in the language abilities of successive bilingual children has shown that at an early stage of development when they have very little exposure to their second language (L2), L2 children have difficulties with the inflectional morphology of their L2 and show a similar pattern of performance to children with Specific Language Impairment (SLI) (Hakansson & Nettelbladt, 1996; Paradis, 2005; Paradis & Crago, 2000). This similarity causes difficulties in discriminating between successive bilingual children with and without language impairment (Crutchley, Conti-Ramsden, & Botting, 1997). This may lead to non-impaired L2 children being inappropriately diagnosed as language impaired and L2 children with SLI not receiving appropriate services. Given the lack of norms for L2 children and the very limited number of assessment resources for the L1 of the successive bilingual children, it is difficult to distinguish between children whose lower language abilities are due to a late onset or to a combination of a late onset and language impairment (Gutierrez Clellen, 2000). The present paper contributes to our understanding of the profile of L2 children by investigating tense marking in non-impaired Turkish-English L2 children who have a considerable amount of exposure to English.

English-speaking children with SLI have severe problems with tense marking morphemes, i.e. third person singular *–s* (The boy kick*s* the ball), past tense *–ed* (The boy kick*ed* the ball), auxiliary BE (The boy *is* kicking the ball), copula BE (The boy *is* tall), and auxiliary DO (*Does* the boy like football?) and tense marking has been proposed as a clinical marker for SLI (Conti-Ramsden, 2003; Rice & Wexler, 1996). However, difficulties with tense marking are not confined to the group of children with SLI, but are attested also in children with English as a L2 (Gavruseva & Lardiere, 1996; Haznedar, 2001; Ionin & Wexler, 2002; Lakshmanan, 1994; Paradis, 2005, 2008).

Non-impaired children acquiring English as a L2 have greater difficulties with tense compared to non-tense morphemes (Jia & Fuse, 2007; Paradis, 2005) and they show an uneven profile in the production of tense morphemes; third person singular *–s* and past tense forms seem to be more affected than auxiliary and copula BE and children have been reported to use BE as an all-purpose finiteness marker (Ionin & Wexler, 2002). Paradis (2005) using the Test of Early Grammatical Impairment (TEGI) (Rice & Wexler, 2001) showed that 4-to-7 year-old L2 children had a lower accuracy in third person singular *–s* (16.6%), regular past tense *–ed* (22.6%), and irregular past tense forms (12.7%) compared to BE (60.2%), and Paradis et al. (2008) showed that the same L2 children were less accurate than L1 children with SLI matched on age and MLU and L1 younger controls matched on MLU in the production of third person singular *–s* and past tense. In addition, these two studies by Paradis and colleagues revealed that the amount of exposure to English did not correlate with the use of tense morphemes.

Most studies investigating tense marking in L2 children have focused on L2 children with very limited exposure to English (Paradis, 2005; Paradis et al., 2008: 2-18 months; Ionin & Wexler, 2002: less than 24 months to 3 years). A recent study by Paradis (2008) addresses the issue of exposure by reporting how nine Mandarin and Cantonese non-impaired L2 children and two L2 children with SLI perform on the TEGI after one, two and three years of exposure to English. The asymmetry between third person singular and past tense vs. BE attested in the previous studies disappears in the non-impaired L2 children after two years of exposure, but even after three years of exposure their accuracy in tense marking was still 1.5 SD below the monolingual norms. It is, thus, unclear how long it takes for L2 children to converge with L1 children and meet the criterion scores on the TEGI.

Finally, research on past tense formation on irregular verbs in L1 children has revealed that accuracy and error patterns are affected by frequency and phonological properties of the

verbs (Bybee & Slobin, 1982; Marchman, 1997; Matthews & Theakston, 2006). High frequency verbs show higher accuracy and less overregularizations than low frequency verbs. To address the effect of phonological properties on the response pattern of irregular verbs, Bybee & Slobin classified verbs into eight categories based on: 1) whether or not there is final *t/d* in the past tense form and in the stem, and 2) whether or not there is an internal vowel change in the past tense form, as shown in table 1.

INSERT TABLE 1 HERE

In terms of their phonological properties, the above studies have revealed that in school-aged children, verbs ending in an alveolar /t/, /d/ in the stem and in the past tense, (Class V, *write/wrote*) elicit fewer overregularizations and more errors of omission or internal vowel change than all other categories involving internal vowel change (III-VIII). Verbs ending in an alveolar /t/, /d/ in the past tense but not in their stem form (Class III, IV: *make/made, catch/caught*) elicit more overregularizations than errors of omission and vowel change errors compared to the other categories involving internal vowel change, and verbs that do not end in an alveolar /t/, /d/ in their past tense but end in another consonant (Class VI, VII: *dig/dug, give/gave*) or vowel (Class VIII: *blow/blew*) elicit the largest number of overregularizations from all categories. To date, there are no data on the effect of frequency and phonological properties on L2 children's response pattern of irregular verbs. Thus, it is unclear whether or not L2 children are sensitive to input frequency and phonological properties of irregular verbs and whether their production of irregular verbs is affected by these factors. If the frequency and phonological properties of irregular verbs do affect the L2 children's production, this may have implications for the evaluation of results from existing language assessments that do not

control for those factors and also these factors should be taken into consideration in the development of new assessment material.

The present paper addresses the production of tense marking in non-impaired L2 children who have been exposed to English between 12 and 84 months compared to L1 age-matched controls addressing the following research questions:

1. Do L2 children perform within norms in assessments normed with L1 children?
2. Do L2 children converge with L1 children in the production of tense marking?
3. Do L2 children show a similar error pattern with irregular verbs as L1 children and are their responses affected by the phonological properties of verbs?

METHODOLOGY

Participants

Thirty eight typically developing (TD) successive bilingual Turkish-speaking children and thirty three TD monolingual English-speaking children participated in the study. The two groups were matched on age. 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 (SD: 8 months; range: 6;1 – 8;11) ($F(1, 68) = 2.806$, $p > 0.05$). Both groups of children attended schools whose percentage of free school meals was well above the national average indicating a low socio-economic status.

All L2 children were from the Turkish community in London and were recruited from schools with a high density of 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, the onset of systematic exposure to English should be 3 years or older, as indicated by nursery attendance, 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. This confirmed that none of the children had any history of speech and/or language delay or impairment, their parents were not concerned about their language development, and most children had an age of onset (AoO) around three years (mean: 3;4; SD: 9 months; range: 12 – 60 months). 31 children had an AoO between 2;6 and 3;6, one child had AoO 1;0, three children had AoO 4;0 and three children had AoO 5;0. The children's mean exposure to English was 4 years (SD: 16 months; range: 12 - 72 months).

Procedures

Children participated in a battery of standardised and non-standardised assessments and experimental tasks examining various linguistic phenomena. The present paper reports data from two background tasks assessing comprehension of grammar, i.e. TROG2 (Bishop, 2003), and single-word vocabulary, i.e. BPVSII (Dunn, Dunn, Whetton, & Burley, 1997), and data from the screening task from the Test of Early Grammatical Impairment (TEGI) (Rice & Wexler, 2001) that assesses tense marking. The TEGI has been standardized with 393 TD children and 444 children known to have SLI, between the ages of 4 to 9 years. The screening task from the TEGI elicits third person singular *–s*, regular and irregular past tense forms. For the elicitation of *-s* children are shown pictures of professionals engaging in various activities and are asked to produce a verb form following a probe such as 'Here is a teacher. Tell me what she does'. The expected answer should be something like 'A teacher teaches'. For the elicitation of past tense, children are shown two pictures, one with children engaged in an activity and one with the activity being completed. Past tense forms are prompted by the following lead-in sentences 'Here the boy is painting the fence. Now [pointing to the other picture] he is done. Tell me what he did'. The expected answer is 'The boy painted the fence'. The task includes ten trials for the production of *–s* and eighteen for the production of past tense, ten for regular and eight for irregular verbs.

Coding and scoring

Responses were transcribed during the testing session on the TEGI scoring form and were later scored following different scoring procedures depending on the research question. In order to have an overview regarding the L2 children's performance on tense morphemes and comparable scores to the L1 and SLI norms available in the TEGI manual, the children's responses were first scored according to the procedure described in the manual. According to this procedure, the scores on the individual elicitation probes are calculated in percentage correct. Responses that are irrelevant or do not attempt the target form (e.g. present continuous for third person *–s* or present tense for the past tense) are considered 'unscorable' and are excluded from the denominator for the score. For both probes, target-like responses are coded as correct and omissions of *–s* or *–ed* as incorrect. For irregular past tense forms, overregularizations are also coded as correctly inflected forms. This scoring procedure follows theoretical assumptions according to which overregularizations denote knowledge of tense inflection and past tense rule formation.

To compare the accuracy of regulars vs. irregular verbs, we followed a different scoring procedure for the irregulars to the one on the TEGI. Overregularizations were scored as incorrect together with uninflected forms. In addition, we scored suppletive forms of irregular verbs with non-target internal vowel change (ride→rid) also as incorrect. For this analysis, correct and erroneous (omissions, overregularizations, vowel change) responses were summed up and constituted the denominator for calculating proportion correct. For the error analysis of irregular verbs, the denominator was the sum of the types of errors produced (uninflected, overregularized and vowel change).

RESULTS

To examine how the L2 children perform in assessments standardized with monolingual children, we first analyzed the results from TROG2 and BPVSII, as shown in table 2.

INSERT TABLE 2 HERE

To examine differences between the two groups and the two tasks, we entered the standard scores into a repeated-measures ANOVA with Group (L1, L2) as a between-subjects and Task (TROG2, BPVS) as a within subjects factor. This showed a main effect of Group ($F(1, 68) = 50.342$, $p < 0.001$, $\eta^2 = 0.425$), a main effect of Task ($F(1, 68) = 8.28$, $p < 0.01$, $\eta^2 = 0.116$), and an interaction between Group and Task ($F(1, 68) = 6.888$, $p = 0.01$, $\eta^2 = 0.083$). One-way ANOVAs showed that the L2 children were less accurate than the L1 children on both tasks (TROG2: $F(1 ,68) = 14.206$, $p < 0.001$, $\eta^2 = 0.131$); BPVSII: $F(1, 68) = 99.106$, $p < 0.001$, $\eta^2 = 0.131$); pairwise comparisons using Bonferroni correction showed that L1 children were equally good on TROG2 and BPVSII, but L2 children were significantly better on TROG2 compared to BPVSII ($p = 0.001$). Overall, 17 of the 38 children were at or below 1 SD from the mean in TROG2 and 26 of the 38 in BPVSII. When we take YoE into consideration, more than half (n=57%) of children with 1 to 3 YoE to English were at or below 1SD from the mean in TROG2 and more than two thirds (n=71%) were at or below 1SD in BPVSII. This discrepancy between TROG2 and BPVSII becomes even more pronounced in children that have 4 to 6 YoE to English; around one third of these children (n=36%) were at or below 1SD from the mean in TROG2, but more than two thirds (n=73%) were at or below 1SD in BPVSII.

To address whether the L2 children in our study converge with L1 children in the production of tense marking, we first calculated the children's accuracy for third person singular and past tense according to the scoring procedure of the TEGI, as shown in figure 1.

INSERT FIGURE 1 HERE

To examine differences between the two groups and the two morphemes, we entered the data into a repeated-measures ANOVA with Group (L1, L2) as a between-subjects and Morpheme Type (third person singular, past tense) as a within-subjects factor. This revealed a main effect of Group ($F (1, 68) = 6.409$, $p = 0.01$, $\eta^2 = 0.093$), a main effect of Morpheme Type ($F (1, 68) = 1.186$, $p < 0.01$, $\eta^2 = 0.105$), and a significant interaction between Group and Morpheme Type ($F (1, 68) = 5.259$, $p < 0.05$, $\eta^2 = 0.074$). One-way ANOVAs showed that in the production of third person *–s*, L2 children had a significantly lower accuracy (78.2%), than L1 children (92.3%) ($F (1, 68) = 7.096$, $p = 0.010$, $\eta^2 = 0.094$), but there was no significant difference between the two groups in the production of past tense (L2: 90.1%; L1: 93.5%; ($F (1, 68) = 2.363$, $p > 0.1$, $\eta^2 = 0.034$). Pairwise comparisons between the two morphemes using Bonferroni correction for each group separately showed that L1 children performed equally well in the two morphemes ($p > 0.1$), but L2 children performed significantly better in past tense compared to third person singular ($F (1, 37 = 12.024$, $p = 0.001$, $\eta^2 = 0. 252$).

The TEGI scores are criterion referenced, i.e. the scores for each morpheme are compared to a cut off score, above which performance can be considered in the range of TD children. Criterion scores vary by age according to six-month intervals up to the age of 6;11 and yearly intervals up to the age of 8;11. Figure 2 presents the percentage of L2 children that performed at and above the criterion score for their age in the TEGI screening test.

INSERT FIGURE 2 HERE

Between the ages of 6;0 and 6;5 none of the L2 children (n=5) were at or above the criterion score for either morpheme. Between 6;6 and 6;11 all children were also below the criterion score for third person singular (n=4), but three out of four children were at or above the criterion score for past tense. At the age of 7, more than half of the children (n=9/15) were above the criterion score for the production of the two morphemes. At the age of 8, five out of seven children were above the criterion score for third person singular and four out of seven for past tense. Finally, at the age of 9, six out of seven children were above the criterion score for both morphemes.

To investigate how YoE affect L2 children reaching criterion scores for each morpheme, we calculated the percentage of L2 children who performed at or above criterion score for their age, as shown in figure 3 below.

INSERT FIGURE 3 HERE

The asymmetry between past tense and third person singular is attested also when we group children on the basis of YoE to English. Six out of fourteen children met criterion for past tense before four YoE to English, but only one out of fourteen met criterion for third person singular with the same YoE.[1] From four years of exposure onwards this discrepancy disappears and more than half of the children met criterion for both morphemes. However,

[1] This child had age of onset 5;0.

even after 6 YoE three out of nine children did not meet criterion score for third person singular and two out of nine children did not meet criterion for past tense.

To investigate what combination of age and YoE produces age-expected performance on the TEGI, we focused on the eight and nine year olds because these groups had the highest proportion of children with age-expected performance and then we looked at their YoE. All eight and nine year-old children had 4 or more YoE to English. In the group of eight year olds, four children had 4 YoE and three had 5 YoE. Three out of the four children with 4 YoE met criterion for third person singular and two met criterion for past tense. Two out of the three children with 5 YoE met criterion for third person singular and the same two children met criterion also for past tense. In the group of the nine year olds, one had 4 YoE and six had 6 YoE. The child with 4 YoE met criterion for both third person singular and past tense. Five out of the six children with 6 YoE met criterion for third person singular and the same children met criterion also for past tense.

To investigate differences between regular and irregular past tense forms in the two groups, we conducted a repeated-measures ANOVA with Group (L1, L2) as between-subjects and Past Tense Type (regular, irregular) as a within-subjects factor. This showed a main effect of Group ($F (1, 68) = 56.965$, $p < 0.001$, $\eta^2 = 0.456$), a main effect of Past Tense Type ($F (1, 68) = 587.461$, $p < 0.001$, $\eta^2 = 0.896$), and a significant interaction between Group and Past Tense Type ($F (1, 68) = 28.140$, $p < 0.001$, $\eta^2 = 0.293$). One-way ANOVAs revealed that both groups had a higher accuracy in regulars compared to irregulars (L2: ($F (1, 37) = 372.083$, $p < 0.001$, $\eta^2 = 0.908$; L1: ($F (1, 32) = 254.755$, $p < 0.001$, $\eta^2 = 0.888$). The L2 children were significantly less accurate than the L1 children in both regulars ($F (1, 68) = 7.408$, $p < 0.01$, $\eta^2 = 0.098$) and irregulars ($F (1, 68) = 50.872$, $p < 0.001$, $\eta^2 = 0.428$). The difference between the two groups in the irregulars was larger than in the regulars, as shown in figure 4 below, which seems to have caused the Group by Past Tense Type Interaction.

INSERT FIGURE 4 HERE

To investigate differences between the two groups on the error types of irregular verbs, we calculated the percentage of each error type (uninflected, overregularizations, vowel change) for each group separately as a proportion of the total number of errors, as shown in figure 5. A repeated-measures ANOVA with Group (L1, L2) as a between-subjects and Error type (uninflected, overregularization, vowel change) as a within-subjects factor showed a main effect of Error Type ($F (1.2, 82.1) = 183.844$, $p < 0.001$, $\eta^2 = 0.890$) and a significant interaction between Error Type and Group ($F (1.2 ,82.1) = 5.108$, $p < 0.05$, $\eta^2 = 0.188$). One-way ANOVAs showed that the two groups did not differ on the proportion of uninflected forms they produced ($F (1, 68) = 1.022$, $p > 0.1$, $\eta^2 = 0.015$), but the L2 children made significantly more overregularizations than the L1 children ($F (1, 68) = 5.836$, $p < 0.05$, $\eta^2 = 0.079$), and only the L1 children made errors of vowel change ($F (1, 68) = 15.647$, $p < 0.001$, $\eta^2 = 0.187$). Finally, pairwise comparisons between the three error types for each group using Bonferroni correction showed that both groups made more errors of overregularizations compared with the use of uninflected stems (L1: $p < 0.001$; L2: $p < 0.001$) and errors of vowel change (L1: $p < 0.001$; L2: $p < 0.001$), and L1 children used more uninflected stems than errors of vowel change (L1: $p < 0.01$).

INSERT FIGURE 5 HERE

To investigate effects of phonological properties of the irregular verbs, we classified the irregular verbs from the TEGI (*make, catch, write, eat, ride, dig, give, blow*) according to Bybee & Slobin (1982), as shown in table 3.

INSERT TABLE 3 HERE

All classes apart from class V are represented by only one verb, therefore, it is difficult to ascertain whether the children's responses for each verb are due to the properties of the verb class or the specific verb used on the TEGI. For this reason, the following analyses are tentative and cannot lead to generalizations about verb categories.

To control for a possible confounding factor of frequency, we obtained the verbs' log frequencies from the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993)[2] and we analyzed the two most frequent verbs (make, give) separately from the rest. Tables 4 and 5 show the response types of the L1 and L2 children for each verb separately.

INSERT TABLES 4 AND 5 HERE

For the high frequency verbs *make* and *give*, L2 children were accurate half of the time and the majority of their errors were overregularizations. In contrast, L1 children showed an accuracy rate of over 80% and equal proportions of overregularizations and uninflected forms.

To explore a possible confounding factor of frequency in the mid/low frequency verbs, we conducted a preliminary ANCOVA with Group as a between-subjects factor, Response Type as within-subjects factor, and Frequency as a covariate. This showed no main effect of Frequency ($p > 0.1$) and no interaction of Frequency with Group ($p > 0.1$) and with Response Type ($p > 0.1$), which suggests no frequency effect for this set of verbs.

[2] It should be noted that although the CELEX database includes spoken language, it consists predominantly of written texts. Therefore, we fully agree with one of the reviewers that the frequencies should be treated with caution in terms of how representative they might be of what children of this age are exposed to.

To investigate possible differences between the verbs in the response type of the children and to explore whether individual verbs showed a different response type per group, we conducted a repeated-measures ANOVA with Group as between-subjects factor, Verb and Response Type as within-subjects factors. This showed a main effect of Group ($F (1, 42) = 22.902$, $p < 0.001$, $\eta^2 = 0.026$), a main effect of Response Type ($F (1.921, 80.697) = 64.817$, $p < 0.001$, $\eta^2 = 0.939$), a significant interaction between Group and Response Type ($F (3, 126) = 21.832$, $p < 0.001$, $\eta^2 = 0.614$), Response Type and Verb ($F (15, 630) = 9.028$, $p < 0.001$, $\eta^2 = 0.903$), and a 3-way interaction between Group, Response Type and Verb ($F (15, 630) = 4.580$, $p < 0.001$, $\eta^2 = 0.730$). To unpack this 3-way interaction, we conducted separate ANOVAs for each group with the within-factors Verb and Response Type. For the L2 group, this showed only a main effect of Response Type ($F (1.74, 34.92) = 69.320$, $p < 0.001$, $\eta^2 = 0.954$), but no interaction between Response Type and Verb indicating that the L2 children showed the same pattern of response for all verbs. However, the analysis of the L1 group showed a main effect of Response Type ($F (1.39, 23.71) = 21.663$, $p < 0.001$, $\eta^2 = 0.967$) and an interaction between Response Type and Verb ($F (15, 255) = 10.670$, $p < 0.001$, $\eta^2 = 0.433$). This indicates that the L1 children followed a different response for some of the verbs.

To examine the source of these differences for the L1 children, we conducted one repeated-measures ANOVA for each response type with Verb as the within-factor followed by post-hoc pairwise comparisons using Bonferroni correction. In terms of correct responses, there was a significant effect of Verb ($F (5, 210) = 9.736$, $p < 0.001$, $\eta^2 = 0.806$) because the verb *eat* showed the highest accuracy compared to all other verbs ($p < 0.05$). In terms of overregularizations, there was an effect of Verb ($F (5, 95) = 12.106$, $p < 0.001$, $\eta^2 = 0.457$) because the verb *blow* showed more overregularizations than the verb *eat* ($p < 0.05$) and the verbs *catch* and *dig* showed more overregularizations than the verbs *write* and *eat* ($p < 0.01$). In terms of the uninflected responses, there was no difference between the verbs, but the

analysis of vowel change errors showed a main effect of Verb ($F (5, 95) = 13.157$, $p < 0.001$, $\eta^2 = 0.437$) because only the verbs *write* and *ride* showed errors of vowel change ($p < 0.05$).

To examine possible relationships between the accuracy in the two grammatical morphemes, and the raw scores of TROG2 and BPVSII in the L2 children we conducted non-parametric correlations (Spearman's rho), as shown in table 6.

INSERT TABLE 6 HERE

These showed moderate positive correlations between the children's accuracy in third person singular, TROG2, BPVSII, and regular past tense, and a weak positive correlation between third person singular and past tense. There was a high positive correlation between the children's accuracy in past tense and accuracy in regular past tense forms, a moderate positive correlation between past tense and TROG2, and a weak positive correlation between past tense and BPVSII. There was a moderate positive correlation between the children's accuracy in regular and irregular past tense and BPVSII, and a weak positive correlation between the children's accuracy in irregular past tense and TROG2. Finally, there was a moderate positive correlation between the children's scores in TROG2 and BPVSII.

To examine possible relationships between YoE, the accuracy in the two grammatical morphemes and the raw and standard scores of TROG2 and BPVSII we conducted non-parametric correlations (Spearman's rho). These showed a strong positive correlation between YoE and the raw score of BPVSII ($r (36) = 0.707$, $p < 0.01$), moderate positive correlations between YoE and third person singular ($r (36) = 0.483$, $p < 0.01$) and YoE and the raw score of TROG2 ($r (36) = 0.444$, $p < 0.01$), and a weak positive correlation between YoE and past tense ($r (36) = 0.343$, $p < 0.05$). No correlations were found between YoE and the standard scores on TROG2 and BPVSII.

DISCUSSION

This study investigated the production of third person singular *–s* and past tense in thirty eight 6;2-9;8 year-old Turkish-English children whose mean onset was 3;4 years and had an average of 4 years of exposure to English, compared to thirty three L1 children matched on age. The aim of the study was to investigate: 1) how L2 children perform in assessments normed for monolingual children, 2) whether they converge with monolingual children in tense marking, and 3) whether they show a similar error pattern to L1 children in the production of irregular verbs and whether their errors are affected by the phonological properties and frequency of the verbs. To date, this is the only child L2 study using such a large sample size in a relatively narrow age range of children with mean age of exposure of more than 18 months, and investigating not only the L2 children's performance on third person singular *–s*, and past tense, but also addressing their error patterns in irregular verbs.

The results showed that the L2 children were less accurate than the L1 children in two assessments widely used in Speech and Language Therapy clinics tapping the comprehension of grammar (TROG2) and single-word vocabulary (BPVSII), and overall, L2 children had a higher score in the comprehension of grammar than in the comprehension of vocabulary. This is partially in line with previous studies showing that L2 children are less accurate than L1 children in standardized assessments (Crutchley, Conti-Ramsden, & Botting, 1997; Mahon & Crutchley, 2006), but also shows that L2 children converge to the norms of TROG2 earlier than on the norms of BPVSII. Importantly, as a group, the scores from TROG2 are within 1SD of the population mean, whereas the scores from BPVSII are below 1SD of the population mean. In addition, although the standard scores on both TROG2 and BPVSII did not correlate with YoE, we found an effect of YoE for TROG2 when we used four YoE as a cut-off point. With 1 to 3 YoE, 57% of the L2 children were at or below 1SD from the mean,

but with 4 to 6 YoE, the proportion of children at or below 1SD was only 36%. In contrast, the proportion of children at or below 1SD in BPVSII did not change with YoE (1 to 3 YoE: 71%, 4 to 6 YoE: 73%). This shows that L2 children catch up in the comprehension of grammar at a faster pace than in the comprehension of single-word vocabulary. Therefore, interpreting the performance of L2 children on TROG2 using monolingual norms is less likely to result in misidentification than doing so with BPVSII norms. Finally, the significant correlations between YoE and the raw scores of TROG2 and BPVSII show that the L2 children's vocabulary and grammar do develop with extended exposure, although in an asymmetric manner and without reaching overall age-appropriate norms, as the lack of correlations of YoE with the standard scores indicate.

In the production of tense marking using the scoring of the TEGI, when we looked at the L2 children as a group, they were as accurate as the L1 children in past tense, but they were significantly less accurate in third person singular *–s*. However, this does not mean that all L2 children performed similarly to L1 children. Age and YoE seemed to affect their performance. When we analysed the data based on the age of the children, none of the 6 year-old children met the criterion score for *-s* or past tense, but the majority of 6;5 -to-9-year-old children met the criterion score for past tense, and from the age of 7, the majority of children met criterion also for third person singular *–s*. When we analysed the data based on the YoE, the discrepancy between *–s* and *–ed* was present in children with less than 4 YoE, but disappeared in children with more than 4 YoE to English. This is a novel finding and shows that L2 children can meet the criterion score for past tense earlier than for third person singular in terms of YoE. YoE correlated with the children's performance on both third person singular and past tense and most children in our study met criterion for both after 4 YoE to English. However, it should be noted that even after 6 YoE, 22% of the L2 children did not meet criterion for past tense and 33% did not meet criterion for third person singular.

The highest age-expected performance on the TEGI was attested in eight and nine year-old children who had 4 to 6 YoE to English.

The asymmetry between third person singular *–s* and past tense could be related to the properties of the morpho-phonological paradigm of agreement and tense in English. Whereas the English past inflection has a consistent morpho-phonological form, the regular suffix *–ed,* which does not vary depending on the form of the subject, the agreement paradigm is morpho-phonologically poor with only third person singular*–s* being overtly marked. Whether *–s* is also morpho-syntactically marked, with the zero Ø forms on all other persons and numbers being the unmarked ones and indicating lack of agreement, is much less clear. It may well be the case that third person *–s* indicates non-agreement with a personal subject, whereas the zero suffix signals the opposite (Blevins, 2006). In this respect we could argue that the child L2 learners find the impoverished English agreement paradigm problematic, but not that they have problems establishing an agreement relation between the subject and the verb.

Analyses for regular and irregular past tense separately revealed that both groups were more accurate in regular compared to irregular verbs, and this discrepancy was larger in L2 children. Analyses of error types for irregular verbs showed that both groups made more errors of overregularization than of uninflected forms, but the rate of overregularizations was higher in L2 compared to L1 children, and L1 children made errors of vowel change that were not attested in L2 children. Both groups showed a higher accuracy rate in irregular verbs with high compared to mid/low frequency, but only L1 children were sensitive to phonological properties of the irregular verbs showing: a) errors of vowel change in verbs ending in an alveolar /*t*/, */d/* in the stem (*write, ride*), and 2) more overregularization errors than errors of omission or vowel change in verbs ending in an alveolar /*t*/, */d/* in the past tense, but not in the stem (*catch*), and in verbs that do not end in a alveolar /*t*/, */d/* in the past tense, but end in a consonant (*dig*) or vowel (*blow*). This is also a novel finding demonstrating that L2 children

have acquired the suffixation rule for past tense and are sensitive to the frequency of the input, and therefore, they are more accurate in high frequency compared to mid/low frequency verbs, which they overregularize using the suffixation rule. Their lack of sensitivity to phonological properties of irregular verbs suggests that the suffixation rule overrides the use of schemas for the production of irregular past tense forms. Thus, the properties of individual items, such as frequency and phonological properties, should be taken into consideration in the evaluation of results from existing language assessments that do not control for those factors, and in the development of assessment material.

The positive correlation between TROG2 and BPVSII shows that comprehension of grammar develops in relation to comprehension of single-word vocabulary. The positive correlation between third person singular *–s* and regular past tense *–ed* suggests that although L2 children are less accurate in *–s* than in *–ed*, the two morphemes are acquired in tandem as a reflection of the acquisition of tense. The correlations between regular and irregular past tense, TROG2 and BPVSII suggests that children with better comprehension of grammar and vocabulary are better in acquiring specific grammatical phenomena, such as the marking of past tense. Finally, the positive correlations between YoE third person singular *–s* and past tense confirm that L2 children's production of tense morphology increases in relation to YoE to English.

These findings have important clinical implications for the assessment of children with English as L2. The age of the children, the years of exposure, and the type of task seem to be crucial for the validity of the monolingual norms for L2 children. L2 children require more years of exposure to reach monolingual norms in the comprehension of vocabulary than comprehension of grammar and production of tense morphemes. The delay in the acquisition of tense morphemes in L2 children attested in previous studies testing children with a relatively short age of exposure (e.g., Ionin & Wexler, 2002; Paradis, 2005; Paradis et al.,

2008) does not seem to be present in the majority of the children of the present study who were eight and nine years old and had 4 or more YoE to English. This suggests that four YoE to English are sufficient for most eight and nine year-old L2 children to score within norms for L1 children on the TEGI. Thus, tense marking and the screening test of the TEGI may be a promising measure for differential diagnosis in eight and nine year-old children with English as a L2 that have at least four years of exposure to English.

Acknowledgments

This research was supported by the Economic and Social Research Council research grant 'Real-time processing of syntactic information in children with English as a Second Language & children with Specific Language Impairment' awarded to Theo Marinis (RES-061-23-0137). We would like to thank Halit Firat for collecting the data from the L2 children, the schools and families for participating in this project, and two anonymous reviewers for their constructive comments and suggestions.

References

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX Lexical Database, Release 1 (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bishop, D. V. M. (2003). *Test for Reception of Grammar, Version 2 (TROG-2)*. London: The Psychological Corporation.

Blevins, J. P. (2006). English inflection and derivation. In B. Aarts & A. M. McMahon (Eds.), *Handbook of English Linguistics* (pp. 507-536). London: Blackwell.

Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language, 58*, 265-289.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 46*, 1029-1037.

Crutchley, A., Conti-Ramsden, G., & Botting, N. (1997). Bilingual children with Specific Language Impairment and standardised assessments: preliminary findings from a study of children in language units. *International Journal of Bilingualism, 1*, 117-134.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale: Second Edition*. Windsor: NFER-Nelson.

Gavruseva, L., & Lardiere, D. (1996). The emergence of extended phrase structure in child L2 acquisition. In A. Stringfellow, D. Cahama-Amitay, E. Hughes & A. Zukowski (Eds.), *Proceedings of the 20th annual Boston University Conference on Language Development, Vol. 1* (pp. 225-236). Somerville, MA: Cascadilla Press.

Gutierrez Clellen, V. F. (2000). Language sample analysis in Spanish-speaking children: methodological considerations. *Language, Speech and Hearing Services in Schools, 31*(1), 88-98.

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. (2001). The acquisition of the IP system in child L2 English. *Studies in Second Language Acquisition, 23*(1), 1-39.

Ionin, T., & Wexler, K. (2002). Why is "Is" easier that "-s"?-Acquisition of tense/agreement morphology by child L2-English learners. *Second Language Research, 18*(2), 95-136.

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.

Lakshmanan, U. (1994). *Universal grammar in child second language acquisition: Null subjects and morphological uniformity*. Amsterdam: Benjamins.

Mahon, M., & Crutchley, A. (2006). Performance of typically-developing school-age children with English as an additional language (EAL) on a test of receptive vocabulary *Child Language Teaching and Therapy, 22*, 333-351.

Marchman, V. A. (1997). Children's productivity in the English past tense: The role of frequency, phonology, and neighborhood structure. *Cognitive Science, 21*, 283-304.

Matthews, D. E., & Theakston, A. L. (2006). Errors of omission in English-speaking children's production of plurals and past tense: the effects of frequency, phonology, and competition. *Cognitive Science, 30*, 1027-1052.

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. Gavruseva & B. Haznedar (Eds.), *Current trends in child second language acquisition: a generative perspective* (pp. 337-356). Amsterdam: John Benjamins.

Paradis, J., & Crago, M. (2000). Tense and temporality: A comparison between children learning a second language and children with SLI. *Journal of Speech and Hearing Disorders, 43*(4), 834-848.

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.

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, Hartcourt Assessment Company.

TABLES

Table 1: Bybee & Slobin's classification of irregular verbs

	Verb Class	Example
I	Verbs with no change in the past tense	*cut-cut*
II	Verbs that change a final /*d*/ to /*t*/	*send-sent*
III	Verbs with an internal vowel change and a final /*t*/ or /*d*/	*feel-felt*
IV	Verbs with an internal vowel change, deletion of a consonant and addition of a final /*t*/	*catch-caught*
V	Verbs with an internal vowel change and whose stem end in a */t/* or */d/*	*ride-rode*
VI	Verbs with an internal vowel change of /i/ to /ae/ or to /^/	*sing-sang*
VII	All other verbs with an internal vowel change	*give-gave*
VIII	All verbs with a vowel change that end in a diphthongal sequence	*blow-blew*

Table 2: Results from the Test for Reception of Grammar 2 (TROG2) and the British Picture Vocabulary Scale II (BPVSII)

		L1 *(N = 33)*			L2 *(N =38)*		
		Mean	SD	Range	Mean	SD	Range
TROG2	Raw score	13.5	(2.0)	10-18	11.1	(3.7)	1-18
	Standard score	100.1	(9.3)	85-120	88.8	(15.4)	55-113
	Z-score	0	(0.6)	-1.0 – 1.3	-1.0	(-0.7)	-3.0 – 0.9
BPVS II	Raw score	75.4	(11.2)	49-91	54.3	(11.2)	31-83
	Standard score	99.7	(8.4)	87-115	81.1	(7.2)	65-101
	Z-score	-0.5	(2.0)	-0.9 – 1.0	-1.2	(0.5)	-2.3 – 0.1

Table 3: Verb classes, frequency and phonological properties of verbs on the Rice/Wexler Test of Early Grammatical Impairment (TEGI)

Verb	Verb Class	Log frequency (past tense)	frequency	Stem *t* or *d*	Past *t* or *d*
make	III	2.2989	High	No	Yes
catch	IV	1.3424	Mid	No	Yes
write	V	1.5051	Mid	Yes	Yes
eat	V	1.0414	Mid	Yes	Yes
ride	V	.6021	Low	Yes	Yes
dig	VI	.4777	Low	No	No
give	VII	1.8573	High	No	No
blow	VIII	.6990	Low	No	No

Table 4: L1 children's response types by verb

Frequency	Verb	Verb class	Correct	Over-regularizations	Uninflected	Vowel change
High	Make	III	93.3	3.3	3.3	0
	Give	VII	81.2	9.4	9.4	0
Mid/Low	Write	V	32.2	35.7	0	32.1
	Eat	V	91.2	3.4	3.4	0
	Ride	V	14.8	33.3	7.4	44.5
	Blow	VIII	53.1	46.9	0	0
	Catch	IV	29	71	0	0
	Dig	VI	29	71	0	0

Table 5: L2 children's response types by verb

Frequency	Verb	Verb class	Correct	Over-regularizations	Uninflected	Vowel change
High	Make	III	51.4	40	8.6	0
	Give	VII	54.3	40	5.7	0
Mid/Low	Write	V	15.6	68.8	15.6	0
	Eat	V	21.9	56.2	21.9	0
	Ride	V	6.9	79.3	13.8	0
	Blow	VIII	5.9	94.1	0	0
	Catch	IV	5.9	82.4	11.7	0
	Dig	VI	5.8	82.5	11.7	0

Table 6: Correlations between tense morphemes, and raw scores of the Test for Reception of Grammar 2 (TROG2) and the British Picture Vocabulary Scale II (BPVSII)

	TROG2	BPVSII	3rd Person	Past tense	Regular past tense
BPVSII	.591(**)				
3rd Person	.583(**)	.600(**)			
Past tense	.594(*)	.368(*)	.384(*)		
Regular Past tense	.301	.424(**)	.484(**)	.751(**)	
Irregular past tense	.333(*)	.463(**)	.026	.095	-.086

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

FIGURE CAPTIONS

Figure 1: Accuracy in third person –*s* and past tense following the Test of Early Grammatical Impairment (TEGI) scoring procedure

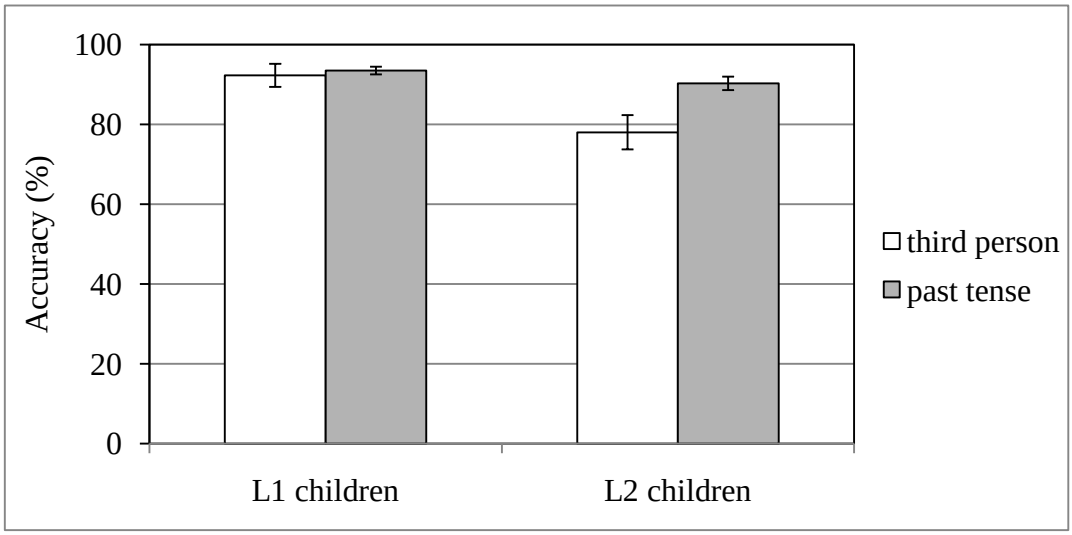

Figure 2: Percentage of L2 children meeting the criterion score for the Test of Early Grammatical Impairment (TEGI) (by age)

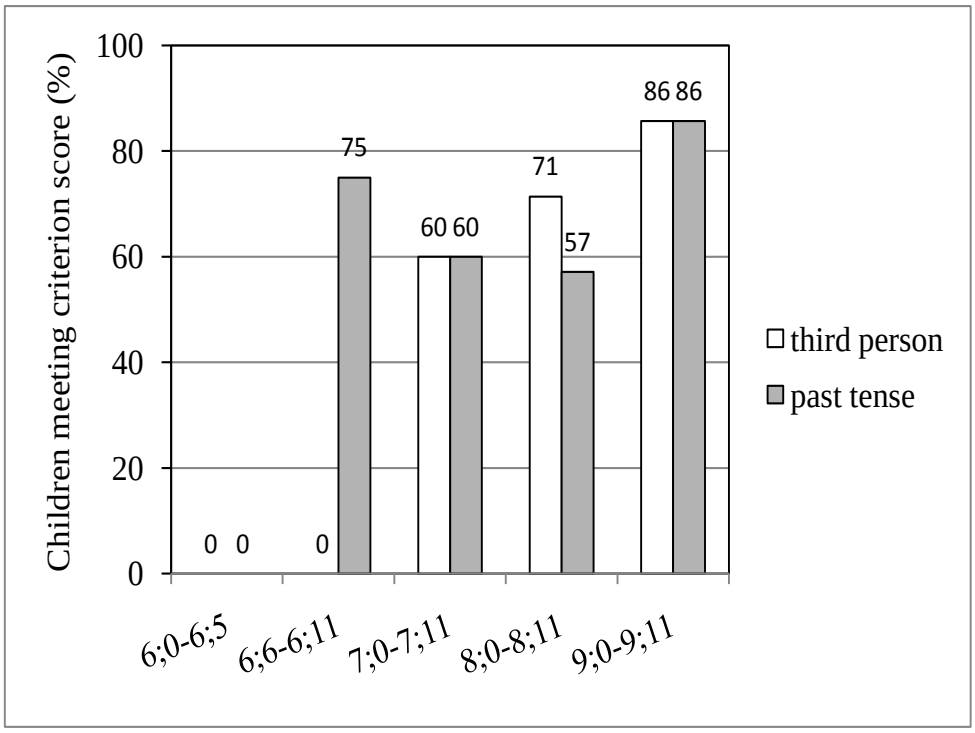

Figure 3: Percentage of L2 children meeting the criterion score for the Test of Early Grammatical Impairment (TEGI) (by YoE)

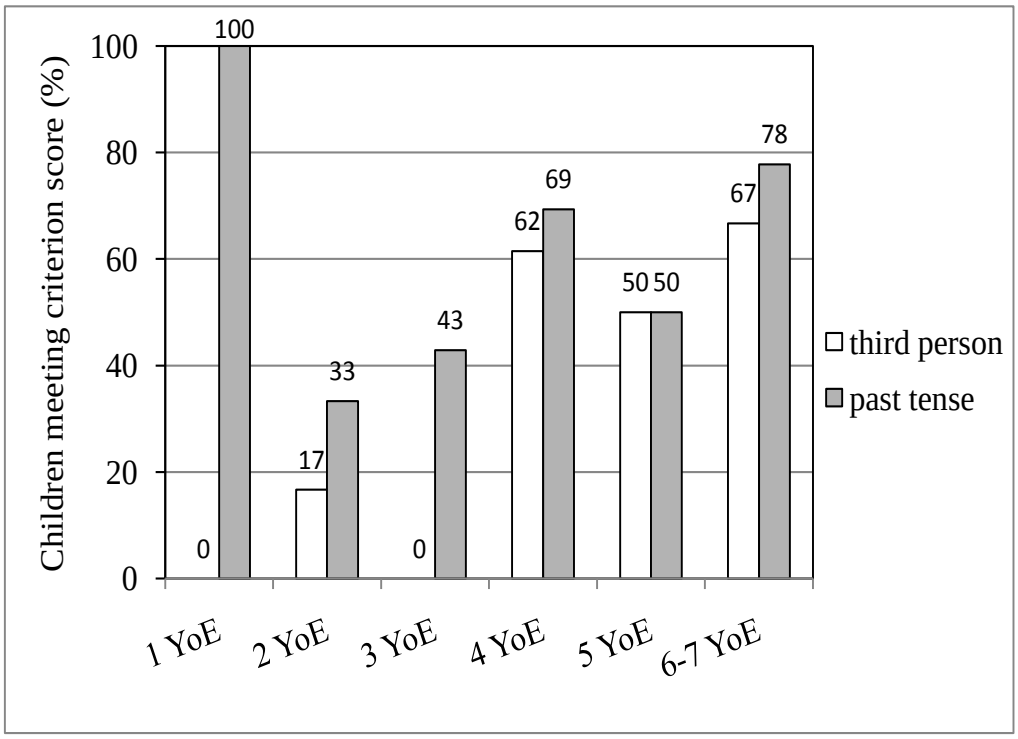

Figure 4: Accuracy of regular vs. irregular past tense forms

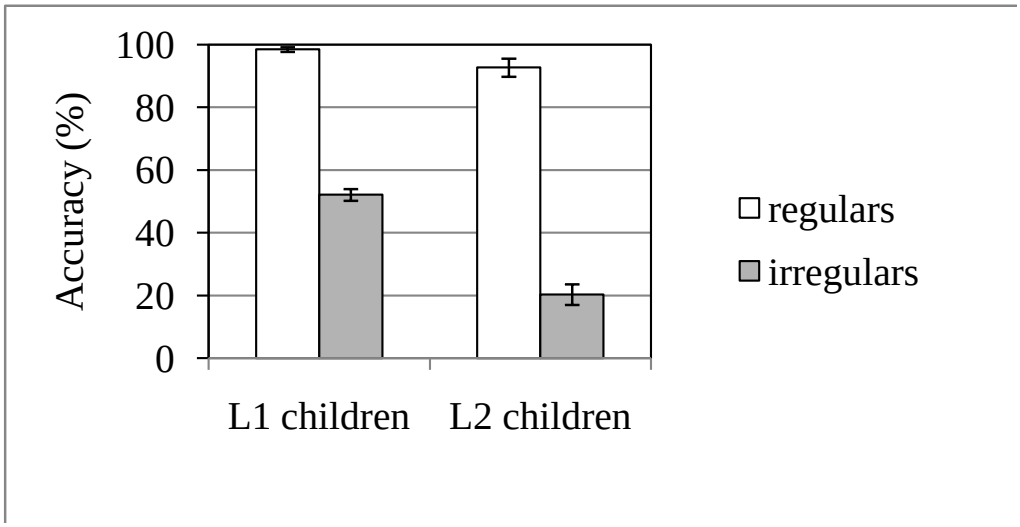

Figure 5: Error analysis of irregular verbs