The Development of Language Learning Aptitude and Metalinguistic Awareness in Primary-School Children: A Classroom Study

Angela Tellier
University of Essex
atelli@essex.ac.uk

Karen Roehr-Brackin
University of Essex
kroehr@essex.ac.uk

Essex Research Reports in Linguistics

Volume 62

Number 1

5 April, 2013

Dept. of Language and Linguistics,
University of Essex,
Wivenhoe Park,
Colchester, Essex, UK,
CO4 3SQ

http://www.essex.ac.uk/linguistics/publications/errl/

Essex Research Reports in Linguistics present ongoing research activities of the members of the Department of Language and Linguistics.

The main purpose of these reports is to provide a quick publication outlet. They have 'pre-publication status', and most will subsequently appear in revised form as research articles in professional journals or in edited books.

If you have technical problems downloading a paper, or for further information about these reports, please contact the editor:

Doug Arnold: doug@essex.ac.uk.

Citation Information:

Angela Tellier and Karen Roehr-Brackin. 'The Development of Language Learning Aptitude and Metalinguistic Awareness in Primary-School Children: A Classroom Study', Essex Research Reports in Linguistics, Vol. 62.1. Dept. of Language and Linguistics, University of Essex, Colchester, UK, April, 2013.
http://www.essex.ac.uk/linguistics/publications/errl/errl62-1.pdf

The Development of Language Learning Aptitude and Metalinguistic Awareness in Primary-School Children: A Classroom Study

Angela Tellier and Karen Roehr-Brackin
University of Essex

Abstract

In the typical foreign language classroom, many learners all over the world find themselves in a minimal-input environment. Existing research suggests that in such a setting, adolescents typically outperform younger children. The greater cognitive maturity of older learners manifests itself in greater language learning aptitude, greater metalinguistic awareness, and enhanced capacity for explicit learning. We examined whether the teaching and learning of either Esperanto or French would facilitate the development of language learning aptitude and metalinguistic awareness in 8-9-year-old children (N=28), thus setting the scene for enhanced explicit learning even at a young age. Following instruction in either Esperanto or French over a school year, children made significant gains on measures of aptitude, metalinguistic awareness, and L2 proficiency. Effect sizes in the Esperanto group were larger throughout, however, with greater homogeneity of performance in evidence and a closer association between aptitude, metalinguistic awareness, and L2 proficiency at the end of the treatment. Moreover, Esperanto proved significantly easier to learn than French, with larger gains in L2 proficiency achieved by the Esperanto group compared with the French group. Finally, we found that language-analytic ability emerged as a significant predictor of L2 achievement in the sample as a whole.

1. Introduction

It is not uncommon for second language (L2) instruction to begin in the early years of a child's schooling. Quite often, classroom-based language learning and teaching takes place in a minimal-input environment, with exposure to the L2 restricted to a small number of hours per week over a limited number of years. At the same time, existing research with classroom

language learners has demonstrated that later starters consistently outperform younger starters on measures of L2 achievement (Cenoz, 2003; García Mayo, 2003; Harley & Hart, 1997; Larson-Hall, 2008; Muñoz, 2006, 2009). Indeed, it appears that in order for L2 learning at an early age to be successful, considerable amounts of input are needed. In the absence of high-quality, large-quantity input, a certain level of cognitive maturity seems to be necessary for learning to be successful. Greater cognitive maturity allows for effective explicit learning, which is associated with (components of) language learning aptitude and metalinguistic awareness (DeKeyser, 2003; García Mayo, 2003; Larson-Hall, 2008; Muñoz, 2009). The aim of the present study was to explore whether the development of language learning aptitude and metalinguistic awareness (and, by implication, effective explicit learning) in children aged 8-9 could be enhanced by means of a specific tool: the teaching and learning of Esperanto.

2. Background to the present study

The concepts of interest to the present study are language learning aptitude, metalinguistic awareness, and explicit learning. In what follows, these concepts are reviewed, with particular reference to child L2 learning and teaching in the classroom.

2.1 Language learning aptitude

Existing research in the field of L2 learning suggests that language learning aptitude, defined as a set of abilities which enables some learners to acquire new language material more quickly and with greater ease than others (Dörnyei, 2005), is a strong and significant predictor of L2 learning success. There is empirical support for this finding from studies with adult learners in different instructional settings (Ehrman & Oxford, 1995; Hummel, 2009; Skehan, 1998). Researchers subscribing to the classic, multi-componential concept of language learning aptitude as exemplified by Carroll's model (1962, 1981; 1990; see below for details) tend to agree that aptitude is relatively stable in adults (Kiss & Nikolov, 2005; Skehan, 1998).

Compared with aptitude research in adults, studies with child learners are less readily available. Kiss and Nikolov (2005) reported that scores on a measure of language learning aptitude explained over 20% of the variance in L2 proficiency scores achieved by 12-year-old Hungarian learners of L2 English. Investigations into aptitude in younger children are even less common, despite the argument that such research is clearly of interest to L2 learning theory as well as educational practice (Kiss, 2009).

When developing a test of language learning aptitude for use with 8-year-old Hungarian children, Kiss (2009) relied on Carroll's classic model of aptitude (1962, 1981, 1990) as her theoretical guide, and the associated tests for English speakers, the adult MLAT (Modern Language Aptitude Test; Carroll & Sapon, 2002b) and the MLAT-E for children (Modern Language Aptitude Test - Elementary; Carroll & Sapon, 2002a) as her empirical guide. In addition, the researcher drew on an existing aptitude measure for Hungarian adults as well as Kiss and Nikolov's (2005) test for 12-year-olds. Kiss (2009) used the test to identify a group of higher-aptitude children for participation in a newly introduced bilingual teaching program. At the end of the first year, the children were tested for oral L2 proficiency and ranked according to teacher judgment for overall in-class performance. A number of significant correlations provided evidence for an association between aptitude and subsequent L2 achievement in the young participants.

Kiss' (2009) approach to studying 8-year-old children's aptitude stands in contrast with the argument put forward by Milton and Alexiou (2006). These researchers posit that aptitude in young children cannot easily be described in terms of the classic conceptualization and operationalization of the construct as exemplified by Carroll's model and tests such as the MLAT-E, since the latter requires highly developed literacy skills. Instead, Milton and Alexiou favour a measure which focuses on more general cognitive abilities such as memory, phonetic, and analytic skills, ideally without making exclusive reference to language-specific abilities. In a study conducted with Greek children aged 5, 6, and 7 years, an aptitude test examining four cognitive abilities via "picture-based and game-like activities appropriate for the learners' ages" was administered (Milton & Alexiou, 2006, p. 185). The measure

comprised a test of rote memory, a semantic integration task, a paired-associates test, and an artificial language game. Results suggest a general trend according to which the older children performed better than the younger children. This developmental pattern is particularly visible with regard to analytic abilities as measured by the artificial language and semantic integration tasks. It is worth noting, though, that these are merely trends, with only few statistically significant differences in evidence. The 5-year-olds achieved significantly lower scores on the aptitude test as a whole compared with the other two cohorts.

Milton and Alexiou (2006) conclude that aptitude in young children cannot be regarded as a fixed quality; instead, it changes over time as children grow older and develop cognitively (see also Suárez & Muñoz, 2011 for a similar conclusion based on a study with 8-12-year-old children). This finding notwithstanding, the researchers believe that aptitude profiles can be useful for diagnostic purposes, i.e. an aptitude measure should allow us to identify a child's specific strengths and weaknesses. As young children's cognitive abilities are still developing, it is plausible to suggest that these abilities can be promoted and enhanced via practice and training. Regarding the observed developmental trend for analytic abilities, Milton and Alexiou highlight the apparent improvement in scores which occurred from around age 6 onwards, that is, perhaps rather earlier than one might have anticipated. They note that "[t]he growth of analytic skills, and their connection with foreign-language learning success, suggests that this learning is explicit; even at very young ages" (2006, p. 190).

2.2 Explicit and implicit learning in relation to age

The suggestion that such young children might engage in explicit L2 learning is not a common argument. The more widely represented view is that child L2 learning relies above all on implicit processes (e.g. DeKeyser, 2003; Muñoz, 2009), that is, learning without conscious awareness. Explicit learning occurs when a learner consciously and deliberately attempts to master language material or solve a language-related problem (Dörnyei, 2009). Indirect evidence for children's successful implicit L2 learning has arisen from research into age-of-onset effects in immersion settings. Such research indicates that although young

children initially learn more slowly than adults, they are likely to eventually outperform older learners and reach higher levels of L2 proficiency, provided that intensive exposure to the L2 continues over a considerable number of years (for reviews, see Birdsong, 2006; Hyltenstam & Abrahamsson, 2003). While this finding constitutes a consensus view among L2 researchers, it is important to bear in mind that it has arisen from – and thus applies to – L2 learning in naturalistic environments only. Research conducted in Canadian immersion classrooms has resulted in more differentiated findings. Harley and Hart (1997), for instance, have argued that L2 learning may depend on different cognitive abilities, depending on whether intensive exposure begins in early childhood or in adolescence.

In contrast with learners experiencing naturalistic exposure or intensive exposure in an immersion classroom, learners in a foreign language classroom typically experience minimal input only, often restricted to two or three lessons a week during term time for a limited number of years. Several recent studies have investigated age-of-onset effects in a foreign language learning environment (e.g. Cenoz, 2003; García Mayo, 2003; Larson-Hall, 2008; Muñoz, 2006, 2009). In large-scale studies undertaken in Spain, classroom learners with different starting ages (4, 8, 11, 14 and 18+ years) were compared on a range of L2 achievement measures at different stages, that is, after about 200, 400, 600 and/or 700 hours of instruction. Overall, older starters consistently and significantly outperformed younger starters on virtually all measures. By way of explanation, it is argued that the older learners' more advanced cognitive development is likely to be the determining factor, especially in the domain of morphosyntactic learning, which appears to boost at around age 12 (Cenoz, 2003; García Mayo, 2003; Muñoz, 2006, 2009).

Acknowledging the argument that older learners' greater cognitive maturity may give them the edge over younger learners in the context of formal achievement measures, Larson-Hall (2008) employed a research design which controlled not only for amount of exposure, but also for learners' age at the time of testing and their language learning aptitude. She examined the long-term effects of an early versus a late start in 18-year-old Japanese learners of English as a foreign language (N=200). All participants had studied the L2 for six years, but had

started either early, that is, before entering junior high school, or late, at age 12-13. The participants completed an oral grammaticality judgment task, a phonemic discrimination test, and a test of language learning aptitude. Overall, Larson-Hall (2008) reports modest effects for an early starting age in both grammatical and phonological abilities. Importantly, however, age interacted with amount of input. On the grammaticality judgment task, early starters scored more highly than late starters once they had had about 1,600 to 2,200 hours of input. Thus, whilst L2 morphosyntactic abilities may benefit from an early start, this only applies after a substantial amount of input that goes well beyond the kind of exposure typically available in a foreign language classroom. Advantages for early starters on the phonemic discrimination test emerged in the range of 1,200-2,200 hours of input, i.e. slightly earlier, but effect sizes were modest. Moreover, an early starting age played only a minor explanatory role, accounting for no more than 3-14% of the variance in scores.

To summarize, studies investigating age-of-onset effects in the context of classroom-based foreign language learning suggest that an early starting age confers no benefits in the minimal-input setting of the average school classroom. Indeed, later starters consistently and significantly outperformed earlier starters on a range of achievement measures in most existing studies. A likely explanation for this pattern of results is the more advanced cognitive development of the later starters. In sum, therefore, existing research suggests that (1) in order for L2 learning at an early age to be successful, considerable amounts of input are needed; (2) in order for L2 learning in a minimal-input setting to be successful, a certain level of cognitive maturity appears to be required.

Cognitive maturity facilitates L2 acquisition in a minimal-input setting because it allows for explicit learning. As mentioned above, explicit processes are conscious and deliberate (Dörnyei, 2009). Requiring attention and effort, they rely on a learner's working memory – a limited resource which has greater capacity in adults than in young children. Whilst taxing in nature, explicit learning is relatively fast and efficient, enabling a learner to benefit from L2 input even if it is provided in relatively small quantities and/or over a relatively short period of time. Indeed, there is ample evidence from research with cognitively mature learners that

explicit teaching and learning effectively promote L2 acquisition in the classroom (see Norris & Ortega, 2001 for a comprehensive review).

Explicit teaching and learning draw on a learner's metalinguistic awareness, i.e. their thinking about language in terms of its nature, function, and form. Metalinguistically aware individuals are able to treat language as an object of reflection (Baker, 2006; Cummins, 1987; Gombert, 1992). "Metalinguistic awareness implies that attention is actively focused on the domain of knowledge that describes the explicit properties of language" (Bialystok, 2001, p. 127). Metalinguistic awareness is related to metalinguistic knowledge, or knowledge about language, and metalinguistic ability, or the capacity to use knowledge about language (Bialystok, 2001). Metalinguistic awareness requires high levels of both analysis of knowledge and control of processing (Bialystok & Ryan, 1985). Analysis of knowledge refers to the structuring and explication of linguistic knowledge representations, while control of processing refers to the direction of attention as well as the selection and integration of information (Bialystok, 1988, 1994). Hence, metalinguistic awareness, knowledge, and ability are cognitively demanding and can thus be expected to be associated with higher levels of cognitive development.

3. The present study

From a language-teaching perspective, the research reviewed above presents a conundrum. In the typical foreign language classroom, many learners find themselves in a minimal-input environment. At the same time, foreign language instruction often starts early, e.g. when children are 6, 7, or 8 years old and thus not yet cognitively mature. In recent years, for instance, several European countries have adopted precisely such an approach, offering language classes not only in primary or elementary school, but sometimes even at pre-school age, in nursery or kindergarten classes (Enever, 2009; European Commission, 2008). This raises the question of whether the L2 learning of young children in a minimal-input setting can be maximized by accelerating children's cognitive development with a view to kick-starting explicit learning processes.

3.1 Esperanto as a propaedeutic tool

The research reviewed above suggests that cognitive maturity and the associated enhanced capacity for explicit L2 learning manifest themselves in higher levels of language learning aptitude in general, higher levels of language-analytic ability in particular, and higher levels of metalinguistic awareness. In child learners who have not yet reached cognitive maturity, it is particularly difficult to ascertain which of these variables might be causative factors and which might be consequences. It seems plausible to suggest that, in young children, (components of) language learning aptitude and metalinguistic awareness influence each other and grow together in a mutually reinforcing cycle. It may even be possible to argue that language-analytic ability (as a component of aptitude) and metalinguistic awareness are partially overlapping constructs, although this argument admittedly depends on exactly how these constructs are operationalized.

Specific operationalizations aside, common conceptual ground can be found. In the most general terms, language-analytic ability and metalinguistic awareness draw on the human capacity to analyse and problem-solve in the linguistic domain. Language-analytic ability is the ability to identify and extrapolate linguistic patterns (Dörnyei & Skehan, 2003; Erlam, 2005; Ranta, 2002; Skehan, 2002). Metalinguistic awareness refers more broadly to the ability to think about language in terms of its nature, function, and form (Baker, 2006; Cummins, 1987; Gombert, 1992). Language-analytic ability and metalinguistic awareness enable the learner to direct attention to the formal properties of language, to make comparisons, to identify regularities, to extract patterns, and to link these with the functional properties of language, i.e. semantic and pragmatic meaning. Moreover, language-analytic ability and metalinguistic awareness enable the learner to move from exemplar to rule and from rule to exemplar, i.e. to think inductively as well as deductively in the domain of language. What better way of honing such abilities than to teach and learn a language whose structures are entirely regular and systematic, and whose form-function mappings are entirely transparent? Esperanto is such a language, and the argument that Esperanto may serve as a propaedeutic tool for subsequent L2 learning in the classroom is not new.

The direct phoneme-grapheme correspondence and simple and regular grammar of Esperanto should result in low learning difficulty (for research on L2 learning difficulty, see DeKeyser, 2005; Ellis, 2006). Early 20th century educationalists predicted benefits in teaching Esperanto before or alongside other languages, such as heightened metalinguistic awareness, more positive attitudes to language learning, improved L1 literacy, and greater self-esteem (Corsetti & LaTorre, 1995; Fantini & Reagan, 1992; Markarian, 1964; Symoens, 1989). However, there has been surprisingly little empirical research which has put these hypotheses to the test. Early work reported promising results (e.g. Fisher, 1921; Halloran, 1952; Williams, 1965a, 1965b), but the studies in question suffered from a number of shortcomings, if considered by today's standards (see, for instance, Corsetti & LaTorre, 1995; Fantini & Reagan, 1992).

More recent studies (Barton & Bragg, 2010; Bishop, 1997) are few and far between and at times appear to lack desirable experimental control, though this is perhaps understandable, given the difficulties of conducting research in schools. Results from a case study examining a curriculum initiative that uses Esperanto as a tool for raising language awareness in primary-school children in Britain are encouraging. The researchers report that children "demonstrated high levels of skill in those tasks which required them to translate unfamiliar languages" (Barton & Bragg, 2010, p. 19). The overall findings suggest that Esperanto may indeed function as a catalyst for the development of both metalinguistic awareness and linguistic competence.

In summary, up-to-date, methodologically sound research into the propaedeutic potential of Esperanto is scarce. What is more, research which has investigated the teaching and learning of Esperanto in connection with language learning aptitude and metalinguistic awareness is not available at all. The present study was designed to begin to fill this gap.

3.2 Research issues

The study reported here was quasi-experimental, employing a pre-test/post-test design. It was carried out over one school year with an intact group of children in an English state primary

school. The children, who were 8-9 years of age at the time of the study, were assigned to two treatment groups (see below for details). Group E was taught Esperanto, Group F was taught French.[1] All children were tested for language learning aptitude, metalinguistic awareness, and L2 proficiency at the beginning and end of the treatment.[2] The following research questions were addressed:

1. Do the children in Group E and Group F make gains on measures of language learning aptitude, metalinguistic awareness, and L2 proficiency?
2. Are there any differences between Group E and Group F on measures of language learning aptitude, metalinguistic awareness, and L2 proficiency?
3. What is the relationship between language learning aptitude, metalinguistic awareness, and L2 proficiency at pre-test and at post-test?
4. Which variables predict L2 proficiency at post-test?

3.3 Instrumentation

Language learning aptitude, metalinguistic awareness, and L2 proficiency were assessed by means of dedicated tests, as follows.

3.3.1 Test of language learning aptitude: MLAT-E(UK)

Children's language learning aptitude was measured by means of a slightly modified version of the MLAT-E(UK), the British English version of the MLAT-E (Carroll & Sapon, 2002a). The test was piloted with a small group of children aged 8-9 (N=10); it showed a good distribution of scores, and reliability (Cronbach's α) was either good or excellent, ranging from .77 to .98 on individual subtests. Thus, the adapted test was retained for use in the main study. It comprised four sections, and the reliability indices remained consistently high (α at pre-test is reported):

- Hidden Words (30 items, α = .82) presents English keywords spelled approximately as pronounced; the child must choose whichever of four correctly spelled words matches the keyword most closely in meaning.
- Matching Words (30 items, α = .89) draws attention to an English keyword in a sentence and asks the child to choose a word that plays the same grammatical role in another sentence given underneath.
- Finding Rhymes (45 items, α = .95) presents an English keyword and four possible rhyming alternatives from which the child must choose the best matching rhyme.
- Number Learning (25 items, α = .96) requires the child to write in figures words for numbers in an invented language spoken aloud by the test administrator; the words are taught immediately before the test.

The MLAT-E(UK) was administered at the beginning and end of the school year by the first author.

3.3.2 Test of metalinguistic awareness: Polyglot

The test of metalinguistic awareness was a short measure entitled Polyglot designed by the first author. The test comprised two task types. The first required children to identify pairs of sentences with the same meaning in seven European languages. The second asked children to translate sentences from three European languages into English. Thus, unlike the measure of language learning aptitude which was based on English (three subtests) or an invented language (one subtest), the Polyglot drew on other European languages, most of which the participating children had not been exposed to before; the two treatment languages, Esperanto and French, were included as well. The tasks encouraged the drawing of comparisons between languages, the identification of similarities in form or meaning, and the transfer of knowledge from one language to another. The maximum score was 30. The test was piloted with a group of children aged 7-9 (N=45) and retained unchanged. Reliability was high (α = .88 at pre-test). The Polyglot was administered at the beginning and end of the school year by the language teacher who taught the treatment languages.

3.3.3 Tests of L2 proficiency

The tests of L2 proficiency were matched tests for the two treatment languages, Esperanto and French, developed by the first author. Subtests focused on core vocabulary and structures taught in the treatment sessions and assessed children's skills in reading, writing, listening, and speaking. Unfortunately, results for the speaking subtest are not available because of an error by the test administrator. In what follows, L2 proficiency refers to results based on the reading, writing, and listening components, each of which carried a maximum score of 6. Varied tasks were used, with colourful picture stimuli and short sentences intended to be both non-threatening to the young test takers and to maintain their attention. The reliability of the Esperanto test was good ($\alpha = .83$ at post-test); following the removal of five unreliable items, the reliability of the French test was likewise acceptable ($\alpha = .75$ at post-test). The L2 proficiency tests were administered at the beginning and end of the school year by the language teacher who taught the treatment languages.

3.4 Experimental treatment

The treatment was based on specifically designed, matched teaching programs for Esperanto and French developed by the first author. Five topic areas were addressed – colours, animals, numbers, families, and self – with each area the focus of instruction for about six weeks. In class, each experimental group learned a similar range of vocabulary items and conversational phrases, sang songs to practise idioms, and completed task sheets. Activities across the two groups were kept as similar as possible, with allowances made for the morphosyntactic differences between Esperanto and French. Lessons covered all four skills (Cameron, 2001). The teaching method was predominantly communicative, with incidental focus on form (Nicholas & Lightbown, 2008). Both treatment groups were taught by the same specialist language teacher for 45 minutes per week over a period of nine months, which translates into a total of 22.5 hours of instruction provided over one school year.

3.5 Participants

The participant sample was drawn from intact mixed-ability classes in an English state primary school. Children in year 4 (age 8-9) took part in the present study, which began with 29 and ended with 33 participants. Data sets from children who were not present for both administrations of the MLAT-E(UK) were excluded from analysis, so the final participant sample comprised 28 children. A total of 26 children were L1 speakers of English; two children reported Polish as their home language, which is in keeping with the fact that two children were registered as bilingual by the school. There was one bilingual child in each of the two treatment groups (see below); the two bilingual children did not produce outlying scores on any of the measures used in the study. Most children had been exposed to more than one language through visits abroad, family, friends, language clubs, and occasional school input. In terms of language experience, the sample can be considered as broadly representative of intact classes in English state primary schools. In order to preserve ecological validity, no further participants were excluded on the grounds of their language history.

The cohort was randomly assigned to two treatment groups by their regular class teacher. Group E followed the Esperanto program and comprised 14 children, 4 boys and 10 girls, mean age = 8;6. Group F followed the French program and likewise comprised 14 children, 4 boys and 10 girls, mean age = 8;7. There were no significant differences in language learning aptitude between Group E and Group F at the beginning of the treatment, $t(26) = .53$, $p = .6$, two-tailed.

4. Results

The first research question asked whether the children in Group E and Group F made gains on measures of language learning aptitude, metalinguistic awareness, and L2 proficiency in the course of the school year. Table 1 shows the descriptive statistics for language learning aptitude.

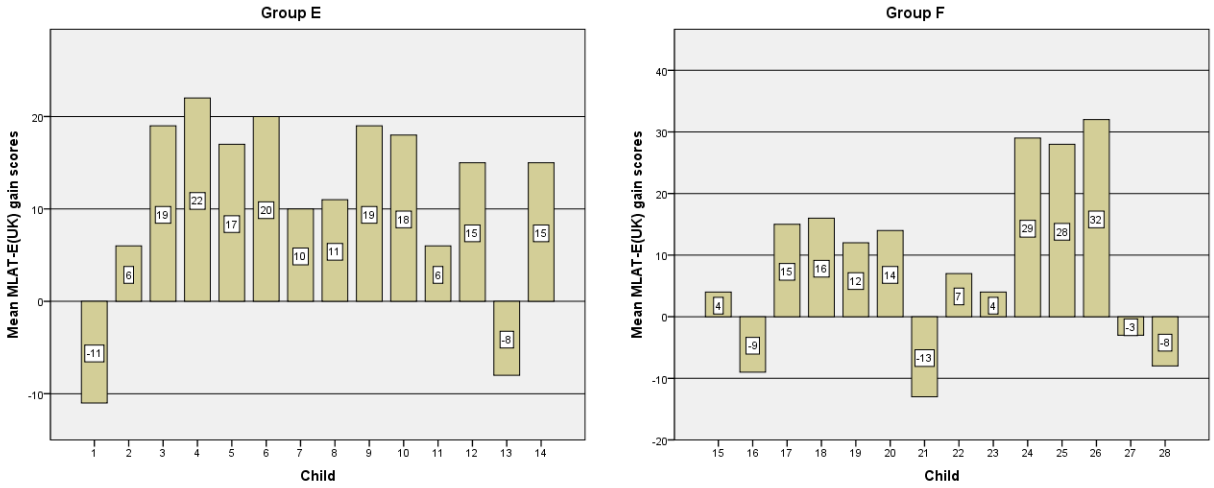
Table 1 Descriptive statistics for language learning aptitude

	Group E MLAT-E(UK) Pre-test	Group E MLAT-E(UK) Post-test	Group F MLAT-E(UK) Pre-test	Group F MLAT-E(UK) Post-test
N	14	14	14	14
Mean % correct	60	69	56	63
Mean	78.0	89.36	72.79	81.93
SD	24.78	29.69	26.87	31.88
Minimum	28	17	29	16
Maximum	113	123	118	122

The results in Table 1 indicate that the MLAT-E(UK) was not too challenging for the children, with mean facility values ranging from 56% to 69%. In fact, the facility values at pre-test are almost identical to the facility value identified by Kiss (2009) in her pilot study with Hungarian 8-year-olds completing a test modelled on the MLAT-E. In the present study, both groups made significant gains between pre-test and post-test, although the effect size was greater in Group E; Group E: $t(13) = -4.18$, $p < .01$, two-tailed, $\eta^2 = .57$; Group F: $t(13) = -2.37$, $p = .03$, two-tailed, $\eta^2 = .30$. This suggests that language learning aptitude as measured by the MLAT-E(UK) was still developing in the young participants. The increasing range in scores provides further support for this finding.

Figure 1 displays individual children's gain scores on the MLAT-E(UK). It is interesting to note that some children made negative gains, i.e. their performance at the end of the school year was worse than at the beginning. In addition, it is worth noting that Group E is characterized by an overall more homogeneous performance pattern, with solid gains made by individual children being the norm, whereas Group F children show more noticeable individual differences, with dramatic gains of around 30 points, less dramatic gains of around 10 points, as well as a number of losses in evidence.

Figure 1 Individual children's gain scores on the MLAT-E(UK)



Tables 2 and 3 respectively show the descriptive statistics for metalinguistic awareness and L2 proficiency.

Table 2 Descriptive statistics for metalinguistic awareness

	Group E Polyglot Pre-test	Group E Polyglot Post-test	Group F Polyglot Pre-test	Group F Polyglot Post-test
N	14	14	14	13
Mean % correct	15	26	15	24
Mean	4.57	7.86	4.5	7.23
SD	3.67	4.57	2.9	6.29
Minimum	0	1	0	0
Maximum	13	18	12	19

The results in Table 2 indicate that the Polyglot test proved very challenging for the children, with mean facility values ranging from 15% to 26%. Nonetheless, as in the case of the

aptitude measure, both groups made significant gains between pre-test and post-test, although the effect size in Group E was again larger than in Group F; Group E: $t(13) = -3.63$, $p < .01$, two-tailed, $\eta^2 = .50$; Group F: $t(12) = -2.47$, $p = .03$, two-tailed, $\eta^2 = .33$. This result demonstrates that the young participants developed in terms of metalinguistic awareness in the course of the school year.

Table 3 Descriptive statistics for L2 proficiency

	Group E L2 proficiency Pre-test	Group E L2 proficiency Post-test	Group F L2 proficiency Pre-test	Group F L2 proficiency Post-test
N	14	13	14	14
Mean % correct	21	66	32	46
Mean	3.79	11.92	4.14	6.0
SD	1.93	3.04	2.11	2.69
Minimum	1	6	2	3
Maximum	8	17	10	10

Table 3 shows that both groups also made significant gains on the respective L2 proficiency measures; Group E: $t(12) = -9.15$, $p < .01$, two-tailed, $\eta^2 = .87$; Group F: $t(13) = -2.31$, $p = .04$, two-tailed, $\eta^2 = .29$. Group E shows a remarkable improvement in performance on the Esperanto test, with a gain of 45% and an impressively large effect size. By contrast, Group F demonstrates a less marked improvement of 14% on the French test; the effect size is only modest in this instance.

The second research question asked whether there were any differences between Group E and Group F on measures of language learning aptitude, metalinguistic awareness, and L2 proficiency. We found that Group E and Group F did not differ significantly in terms of language learning aptitude and metalinguistic awareness either at pre-test or at post-test. The groups were significantly different in terms of proficiency in their respective treatment

languages, although the effect sizes were small. At pre-test, Group F was significantly more proficient in French than Group E in Esperanto; $t(26) = -2.04$, $p = .05$, two-tailed, $\eta^2 = .14$. This is unsurprising, since none of the children had been exposed to Esperanto before, while several of the children had had minimal prior exposure to French through language clubs offered by the school. At post-test, the pattern was reversed, with Group E significantly more proficient in Esperanto than Group F in French; $t(25) = 2.77$, $p = .01$, two-tailed, $\eta^2 = .23$. Analyses conducted on the gain scores achieved by the two groups confirmed these results. Differences in gains in L2 proficiency were significant, with Group E achieving higher gain scores in Esperanto than Group F in French; $t(25) = 3.89$, $p < .01$, two-tailed, $\eta^2 = .38$. There were no significant differences in gains in aptitude or metalinguistic awareness.

The third research question asked about the relationship between language learning aptitude, metalinguistic awareness, and L2 proficiency at pre-test and at post-test. Table 4 shows the relevant bivariate correlation coefficients (Pearson's *r*) for the two groups.

Table 4 Correlations between aptitude, metalinguistic awareness, and L2 proficiency

Pre-test		Polyglot	L2 proficiency
Group E	MLAT-E(UK)	.41 (p = .15)	.05 (p = .85)
	Polyglot		.38 (p = .18)
Group F	MLAT-E(UK)	.55* (p = .04)	.53 (p = .05)
	Polyglot		.44 (p = .12)
Post-test		**Polyglot**	**L2 proficiency**
Group E	MLAT-E(UK)	.58* (p = .03)	.81** (p < .01)
	Polyglot		.57* (p = .04)
Group F	MLAT-E(UK)	.71** (p < .01)	.57* (p = .04)
	Polyglot		.42 (p = .15)

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

The results in Table 4 demonstrate that there are few significant relationships between variables at the beginning of the treatment. In Group F, there is a correlation of medium strength between performance on the MLAT-E(UK) and the Polyglot. Moreover, the correlation between performance on the MLAT-E(UK) and the L2 proficiency test approaches significance. By contrast, at the end of the treatment, several significant correlations are in evidence. In Group F, language learning aptitude is significantly associated with both metalinguistic awareness and L2 proficiency. The relationships are stronger too. In Group E, all variables are significantly associated. The strong correlation between aptitude and L2 proficiency particularly stands out.

The final research question asked which variables predict L2 proficiency at post-test. In order to address this question, a stepwise multiple regression analysis was conducted. The analysis was exploratory because of the relatively small sample size involved. Performance on the L2 proficiency measures at post-test was the dependent variable; performance on the MLAT-E(UK) subtests and the Polyglot at pre-test constituted the independent variables. The results of the analysis are summarized in Table 5.

Table 5 Stepwise multiple regression coefficients, model summary, and significance (ANOVA)

Model	Independent variables	Standardized β coefficients	R	R^2	F	Sig.
1	Number Learning	.65	.65	.43	18.65	$p < .01$
2	Number Learning	.42				
	Matching Words	.38	.72	.52	12.95	$p < .01$

Table 5 shows that two of the independent variables entered (Model 2) emerged as significant predictors of L2 proficiency, explaining 52% of the variance in scores. The predictor variables are performance on the Number Learning subtest of the MLAT-E(UK), which explains 43%,

and performance on the Matching Words subtest of the MLAT-E(UK), which explains a further 9%.

5. Discussion

It is now possible to draw together the threads of the analysis. The aim of the present study was to examine the potentially facilitative role of Esperanto when child L2 learning takes place in a minimal-input setting. We measured language learning aptitude, metalinguistic awareness, and L2 proficiency in 8-9-year-olds who were taught either Esperanto or French during one school year.

The first finding of note is that the treatment groups made gains on all variables measured at the beginning and end of the year. Both Group E and Group F showed statistically significant gains in terms of metalinguistic awareness, although it is worth noting that the effect size was larger in Group E. Gains were not unexpected, since the children were maturing in the course of the year and metalinguistic skills are subject to growth, with higher levels of metalinguistic awareness, knowledge, and ability associated with higher levels of cognitive development (Bialystok & Ryan, 1985; Gombert, 1992). Moreover, as bilingualism is typically associated with greater metalinguistic awareness (Baker, 2006; Bialystok, 1988; Cummins, 1987), it is not surprising that the children improved on the measure of metalinguistic awareness after regular classroom exposure to an L2.

By the same token, both groups made statistically significant gains in terms of proficiency in the L2 they were taught. This was fully expected, of course. More interestingly perhaps, the effect size in Group E was very large, and a statistically significant difference in gains achieved by the two treatment groups was found, with Group E displaying an improvement of 45%, compared with the less dramatic improvement of 14% attained by Group F.

At first glance, one might wish to attribute this difference to some children's prior exposure to French, which, though minimal, may have meant that there was less room for improvement in

Group F than in Group E. As reported above, Group F was significantly more proficient in French than Group E in Esperanto at the beginning of the school year. However, the descriptive statistics show that Group F did not perform particularly well on the L2 proficiency measure at pre-test, so there was still ample room for making gains of a magnitude similar to that achieved by Group E. If anything, Group F was at an advantage compared with Group E, precisely because Group F had a (small) head-start. Instead, however, Group E attained the greater improvement in L2 proficiency, even though Group E children did not differ from Group F children in terms of language learning aptitude at pre-test. Moreover, at the end of the school year, Group E was more proficient in Esperanto than Group F in French. Hence, the most plausible explanation seems to be that learning Esperanto was easier than learning French in the given classroom setting.

The second important finding arising from the present study concerns the apparent dynamicity of language learning aptitude in the participating children. Both treatment groups improved significantly between pre-test and post-test, which suggests that language learning aptitude as measured by the MLAT-E(UK) was still developing in the young participants. The increased range in scores at the end of the school year provides further support for this finding. Indeed, a small number of children exhibited negative gain scores on the MLAT-E(UK), i.e. their performance worsened. Possibly, such a performance pattern may indicate that the children in question either did not put in as much effort as before or were unable to concentrate at the time of the post-test administration. Alternatively, it may constitute further evidence for the dynamic nature of aptitude in the participants. This latter interpretation is consistent with previous research, which suggests that performance on aptitude tests will develop in conjunction with growing cognitive maturity (Milton & Alexiou, 2006; Suárez & Muñoz, 2011) – but development is not necessarily linear. Interestingly, Group E showed a larger effect size than Group F.

The third noteworthy finding is the apparently greater homogeneity in Group E compared with Group F. With regard to language learning aptitude, solid gains by individual children tended to be the norm in Group E. By contrast, Group F children showed more marked

individual differences, with dramatic gains, minor gains, as well as a number of negative gains in evidence. Correlational analyses yielded further evidence for greater homogeneity of performance in Group E. At the beginning of the treatment, few significant relationships between variables were found. In Group F, performance on the measures of aptitude and metalinguistic awareness were correlated, and the correlation between aptitude and L2 proficiency approached significance. This finding can probably be explained by the children's prior knowledge of French, which, though minimal, may have been sufficient for (near-) significant correlations to emerge.

At the end of the treatment, a number of significant relationships, some of them of considerable strength, were uncovered. In Group F, language learning aptitude was significantly associated with both metalinguistic awareness and L2 proficiency. In Group E, all variables were significantly associated, with a strong correlation between aptitude and L2 proficiency in particular standing out. Thus, in Group E, the three variables of aptitude, metalinguistic awareness, and L2 proficiency showed no association at pre-test, but a close association at post-test. One can speculate that the interplay between the three variables may be cyclical: improved aptitude and metalinguistic awareness foster L2 proficiency, while growing proficiency in Esperanto likewise promotes the development of metalinguistic awareness as well as language learning aptitude. In other words, there are indications that Esperanto may have had a facilitative effect on those abilities that underlie successful explicit learning. This interpretation is consistent with the fact that Group E made significant gains on all measures in the course of the year and the fact that effect sizes for Group E were generally large.

The fourth finding refers to the predictive value of language learning aptitude in connection with L2 proficiency. Even though it has been established that aptitude was still developing in the children, it is possible to use a measure of aptitude at a particular point in time for diagnostic purposes or to make predictions about a language learner's future achievement (Kiss, 2009; Kiss & Nikolov, 2005; Milton & Alexiou, 2006). In the present study, children's performance on the Number Learning and Matching Words subtests of the MLAT-E(UK) at

pre-test significantly predicted L2 proficiency at post-test, explaining 52% of the variance in scores. In the present study, then, (two components of) aptitude accounted for a much more substantial proportion of the variance in L2 proficiency than reported in a previous study (Kiss & Nikolov, 2005), although it must be borne in mind that the regression analysis was carried out in an exploratory mode, given the relatively small sample size.

Performance on the Number Learning subtest was the stronger predictor, explaining 43%, while performance on the Matching Words subtest explained a further 9% of the variance in L2 proficiency scores. The Number Learning subtest required children to learn and recall numbers in an invented language; the learning task is performed in an aural modality and strongly relies on memory ability. Conversely, the Matching Words subtest required children to match a keyword in an English sentence with a word that plays the same grammatical role in a second English sentence. This task is performed in a written modality and relies on language-analytic ability. Hence, performance on a memory-oriented task and an analytically-oriented task significantly predicted L2 proficiency, with the memory-oriented task the stronger predictor. This finding indicates that the 8-9-year-old participants may still rely quite strongly on memory in their L2 learning, as one would expect in the case of child learners. However, it is likewise clear that analytic abilities are beginning to play a role as well – a finding which is perhaps surprising. Existing research tends to support the view that language-analytic abilities become relevant at around age 12 (Harley & Hart, 1997; Muñoz, 2009), although some researchers have claimed that such abilities may kick in much earlier than that, i.e. at around age 6 or 7 (Milton & Alexiou, 2006). The results of the current study are consistent with this latter claim.

6. Conclusions, limitations, and implications for further research

At the beginning of this paper, we raised the question of whether the development of language learning aptitude and metalinguistic awareness (and, by implication, effective explicit learning) in children aged 8-9 could be enhanced by means of a specific tool: the teaching and learning of Esperanto. We asked whether it might be possible to maximize the L2 learning of

young children in the minimal-input setting of the typical language classroom by accelerating their cognitive development with a view to kick-starting explicit learning processes. The findings of the present study indicate that the answer to this question may be a cautious 'yes'.

We found that in addition to progressing in the L2 they were taught, both treatment groups made significant gains on measures of language learning aptitude and metalinguistic awareness. Furthermore, performance on a memory-based and an analytically-based component of language learning aptitude at pre-test significantly predicted L2 proficiency scores at post-test, suggesting that language-analytic ability and, by implication, explicit learning have a role to play in the L2 learning of 8-9-year-olds.

In respect of the potential benefits of teaching and learning Esperanto, Group E children's performance was characterized by greater homogeneity than the performance of Group F children. In addition, effect sizes in Group E were consistently larger than in Group F for improvements in aptitude, metalinguistic awareness, and L2 proficiency. Learning a regular and transparent language was not only demonstrably easier than learning an L2 such as French, but proficiency in Esperanto was strongly associated with language learning aptitude and metalinguistic awareness at the end of the treatment. While a correlation does not allow us to make claims about cause and effect, it is reasonable to suggest that the three variables were engaged in interplay, mutually and positively influencing one another, fostering connected growth.

It is acknowledged that the present study had its limitations. Due to the practical difficulties associated with setting up a research study in a school classroom, the number of participating children was relatively small and the duration of the treatment was comparatively short. In a more comprehensive study involving larger numbers of children and longer treatments, the trends that emerged in the present study could be corroborated, or otherwise shown to be specific to the sample at hand and not generalizable. It should be the aim of future work to gain further insight into the development of language learning aptitude and metalinguistic awareness in child L2 learners, and to pinpoint the potentially facilitative role of a language

such as Esperanto in the context of classroom L2 learning and teaching characterized by a minimal-input situation.

Notes

[1] Ideally, there would have been a control group of children who did not learn any L2. In practice, ethical considerations prevented us from attempting to set up such a group in the actual school context: it seemed inappropriate to expect young children to complete a number of tests without offering them any language learning opportunity.

[2] In addition, children's attitudes to languages and language learning were assessed by means of a short questionnaire. Overall, treatment group membership and attitudes were statistically independent of each other.

References

Baker, C. (2006). *Foundations of bilingual education and bilingualism* (4th ed.). Clevedon: Mutilingual Matters.

Barton, A., & Bragg, J. (2010). *Springboard to Languages project: Phase 3 report*: University of Manchester.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology, 24*(4), 560-567.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition, 16*, 157-168.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge: Cambridge University Press.

Bialystok, E., & Ryan, E. B. (1985). Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly, 31*(3), 229-251.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning, 56*(s1), 9-49.

Bishop, A. J. (1997). *EKPAROLI project report 1994-1997*. Retrieved 15/07/2011, from http://web.archive.org/web/20040301110024/http://www.education.monash.edu.au/projects/esperanto/Ekrep97.htm.

Cameron, L. (2001). *Teaching language to young learners*. Cambridge: Cambrigde University Press.

Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87-136). New York, NY: John Wiley & Sons.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83-118). Rowley, MA: Newbury House.

Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. S. Parry & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 11-29). Englewood Cliffs, NJ: Prentice Hall.

Carroll, J. B., & Sapon, S. M. (2002a). *Modern Language Aptitude Test - Elementary: MLAT-E*. Rockville, MD: Second Language Testing Foundation.

Carroll, J. B., & Sapon, S. M. (2002b). *Modern Language Aptitude Test: MLAT*. N. Bethesda, MD: Second Language Testing Inc.

Cenoz, J. (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code-mixing. In M. d. P. Garcia Mayo & M. L. Garcia Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 77-93). Clevedon: Multilingual Matters.

Corsetti, R., & LaTorre, M. (1995). Quale lingua prima. *Language Problems and Language Planning, 19*, 26-46.

Cummins, J. (1987). Bilingualism, language proficiency, and metalinguistic development. In P. Homel, M. Palij & D. Aaronson (Eds.), *Childhood bilingualism: Aspects of linguistic, cognitive, and social development* (pp. 57-73). Hillsdale, NJ: Erlbaum.

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Malden, MA: Blackwell.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning, 55*(s1), 1-25.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Erlbaum.

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589-630). Malden, MA: Blackwell.

Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *Modern Language Journal, 79*(1), 67-89.

Ellis, R. (2006). Modelling learning difficulty and second language proficiency: The differential contributions of implicit and explicit knowledge. *Applied Linguistics, 27*(3), 431-463.

Enever, J. (2009). Can today's early language learners in England become tomorrow's plurilingual European citizens? In M. Nikolov (Ed.), *Early learning of modern foreign languages: Processes and outcome* (pp. 15-29). Bristol: Multilingual Matters.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research, 9*(2), 147-171.

European Commission. (2008). *Key data on teaching languages at school in Europe*. Brussels: Education, Audiovisual and Culture Executive Agency P9 Eurodyce.

Fantini, A. E., & Reagan, T. G. (1992). *Esperanto and education: Toward a research agenda*. Retrieved 08/07/2011, from http://dok.esperantic.org/f-r1.htm.

Fisher, A. (1921). Modern languages by way of Esperanto. *Modern Languages, 5/6*, 179-182.

García Mayo, M. d. P. (2003). Age, length of exposure and grammaticality judgements in the acquisition of English as a foreign language. In M. d. P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 94-114). Clevedon: Multilingual Matters.

Gombert, J. E. (1992). *Metalinguistic development*. Hemel Hempstead: Harvester.

Halloran, J. H. (1952). A four year experiment in Esperanto as an introduction to French. *British Journal of Educational Psychology, 22*(3), 200-204.

Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition, 19*, 379-400.

Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics, 30*, 225-249.

Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 539-588). Malden, MA: Blackwell.

Kiss, C. (2009). The role of aptitude in young learners' foreign language learning. In M. Nikolov (Ed.), *The age factor and early language learning* (pp. 253-276). Berlin: Mouton de Gruyter.

Kiss, C., & Nikolov, M. (2005). Developing, piloting, and validating an instrument to measure young learners' aptitude. *Language Learning, 55*(1), 99-150.

Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situaton. *Second Language Research, 24*(1), 35-63.

Markarian, R. (1964). *The educational value of Esperanto teaching in the schools*. Rotterdam: Libroserve Universala Esperanto-Asocio.

Milton, J., & Alexiou, T. (2006). Language aptitude development in young learners. In C. Abello-Contesse, R. Chacón-Beltrán, M. D. López-Chiménez & M. M. Torreblanca-López (Eds.), *Age in L2 acquisition and teaching* (pp. 177-192). Oxford: Peter Lang.

Muñoz, C. (2006). The effects of age on foreign language learning: The BAF project. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 1-40). Clevedon: Multilingual Matters.

Muñoz, C. (2009). Input and long-term effects of early learning in a formal setting. In M. Nikolov (Ed.), *The age factor and early language learning* (pp. 141-159). Berlin: Mouton de Gruyter.

Nicholas, H., & Lightbown, P. M. (2008). Defining child second language acquisition, defining roles for L2 instruction. In J. Philp, R. Oliver & A. Mackey (Eds.), *Second language acquisition and the younger learner: Child's play?* (pp. 27-51). Amsterdam: John Benjamins.

Norris, J. M., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning, 51*(1), 157-213.

Ranta, L. (2002). The role of learners' language analytic ability in the communicative classroom. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 159-180). Amsterdam: John Benjamins.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69-94). Amsterdam: John Benjamins.

Suárez, M. d. M., & Muñoz, C. (2011). Aptitude, age and cognitive development: The MLAT-E in Spanish and Catalan. In L. Roberts, G. Pallotti & C. Bettoni (Eds.), *EuroSLA Yearbook* (Vol. 11, pp. 5-29). Amsterdam: John Benjamins.

Symoens, E. (1989). *The socio-political, educational and cultural roots of Esperanto*. Antwerpen: Internacia Ligo de Esperantistaj Instruistoj.

Williams, N. (1965a). A language teaching experiment. *Canadian Modern Language Review, 22*, 26-28.

Williams, N. (1965b). *Report on the teaching of Esperanto from 1948 to 1964*. Manchester: Denton Egerton Park County School.