Metalinguistic awareness in children with differing language learning experience

Angela Tellier and Karen Roehr-Brackin

University of Essex

Abstract

Theoretical research concerned with the notion of second language (L2) learning difficulty has resulted in specific criteria that can be used to predict the learning difficulty of different languages in terms of both explicit and implicit knowledge. The characteristics of the constructed language Esperanto suggest that this language has lower explicit and implicit learning difficulty than other languages. It may therefore be a suitable 'starter language' for child L2 learning in the classroom. Specifically, we propose that Esperanto may facilitate the development of metalinguistic awareness and, as a consequence, boost children's budding capacity for explicit learning. This would be particularly advantageous in the minimal-input setting of the average foreign language classroom. We present findings from an empirical study which compared 11 to 12-year-old English-speaking children who had learned Esperanto and a European L2 (N = 35) with children who had learned various combinations of European and non-European L2s (N = 168) in terms of their performance on a measure of metalinguistic awareness. No significant differences in overall level of metalinguistic awareness were identified, but the Esperanto group significantly outperformed the comparison group on one of the eleven metalinguistic tasks included in the measure. Moreover, the Esperanto group displayed a more homogeneous performance than the other groups of children. This suggests that learning Esperanto may have a lasting levelling effect, reducing differences between children with varying metalinguistic abilities.

Introduction

This paper has two principal aims. The first aim is to put forward a theoretical argument concerned with the notion of learning difficulty in the context of explicit and implicit second language (L2) learning. To this end, the constructs of explicit and implicit learning as well as metalinguistic ability and awareness are defined and discussed with reference to L2 learning. Subsequently, criteria for assessing the learning difficulty of linguistic constructions and metalinguistic descriptions are applied to the constructed language Esperanto. It is concluded that Esperanto has lower learning difficulty than other European languages. On the basis of this conclusion, we hypothesise that Esperanto may be a suitable 'starter language' for child L2 learning in the classroom. The second aim is to present initial empirical evidence which speaks to this hypothesis. Findings from an empirical study investigating levels of metalinguistic awareness in 11 to 12-year-old English-speaking children (N = 203) are described and discussed. The study compared children who had learned Esperanto and a European L2 with children who had learned various combinations of European and non-European L2s in terms of their performance on a dedicated measure of metalinguistic awareness.

Theoretical background

Explicit and implicit learning, metalinguistic ability and awareness

Most applied linguists agree that adults draw on both implicit and explicit knowledge when learning an L2. Explicit knowledge refers to knowledge that is represented declaratively, can be brought into awareness and can potentially be verbalised. Conversely, implicit knowledge is knowledge that cannot be brought into awareness or articulated (Anderson, 2005; R. Ellis, 2004; Hulstijn, 2005). Researchers likewise agree that in adults, L2 proficiency is achieved through a combination of implicit and explicit learning processes (Dörnyei, 2009; R. Ellis,

2006). Explicit learning refers to situations when a learner consciously and intentionally attempts to master language material or solve language-related problems (Dörnyei, 2009; N. C. Ellis, 1994). By contrast, implicit learning is learning without conscious awareness and is incidental (DeKeyser, 2003; N. C. Ellis, 1994; Hulstijn, 2003).

Child L2 learners, who have not yet reached cognitive maturity, are thought to learn primarily implicitly. Evidence from age-of-onset studies that have investigated child L2 learning in naturalistic settings suggests that children learn very successfully in such environments (e.g. DeKeyser, 2000; Johnson & Newport, 1989), that is, environments which typically offer large amounts of high-quality input over a prolonged period of time. Although children initially learn more slowly than adults, they eventually reach higher levels of proficiency than older learners with a later starting age (see Birdsong, 2006; Hyltenstam & Abrahamsson, 2003). In classroom settings, however, children do less well. Indeed, research looking at the attainment of foreign language learners exposed to classroom-based instruction has shown that later starters consistently outperform younger starters on measures of L2 achievement. In other words, older children and adolescents do better than younger children on a variety of L2 performance measures after the same amount of L2 exposure. This trend is in evidence throughout, and the majority of findings are statistically significant (Cenoz, 2003; García Mayo, 2003; Harley & Hart, 1997; Larson-Hall, 2008; Muñoz, 2008, 2009, 2006).

The reason why younger learners do less well in classroom settings can be found in the quality and above all the quantity of input such a learning environment typically offers, considered in conjunction with the human learning mechanisms that are best placed to make use of such input. Around the world, classroom learning is characterised by small amounts of input provided over a limited period of time, e.g. one or two hours a week during term time, spread over a few years. In other words, typical classroom foreign language learning is neither intensive nor extensive. In such a minimal-input situation, explicit learning is

particularly useful because explicit processes are fast and efficient, unlike implicit learning, which requires considerable exposure to input in order to be successful (DeKeyser, 2003). Explicit learning is also conscious and intentional, however (Dörnyei, 2009); thus, it requires attention, awareness, and effort, relying on the processing of information in working memory (Ashby & Casale, 2005; Bailey, 2005; Reber, 2005). Put differently, explicit learning is cognitively demanding. Given the demands of explicit learning, it is unsurprising that cognitively more mature learners (e.g. adolescents) are better able to cope than cognitively less mature learners (e.g. young children). Cognitively mature learners can therefore benefit considerably even from minimal L2 input, while younger children tend to benefit less, since their cognitive abilities are not fully developed yet.

At this point it is worth noting that some existing research suggests that younger children may also draw on explicit knowledge and learning (Barton & Bragg, 2010; Milton & Alexiou, 2006; Roehr, 2012a), though certainly to a lesser extent than adults. It appears that the metalinguistic abilities which are used during explicit learning develop most visibly from around age 6 or 7, in parallel with the onset of literacy in children who are exposed to schooling (Birdsong, 1989; Karmiloff & Karmiloff-Smith, 2002). Metalinguistic ability refers to "the capacity to use knowledge about language" (Bialystok, 2001: 124). It is closely related to metalinguistic awareness, which "implies that attention is actively focused on the domain of knowledge that describes the explicit properties of language" (Bialystok, 2001: 127).[1] Put differently, metalinguistic awareness refers to an awareness of the nature, function, and form of language; metalinguistically aware individuals can treat language as an object of inspection, reflection, and analysis (Baker, 2006; Birdsong, 1989; Cummins, 1987; Gombert, 1992).

Metalinguistic ability can be understood in terms of the development of analysed knowledge and cognitive control (Bialystok, 1988, 1994a, 1994b, 2001; Bialystok & Ryan, 1985).

Analysis of knowledge and cognitive control are situated on a continuum; they gradually emerge as children mature. Analysis of knowledge refers to how knowledge is structured; more analysed linguistic and conceptual representations are more explicit, more abstract, and more accessible to introspection than less analysed representations. Through analysis of knowledge, the "basic categories of language and thought" may be uncovered (Bialystok, 1994b: 561). Cognitive control of processing refers to executive function in working memory (Bialystok & Craik, 2010; Bialystok, Craik, & Luk, 2012); it enables the individual "to selectively attend to specific aspects of a representation" (Bialystok, 2001: 131), to integrate information, and to inhibit irrelevant information (Bialystok, 1988). Though theoretically independent, analysis and control are interrelated in practice in the sense that they develop interactively as individuals mature. Increasingly developed analysis of knowledge and control of processing allow children to gradually move towards the use of more complex linguistic skills, that is, to progress from simple conversation to skills such as reading and metalinguistic problem-solving (Bialystok, 2001; Bialystok & Ryan, 1985). Metalinguistic tasks typically make high demands on both analysis of knowledge and control of processing (Bialystok, 1988), requiring individuals to draw on and selectively attend to specific aspects of analysed representations.

Existing research suggests that children who grow up as balanced bilinguals typically outperform monolingual peers of the same age on cognitive tasks requiring high levels of analysis of knowledge and/or control of processing (e.g. Bialystok, 1988; Bialystok & Martin, 2004). In this sense, bilingualism appears to confer cognitive advantages on young children. Comparable research with multilingual children is still in short supply; the evidence that is available to date indicates that trilinguals have no additional advantages over bilinguals in terms of cognitive control, but instead perform at similar levels (Poarch & van Hell, 2012).

Applied linguistics researchers have highlighted the cognitive advantages resulting from the learning of multiple L2s in learners of all ages. It has been proposed, for instance, that multilinguals have particularly well-developed self-monitoring skills, enhanced sociolinguistic awareness, as well as superior metalinguistic abilities that can be brought to bear during performance on both metalinguistic and linguistic tasks, e.g. the ability to draw comparisons between different languages (Jessner, 1999, 2006, 2008). In a study of metalinguistic knowledge in university-level L2 learners, it was found that language learning experience, operationalised as years of formal study of the L2 under investigation (German or Spanish) together with cumulative years of study of other L2s accounted for an impressive 45% of the variance in participants' level of metalinguistic knowledge as measured by a dedicated test (Roehr & Gánem-Gutiérrez, 2009b). In other words, language learning experience in terms of both length of L2 exposure and number of L2s studied was a powerful predictor of metalinguistic ability. To date, to the best of our knowledge, there is no comparable empirical research which has examined by means of a dedicated measure metalinguistic knowledge/ability/awareness in child learners who have had minimal L2 exposure in a foreign language classroom setting.

Indeed, as the use of metalinguistic abilities is cognitively demanding (Cummins, 1987), requiring high levels of both analysis of knowledge and control of processing, heightened metalinguistic awareness is typically associated with either bilingualism from birth (in studies with naturalistic child learners) or higher levels of cognitive development and, by implication, greater cognitive maturity (in studies with classroom L2 learners). If, however, young children's budding metalinguistic awareness and their developing capacity to learn explicitly could be enhanced, their classroom-based L2 learning could potentially be made more successful. In other words, children who are better able to learn explicitly at an early

age would be better able to benefit even from minimal L2 input, that is, the kind of input available in most foreign language classrooms.

Learning a language that lends itself especially well to metalinguistic inspection, reflection, and analysis may help sharpen a learner's metalinguistic abilities. In this way, it may help accelerate the development of explicit learning capacity. This line of argument led us to the following hypothesis: Through learning an 'easy' language, the abilities that facilitate learning other, more 'difficult' languages in a minimal-input environment might be fostered particularly effectively [2].

Learning difficulty

The constructed language Esperanto meets many of the criteria that have been associated with low learning difficulty, in terms of both implicit and explicit knowledge. Existing recent research has sought to identify characteristics that can help account for learning difficulty (e.g. Collins, Trofimovich, White, Cardoso, & Horst, 2009; DeKeyser, 2005; Housen, Pierrard, & Van Daele, 2005; Spada & Tomita, 2010). In an instructed L2 learning context, the characteristics of specific linguistic constructions can be used to predict implicit learning difficulty, while the characteristics of metalinguistic descriptions can be used to predict explicit learning difficulty (R. Ellis, 2006; Roehr & Gánem-Gutiérrez, 2009a). Metalinguistic descriptions in the sense of pedagogical grammar rules are tied to the particular linguistic constructions they describe.

To illustrate with reference to a recently proposed taxonomy (Roehr & Gánem-Gutiérrez, 2009a), it can be argued that linguistic constructions have low implicit learning difficulty if they are characterised by transparent form-meaning mappings, high perceptual salience, and low communicative redundancy. Transparent form-meaning mappings refer to the association of one form with one meaning only, and vice versa. High perceptual salience means that a

construction can be perceived easily in (auditory) input. Communicative redundancy characterises the communicative value of a construction; if redundancy is high, inaccurate use of the construction is unlikely to result in misunderstanding on the part of the interlocutor. As a consequence, erroneous use may go unnoticed and learning difficulty is increased (see also Goldschneider & DeKeyser, 2001). Along similar lines, it can be argued that metalinguistic descriptions have low explicit learning difficulty if they are characterised by low conceptual complexity and high truth value. Conceptual complexity is defined by the number of categories and relations included in a pedagogical grammar rule; a rule with low conceptual complexity draws on few categories and relations and is thus easier to process explicitly. Truth value refers to the extent to which a pedagogical grammar rule has exceptions; high truth value means that there are no or only very few exceptions to the rule; this results in low explicit learning difficulty (for a full discussion, see Roehr, 2012b; Roehr & Gánem-Gutiérrez, 2009a) [3].

Esperanto is a constructed language that was designed to be easy to learn and use, and the linguistic constructions that constitute Esperanto do indeed satisfy the criteria of low implicit learning difficulty mentioned above to a greater extent than most – or perhaps all – other European languages. In addition, metalinguistic descriptions that satisfy the criteria of low explicit learning difficulty are likewise more readily available for Esperanto than for most – or all – other European languages. Unlike most languages, Esperanto has highly regular and transparent morphology and syntax, and it is characterised by direct phoneme-grapheme correspondence. The language has 16 key rules of grammar and a transparent morphological system. In lexical terms, Esperanto draws on the main European languages, with a particularly heavy influence from the Romance languages.

As a matter of fact, a recent study comparing the classroom-based learning of Esperanto with the learning of French in 8 to 9-year old English-speaking children over a school year showed

that in the given minimal-input setting Esperanto was significantly easier to learn than French, with larger gains in L2 proficiency achieved by the Esperanto group compared with the French group (Tellier & Roehr-Brackin, 2013).

In the context of classroom-based child L2 learning in the 21st century, it is also worth noting that Esperanto is arguably more 'egalitarian' a tool than a language like Latin which has traditionally been regarded as a suitable tool for raising language learners' metalinguistic awareness (e.g. Sparks, Ganschow, Fluharty, & Little, 1995-6). Unlike Latin, Esperanto is a continually evolving language with third- and fourth-generation speakers. It therefore lends itself more naturally to a focus on oral and aural skills and the avoidance of heavy reliance on fully developed literacy skills that may not be available to younger children and to children of lower ability.

Esperanto as a 'starter language'

The idea that Esperanto may serve as a preparatory tool for subsequent L2 learning in the classroom is not new. In the early 20th century, academics began to speculate on the possible pedagogical advantages of Esperanto, if taught and learned before or alongside other languages (Lodge, 2004/1905; see also Masson, 2006). Benefits such as heightened metalinguistic awareness, more positive attitudes to language learning, improved L1 literacy, and greater self-esteem were predicted (Corsetti & LaTorre, 1995; Fantini & Reagan, 1992; Markarian, 1964; Symoens, 1989). However, there has been surprisingly little empirical research to date which has put these hypotheses to the test. Early work by Fisher (1921) and Halloran (1952) conducted in Britain reported that children who learned Esperanto for a year achieved higher results after four years of learning French than children who had studied French for five years. Similarly, Williams (1965a, 1965b), over an 18-year period, found that secondary school-age children performed better in French if they had learned Esperanto for a

year first. Like Halloran, Williams posited that the effects of learning Esperanto might be strongest in children whose verbal intelligence scores were low. He warned, however, that his conclusions were impressionistic (Williams, 1965a, 1965b).

It is indeed worth noting that most early studies on the teaching and learning of Esperanto in schools suffered from a number of shortcomings, if considered by today's standards. Not only is reporting often brief or anecdotal, but studies were conducted in a wide range of educational contexts which cannot be compared easily (Symoens, 1992), including selective grammar schools (Halloran, 1952) and non-selective secondary modern schools (Williams, 1965b), elementary schools (Formaggio, 1990) and secondary-level schools (Thorndike & Kennon, 1927). Researchers reviewing early work from today's perspective argue that studies often had poorly defined aims (Fantini & Reagan, 1992) or were somewhat superficial (Corsetti & LaTorre, 1995). Moreover, several of the early studies referred to above employed an experimental research design, but lacked clear objectives or methodological rigour, thereby compromising findings. For instance, modern reviewers agree that experimental and control groups were not always comparable (Maxwell, 1988).

More recent studies are few and far between. Bishop (1997), in his research in the Australian school context, gathered teacher ratings of pupil performance. Participating secondary-school teachers considered language students who had been exposed to Esperanto in primary school as more motivated than language students who had not learned Esperanto prior to commencing secondary school. The Esperanto students' L2 speaking skills and their overall L2 achievement were also rated more highly by the teachers when compared with the ratings allocated to other students. The participating teachers did not know which students had and which students had not learned Esperanto (Bishop, 1997). Results from a case study

examining a curriculum initiative that used Esperanto as a tool for raising language awareness in primary-school children in Britain are likewise encouraging (Barton & Bragg, 2010). The researchers report that the 7 to 8-year-old pupils had evidently enjoyed the Esperanto programme and were motivated to learn languages in the future. Moreover, the children "demonstrated high levels of skill in those tasks which required them to translate unfamiliar languages" (Barton & Bragg, 2010: 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, while theoretical considerations lead to positive conclusions, up-to-date empirical research into the potential of Esperanto as a 'starter language' is still scarce. Hence, there is as yet little concrete evidence available as to whether the teaching and learning of Esperanto really does convey any cognitive benefits that could result in enhanced subsequent (explicit) L2 learning in the classroom. The empirical study reported in the following was conducted to begin to address this issue.

Empirical study

In order to take a first step towards addressing the hypothesis that learning Esperanto prior to learning other languages may foster metalinguistic awareness in children and may thus contribute to (accelerated) development of the capacity for explicit learning, an empirical study was conducted. The study had a cross-sectional design and was aimed at comparing children's performance on a dedicated test of metalinguistic awareness following differential L2 input. The following research question was formulated:

What are the long-term effects of being exposed to Esperanto and a European L2, compared with being exposed to European and non-European L2s, on children's metalinguistic awareness?

We hypothesised that children who had been exposed to Esperanto and a European L2 would outperform children who had been exposed to European and non-European L2s on the measure of metalinguistic awareness that was used in the study.

Participants

An intact cohort of children in Year 7 of an English state secondary school [4] (N = 225) was recruited for participation in the study. Year 7 is the first year of secondary school in the English education system. All children were 11 to 12 years old. Children who reported to be bilingual in English and another language (N = 15) were excluded from the sample because their bilingualism was likely to confer advantages in metalinguistic ability that would have skewed the results. For the purpose of the present study, bilingualism was defined as speaking a language other than English at home whilst coping without difficulty with normal classroom work in English. Children who reported having spent more than three months in a non-English-speaking country, the U.S., or South Africa (N = 7) were also removed from the sample since it could not be assumed that these children had not developed a certain level of bilingualism through naturalistic exposure [5]. This resulted in a final sample of 203 participants.

In Year 7, all children were exposed to classroom instruction in French. In Years 3 to 6 of primary school, the children had learned different combinations of languages, depending on the primary schools they had attended. A total of 39 children additionally reported attendance of extra-curricular language clubs at some point during primary school. Languages they were exposed to included French, German, Italian, Japanese and Spanish. Typically, children could

not remember precisely when or for how long they had attended a language club (e.g. for an entire school year or less), how old they were at the time, or indeed what the session length was (e.g. 30 minutes or less/more). Therefore, no further participants were excluded from the sample due to reporting language club attendance.

For the purposes of data analysis, the children were divided into seven groups, based on the primary schools they had previously attended and, thus, the prior L2 input they had received in the context of curriculum-based language instruction. Group 1 (N = 35) had learned Esperanto and a European L2. Groups 2-7 (N = 168) had learned different combinations of European and non-European L2s, but not Esperanto. Groups 1 to 6 were intact classes from six different primary schools; Group 7 comprised children from 20 other primaries. The children's prior L2 exposure in primary school is summarised in Table 1.

Table 1. Children's prior L2 input in primary school (Years 3-6)

Group	**N**	**Languages**
1	35	Esperanto in Y3 + Y4, Spanish in Y5 + Y6
2	35	French in Y5 + Y6
3	32	French in Y3, Y4, Y5, French and Latin in Y6
4	18	French in Y3 + Y4, language taster programme (German, Japanese, Latin) in Y5, Spanish in Y6
5	23	French in Y3 + Y4, French and Japanese in Y5 + Y6
6	12	French in Y3 + Y4, French and Spanish in Y5 + Y6
7	48	A range of languages or a single language; some children had input over 2 years, some over 4 years

Note: Y = Year

Test of metalinguistic awareness

All children completed a test of metalinguistic awareness which had been designed for group administration to mixed-ability classes of English-speaking children aged 8 to 11. Although the sample in the present study included 12-year-olds, there was no ceiling effect (see results

section below). The test is a paper-and-pen measure comprising eleven tasks; the maximum possible score for the test as a whole is 88. The test covers domains relevant to both L1 and L2 learning, i.e. lexical semantics, morphology and syntax, ambiguity, and basic metalinguistic terminology. The test addresses concepts specific to the L2 learning of English-speaking children, who typically encounter other European languages in the language classroom. The concepts in question include grammatical gender, case, verbal and adjectival agreement, cognates, and similarities and differences between languages. Translation also features in the test; translation can be considered a specific metalinguistic skill that requires deliberate analytic comparison of two languages. The translation tasks were designed to encourage children to employ knowledge of any language(s) they had (including their L1) and make use of opportunities for positive transfer.

Section 1 of the test (Tasks 1-5) is based on a number of European languages including Esperanto; Section 2 (Tasks 6-11) is based on a constructed language designed for the purpose of the test. Tasks 1-4 assess children's ability to make comparisons between different L2s via cognate recognition and/or translation; Task 5 asks children to match a syntactically ambiguous English sentence with appropriate pictures illustrating the meaning expressed by the sentence. Task 6 tests children's understanding of metalinguistic terminology, focusing on parts of speech. Task 7 deals with accusative case marking, while Task 8 addresses children's understanding of a syntactic rule pertaining to word order. Task 9 requires children to spot the common features in lists of words and create two more words which could belong to the same word class. Task 10 focuses on subject-verb agreement, while Task 11 deals with grammatical gender and gender marking. Example tasks are shown in Appendix A (Task 1) and Appendix B (Task 8). The test had been piloted extensively prior to its use in the current study. In Years 3-5 of primary school, i.e. several years prior to the current study, children in Group 1 had been exposed to earlier, much simpler and shorter versions of Tasks 1, 3 and 4

as well as one of the sentences from Task 2 when those tasks were first trialled [6]. For results of the pilot studies as well as detailed information regarding the reliability and validity of the measure, the reader is referred to Tellier (2013).

The test was administered by the children's respective class teachers during regular class time in June of the school year in question, i.e. after all children had been exposed to French, their current L2, for eight months. All children completed the test on the same day and finished within an hour. The test papers were collected and scored by the first author according to a prepared scoring key (for details, see Tellier, 2013). In the present study, the test proved to be highly reliable with a Cronbach's alpha coefficient of 0.91.

Results

Table 2 shows the descriptive statistics in terms of overall test performance for Group 1 and Groups 2-7 taken together. Recall that Group 1 comprises children who had been exposed to Esperanto and another European L2 in primary school, while children in Groups 2-7 had been exposed to different combinations of European and non-European L2s.

Table 2. Descriptive statistics: Test of metalinguistic awareness

Test total	**Group 1 (N = 35)**	**Groups 2-7 (N = 168)**
Mean % correct	50.6%	51.2%
Mean	44.54	45.04
SD	9.02	13.34
Min.	24	13
Max.	62	71
Section 1		
Mean % correct	50.7%	48.5%
Mean	28.91	27.65
SD	6.19	7.11
Min.	16	9
Max.	43	46
Section 2		

Mean % correct	50.4%	56.1%
Mean	15.63	17.39
SD	6.68	7.72
Min.	0	0
Max.	27	31

The descriptive statistics in Table 2 suggest that the test as a whole was fairly challenging for the participants with mean facility values of around 50% throughout. The scores achieved by Group 1 and Groups 2-7 taken together are rather similar, both for the test as a whole and for Section 1 and Section 2 individually. Independent samples t-tests confirmed that there were no statistical differences (Test total: $t(201) = 0.271$, $p = 0.79$, two-tailed; Section 1: $t(201) = -0.973$, $p = 0.33$, two-tailed; Section 2: $t(201) = 1.253$, $p = 0.21$, two-tailed). Our hypothesis that children who had been exposed to Esperanto and a European L2 would outperform children who had been exposed to European and non-European L2s on the measure of metalinguistic awareness was thus not supported.

In order to obtain more detailed information about children's performance on the test of metalinguistic awareness, analyses based on the eleven individual tasks as well as analyses considering the seven participant groups separately were carried out. Table 3 shows the descriptive statistics for the eleven individual tasks constituting the test of metalinguistic awareness.

Table 3. Descriptive statistics: Individual metalinguistic awareness tasks

		Group 1 (N = 35)	**Groups 2-7 (N = 168)**
Task 1	Mean % correct	74.6%	52.9%
	Mean	6.71	4.76
	SD	2.257	2.033
Task 2	Mean % correct	24.4%	27.4%
	Mean	6.60	7.40
	SD	3.875	4.240

Task 3	Mean % correct	76.1%	77.1%
	Mean	6.09	6.17
	SD	1.961	1.807
Task 4	Mean % correct	75.2%	74.9%
	Mean	6.77	6.74
	SD	1.784	2.010
Task 5	Mean % correct	68.5%	64.5%
	Mean	2.74	2.58
	SD	0.561	0.777
Task 6	Mean % correct	49.8%	55.2%
	Mean	2.49	2.76
	SD	1.755	1.738
Task 7	Mean % correct	31.5%	27.5%
	Mean	0.63	0.55
	SD	0.910	0.867
Task 8	Mean % correct	57.8%	65.8%
	Mean	2.31	2.63
	SD	1.471	1.446
Task 9	Mean % correct	53.6%	59%
	Mean	6.43	7.08
	SD	3.806	3.814
Task 10	Mean % correct	48.7%	60.3%
	Mean	1.46	1.81
	SD	1.146	1.158
Task 11	Mean % correct	46.2%	51.2%
	Mean	2.31	2.56
	SD	1.778	1.801

The scores shown in Table 3 indicate that Tasks 3 and 4 were generally easiest, while Tasks 2 and 7 were generally very challenging for the children. Task 3 requires the identification and application of a morphological rule focusing on singular and plural nouns in four different languages (French, Irish, Romanian, Spanish); Task 4 targets recognition of cognates among simple and compound nouns in four different languages (English, Esperanto, French, German). Clearly, children were able to handle these noun-based tasks well, achieving a mean correctness score of around 75% throughout. By contrast, Tasks 2 and 7 had facility

values around or even below the 30% mark, thus proving very taxing. Task 2 targets cognate recognition and translation, presenting children with three fairly long sentences in three languages (Esperanto, German, Italian) which they are asked to translate into L1 English. The length of the sentences to be translated (11-14 words) may have proved too much of a challenge for many children. Task 7 requires the identification and application of a morphological rule focusing on case marking in the constructed language that is the basis of Section 2 of the test. Case marking is not obvious in L1 English, so this may have contributed to the difficulty of this task.

Comparing between groups, Table 3 suggests that on most tasks, Group 1 and Groups 2-7 taken together showed broadly similar performance patterns. Overall, Groups 2-7 attained higher percentage scores than Group 1 on seven of the eleven tasks, while Group 1 outperformed Groups 2-7 on four tasks. Specifically, Group 1 achieved a much higher score than Groups 2-7 on Task 1. Conversely, Groups 2-7 achieved a noticeably higher score than Group 1 on Task 10. An independent samples t-tests confirmed that Group 1 significantly outperformed Groups 2-7 taken together on Task 1 ($t(201) = -5.069$, $p < 0.01$, two-tailed). There were no other statistical differences. Task 1 focuses on cognate recognition and translation. Specifically, it presents children with seven short sentences in seven European languages (Dutch, English, Esperanto, Italian, Portuguese, Romanian, Spanish); children are asked to match three pairs of sentences which have the same meaning and translate the remaining sentence into English (see Appendix A).

Table 4 shows the descriptive statistics for the metalinguistic awareness test as a whole for the seven individual participant groups.

Table 4. Descriptive statistics: Individual groups' performance on the test of metalinguistic awareness

	Group 1	**Group 2**	**Group 3**	**Group 4**	**Group 5**	**Group 6**	**Group 7**
N	35	35	32	18	23	12	48
Mean	44.54	43.23	49.56	37.72	45.83	49.25	44.67
SD	9.02	12.32	13.86	12.07	12.25	15.63	13.18
Min.	24	13	20	16	28	20	16
Max.	62	66	71	55	63	68	68

Despite the apparently quite different means and standard deviations exhibited by some of the groups, a one-way ANOVA revealed that there were in fact no statistical differences between groups, although the result was approaching significance ($F(6, 196) = 2.114$, $p = 0.053$) [7]. It is worth noting that Group 1 showed by far the lowest standard deviation of all the groups ($SD = 9.02$). In other words, performance on the test of metalinguistic awareness was more homogeneous in Group 1 than in any of the other groups.

An analysis taking into account the performance by individual participant groups on the eleven individual metalinguistic tasks yielded significant results for Task 1 ($F(6, 196) = 5.317$, $p < 0.01$), Task 2 ($F(6, 196) = 2.342$, $p = 0.03$) and Task 8 ($F(6, 196) = 3.475$, $p < 0.01$). A post-hoc Tukey test revealed that on Task 1, Group 1 significantly outperformed Groups 2, 4, and 7. This reflects the result reported above which showed that Group 1 outperformed Groups 2-7 taken together on this task. Post-hoc Tukey tests further revealed that on Task 2, Group 4 was significantly outperformed by Group 3, and on Task 8, Group 4 was significantly outperformed by Groups 3, 5, 6, and 7. This pattern of results indicates that Group 4 showed a rather weak performance overall, as can also be seen in the descriptive statistics in Table 4. Scrutiny of publicly accessible performance tables for schools revealed that Group 4 included the lowest-performing children in terms of expected standards in

English and Mathematics in Year 6. In addition, Group 4 was exposed to the greatest number of L2s in primary school (see Table 1).

In order to establish whether number of L2s learned in primary school had any effect on children's performance, the participants were re-grouped to allow for a comparison between children who had had four years of primary-school exposure to one L2 only, i.e. French (N = 14), with children who had had four years of primary-school exposure to more than one L2 (N = 134). Children who had had fewer than four years of primary-school exposure (N = 55) were excluded from this analysis to avoid confounding number of languages with length of exposure. The descriptive statistics for the re-grouped participants are shown in Table 5.

Table 5. Descriptive statistics: Children with a single L2 vs. children with multiple L2s

		Single L2 (N = 14)	**Multiple L2s (N = 134)**
Test total	Mean	48.36	45.61
	SD	12.858	12.786
Section 1	Mean	28.50	28.43
	SD	6.478	6.984
Section 2	Mean	19.86	17.19
	SD	7.492	7.840
Task 1	Mean	5.07	5.37
	SD	1.730	2.280
Task 2	Mean	6.57	7.57
	SD	3.736	4.236
Task 3	Mean	6.57	6.18
	SD	1.651	1.699
Task 4	Mean	7.57	6.63
	SD	1.158	2.032
Task 5	Mean	2.71	2.67
	SD	0.726	0.744
Task 6	Mean	2.86	2.67
	SD	1.834	1.768
Task 7	Mean	0.86	0.54
	SD	1.027	0.864
Task 8	Mean	2.71	2.57

	SD	1.267	1.514
Task 9	Mean	8.64	6.98
	SD	3.054	3.867
Task 10	Mean	1.79	1.77
	SD	1.251	1.176
Task 11	Mean	3.00	2.65
	SD	1.664	1.795

The descriptive statistics show that children who had been exposed to a single L2 tended to do slightly better on the test of metalinguistic awareness than children who had been exposed to more than one L2. This applies to the test as a whole, the two test sections, and eight out of the eleven individual tasks. Independent samples t-tests revealed that the differences in performance were not significant for the test as a whole ($t(146) = 0.764$, $p = 0.45$, two-tailed) or the two test sections (Section 1: $t(146) = 0.38$, $p = 0.97$, two-tailed; Section 2: $t(146) = 1.218$, $p = 0.23$, two-tailed). However, the group exposed to a single L2 significantly outperformed the multiple-L2 group on Task 4 of the metalinguistic awareness test ($t(146) = 2.634$, $p = 0.02$, two-tailed). Task 4 requires the recognition of cognates among simple and compound nouns in four different languages (English, Esperanto, French, German); it was one of the easiest tasks on the test for the sample as a whole.

Finally, in order to establish whether length of prior L2 exposure had any role to play with regard to performance on the test of metalinguistic awareness, we compared children who had had two years of L2 input in primary school ($N = 51$) with children who had had four years of L2 input in primary school ($N = 148$). Children who had had less than two years of L2 input ($N = 4$) were excluded from the analysis. The descriptive statistics for the re-grouped participants are shown in Table 6.

Table 6. Descriptive statistics: Children with two years of prior L2 input vs. children with four years of prior L2 input

		Two years of L2 input (N = 51)	**Four years of L2 input (N = 148)**
Test total	Mean	42.51	45.87
	SD	12.532	12.774
Section 1	Mean	26.49	28.43
	SD	7.095	6.917
Section 2	Mean	16.02	17.44
	SD	6.769	7.822
Task 1	Mean	4.45	5.34
	SD	1.993	2.231
Task 2	Mean	6.80	7.47
	SD	4.247	4.190
Task 3	Mean	5.98	6.22
	SD	2.214	1.692
Task 4	Mean	6.80	6.72
	SD	1.990	1.982
Task 5	Mean	2.45	2.68
	SD	0.673	0.740
Task 6	Mean	2.82	2.69
	SD	1.717	1.768
Task 7	Mean	0.57	0.57
	SD	0.878	0.881
Task 8	Mean	2.59	2.59
	SD	1.329	1.489
Task 9	Mean	6.45	7.14
	SD	3.711	3.820
Task 10	Mean	1.63	1.77
	SD	1.131	1.179
Task 11	Mean	1.96	2.68
	SD	1.697	1.781

The descriptive statistics in Table 6 show that children with shorter L2 exposure tended to obtain lower scores, but independent samples t-tests revealed that there were no significant

differences in performance on the metalinguistic awareness test as a whole ($t(197) = -1.629$, $p = 0.11$, two-tailed) or the two test sections (Section 1: $t(197) = -1.718$, $p = 0.09$, two-tailed; Section 2: $t(197) = -1.155$, $p = 0.25$, two-tailed). Looking at individual tasks, it was found that children with four years of prior L2 input significantly outperformed children with two years of prior L2 input on Task 1 ($t(197) = -2.533$, $p = 0.01$, two-tailed) and Task 11 ($t(197) = -2.526$, $p = 0.01$, two-tailed). Task 1 focuses on cognate recognition and translation, as reported above. Task 11 is the most demanding task on the test as judged by the test designer (Tellier, 2013), requiring the identification and application of a morphological rule pertaining to grammatical gender in the constructed language that is the basis of Section 2 of the test.

Discussion and conclusions

Bringing together the various results, the main findings of the empirical component of the present paper can be summarised as follows. First and foremost, there were no significant differences in performance on the test of metalinguistic awareness between Group 1 and Groups 2-7 taken together. Thus, our hypothesis that children who had been exposed to Esperanto and a European L2 would outperform children who had been exposed to European and non-European L2s on the measure of metalinguistic awareness was not supported. It appears, then, that exposure to Esperanto did not convey any (long-term) advantages in terms of metalinguistic ability that could not be attained equally well via input in other L2s.

This cautious conclusion in support of a null hypothesis can be complemented by a slightly different interpretation, however. In principle at least, it is possible that the more recent input in Spanish (Y5 + Y6) and French (Y7) may have obscured any advantages the children in Group 1 might have gained; in other words, almost three years had passed since the children were last exposed to the potentially facilitative input in Esperanto. Two further results seem to add weight to this alternative interpretation. It was found that Group 1 significantly

outperformed Groups 2-7 taken together on one of the eleven metalinguistic tasks, i.e. Task 1. In addition, and arguably more importantly, Group 1 exhibited by far the lowest SD of all groups for the metalinguistic awareness test as a whole. This finding suggests that having been exposed to Esperanto for two years may have had a lasting levelling effect, making children of different abilities more equal in terms of metalinguistic awareness (see also Tellier & Roehr-Brackin, 2013). A possible explanation for this can be found in the rationale for our original hypothesis: learning a language with low learning difficulty such as Esperanto may help foster children's developing capacity for explicit learning. If morphosyntactic, phonological, and orthographic regularities can be identified readily and if lexical analogies can be spotted easily, children may be prompted to engage in deliberate analysis and metalinguistic reflection of their own accord, taking the initiative as they realise that they can make interesting discoveries about the language. Reliance on teacher support would be less important; children would potentially gain in confidence and may be more motivated to engage with both linguistic and metalinguistic tasks – a motivation triggered by feelings of success ('I can do it!'), greater self-efficacy, cognitive control, insight ('I understand!'), and possibly even enjoyment ('This is fun!'). Thus, through a 'gentle' introduction into metalinguistic thought and problem-solving, children's willingness and ability to analyse language and treat it as an object of reflection may be honed, resulting in cognitive as well as affective benefits.

What is more, if it is the case that learning Esperanto is beneficial for lower-ability children (Halloran, 1952; Williams, 1965a, 1965b), allowing them to catch up with their more able peers, this would clearly be a desirable effect from an educational perspective. It would of course be undesirable if the opposite scenario applied, that is, if the learning of Esperanto slowed down more able children. However, there is no evidence in our results pointing in this

direction, since Group 1 did not do significantly worse than any of the other groups. In other words, our results do not suggest a debilitating influence of Esperanto input.

The second main finding worthy of discussion refers to the role of prior exposure to a single L2 vs. multiple L2s in relation to level of metalinguistic awareness. Perhaps surprisingly, children who had been exposed to a single L2 showed a slight trend for better performance on the test of metalinguistic awareness compared with their multiple-L2 peers. Differences were not statistical except for performance on one of the eleven metalinguistic tasks, but nonetheless this trend was somewhat unexpected. In a study with university-level learners, the cumulative years of study of L2s not targeted (as opposed to the L2s targeted, German and Spanish) was a significant predictor of level of metalinguistic knowledge, accounting for no less than 26% of the variance (Roehr & Gánem-Gutiérrez, 2009b), although it must be borne in mind that this variable was in fact a combination of number of L2s learned and length of exposure. The literature on multilingualism emphasises the benefits of learning multiple L2s (e.g. Jessner, 1999, 2006, 2008). By contrast, a recent investigation into the effects of bilingualism and multilingualism on executive function in 5 to 8-year-old children showed no differences between bilinguals and trilinguals (Poarch & van Hell, 2012), although it should be noted that the bilingual children had grown up bilingually from birth, while the trilingual children had grown up with dual language exposure and had been immersed in their L3 outside the home. The study also included a group of L2 learners, who generally performed less well than the bilingual and trilingual children, with some significant differences in evidence. However, even the L2 learner group had had immersion experience, quite unlike the participants in our study, who had had much less intensive L2 exposure.

Overall, then, the results of the present study suggest that exposure to multiple L2s conveyed no benefits in terms of metalinguistic awareness to 11 to 12-year-old children who had had classroom-based L2 input for nearly five years. A possible interpretation of this finding is

that greater depth of knowledge of a single L2 may be more important than superficial knowledge of a number of L2s. As the quantity of L2 input in the current English primary-school system is truly minimal, typically amounting to no more than 25 hours per school year, one would expect that it will take children several years to obtain anything approaching a working knowledge of an L2. The threshold hypothesis of bilingualism states that a certain level of L2 proficiency needs to be achieved before any positive cognitive effects can be expected (Cummins, 1987). This classic hypothesis in the field of L2 education is supported by current research in experimental child psychology, with Poarch and van Hell concluding that "being subjected to a second language at an early age is insufficient to fully accrue the cognitive control advantages found in bilinguals. Evidently, a specific threshold in exposure and usage needs to be reached before such advantages (…) take full effect" (2012: 549). Our result appears to be entirely in keeping with this (for an opposing view, see Yelland, Pollard, & Mercuri, 1993).

The third main finding of relevance arising from the present study concerns the role of length of L2 exposure in relation to level of metalinguistic awareness. Overall, children with four years of L2 input in primary school tended to obtain higher scores on the test of metalinguistic awareness than children with two years of L2 input. There were no significant differences between the two groups of children, however, except for two of the eleven metalinguistic tasks included in the test. In general, this result is consistent with existing research that has identified benefits for performance on measures of metalinguistic knowledge arising from prolonged language learning experience (Roehr & Gánem-Gutiérrez, 2009b). Yet it is possible to use our result to support an opposing stance. If prolonged experience is so important, why is it that so few significant differences were identified for the children who had had four instead of only two years of primary-school exposure to an L2? This question might be answered by another appeal to the threshold hypothesis of

bilingualism (Cummins, 1987), which appears to apply in the minimal-input setting of the English state school system: four years of L2 lessons are only marginally better than two years of L2 lessons since the cumulative input is still only minimal, especially if we bear in mind that most of the children were exposed to more than one L2 and could thus only acquire superficial knowledge at best during their primary-school years.

A final finding arising from the present study is the rather complex picture that presents itself with regard to significant differences in performance on individual metalinguistic tasks. The individual tasks that resulted in statistical differences in the context of our various analyses were Task 1, Task 2, Task 4, Task 8, and Task 11. These tasks do not have any obvious commonalities and showed quite different facility values, although it is perhaps worth noting that all the tasks from Section 1 (Tasks 1, 2, 4) focus on cognate recognition across a number of European languages, while Task 1 and Task 2 additionally include translation. Task 8 and Task 11 are from Section 2 of the test, which is based on a constructed language. Task 8 tests children's understanding of word order, requiring the identification and application of a syntactic rule (see Appendix B). Task 11 requires the identification and application of a morphological rule pertaining to grammatical gender.

In terms of yielding significant results, children who had been exposed to Esperanto set themselves apart on Task 1, while children with four years of prior L2 input outperformed children with two years of prior L2 input on Tasks 1 and 11. Thus, learning Esperanto may help with cognate recognition and translation in the context of European languages, whereas a greater amount of prior L2 input may have the same effect while additionally also helping children to come to terms with the concept of grammatical gender which has no relevance in English.

The relatively weak Group 4 was outperformed by one other group on Task 4 and by several other groups on Task 8, while the single-L2 group outperformed the multiple-L2 group on Task 4. Hence, children with lower levels of general ability who perform below the expected standards of literacy and numeracy for their age may find it especially challenging to understand word order rules, while children at higher levels find this comparatively easier. Moreover, more in-depth knowledge of a single L2 may be more beneficial in identifying cognates among simple and compound nouns than superficial knowledge of multiple L2s, which is arguably somewhat surprising.

It should be noted that the extensive pilot studies which were undertaken to validate the test of metalinguistic awareness used in the present study yielded coherent results, generally showing moderate positive correlations between individual tasks. A factor analysis suggested a one-factor solution, indicating that the test measured a single construct, with each task making a unique and non-redundant contribution (Tellier, 2013). The results from the present study additionally suggest that some tasks may be more powerful discriminators than others when populations with different background characteristics are compared.

Limitations and suggestions for further research

The empirical study reported in this paper was conducted with an intact cohort of school children, thus ensuring ecological validity. At the same time, this meant that we had no control over the quality or quantity of prior L2 input the children were exposed to. This constitutes a limitation, unfortunately, since we were unable to compare learners who had been exposed to Esperanto only with learners who had been exposed to a different language/different languages. Future research should aim for such a ‘pure’ comparison. In addition, the present study yielded a number of results that were statistically non-significant, but suggestive of certain trends. It is an open question as to whether these trends may yet

cross the threshold of statistical significance as children's L2 knowledge grows in depth over time. Thus, it seems desirable to plan future research not just with 'pure' comparison groups, but also with a more longitudinal design. Measures taken at different points over a longer period of time would allow us to identify with certainty whether and when any trends, as apparent in the present study, turn into significant effects.

Notes

[1] Given the conceptual closeness of the two notions, the terms 'metalinguistic ability' and 'metalinguistic awareness' are used interchangeably in this paper.

[2] A reviewer suggested that one could also argue the exact opposite, i.e. learning a 'difficult' language first might foster metalinguistic awareness. We agree that it is logically possible to argue in this way, and the argument may well apply if the targeted participants are cognitively mature, fully literate, and of high general intelligence. Given the context of our research, this line of argument is less convincing, however. Children aged 11 to 12 are still developing in terms of cognitive maturity, and their literacy skills have not been perfected. Lower-ability children in particular struggle with the demands of mastering the finer points of their L1 as emphasised in a school setting, let alone the challenges posed by an L2, especially in the minimal-input setting of the average school classroom.

[3] The taxonomy of learning difficulty we use is concerned exclusively with what has been termed elsewhere "relative complexity" or, alternatively, "*cognitive complexity* or simply *difficulty*" (Bulté & Housen, 2012 : 23; emphasis in original). We acknowledge that other taxonomies have been proposed (e.g. Bulté & Housen, 2012; Dietz, 2002) and operationalised (e.g. Housen et al., 2005).

[4] For reasons of clarity, we employ the term ‘state school’, which is commonly used and readily understood. Schools that are financially maintained by the government and therefore obliged to follow government guidelines are technically known as ‘maintained schools’ (see, e.g., http://www.education.gov.uk/schools/leadership/typesofschools/maintained).

[5] Children in the U.S. and South Africa could potentially have developed proficiency in Spanish and Afrikaans, respectively. In the event, an analysis comparing the excluded children with the main sample showed that the bilingual children significantly outperformed the main sample on Section 1 of the metalinguistic awareness test ($t(223) = -2.294$, $p = 0.23$, two-tailed) as well as on three of the eleven individual tasks constituting the test. Differences in performance on a fourth task approached significance, with bilinguals achieving higher scores than the main sample.

[6] In Year 3, when the children were 7 to 8 years of age, they saw early versions of Tasks 1, 3 and 4. In Task 1 (see Appendix A for current version), four of the seven sentences were used and children were asked to put a tick next to the two sentences that meant the same; the other sentences were redundant, and no translation was involved. In Task 3, six of the eight nouns from three languages (including English), instead of four languages (not including English) as in the current version, were used, and children were asked to identify whether a noun was singular or plural, as opposed to the current version, which asks children to pair singular and plural nouns from the same language. In Task 4, children were asked to match five words in different languages with three Esperanto cognates, as opposed to the current version which includes nine words and five cognates. In Year 4, when the children were 8 to 9 years of age, 33 of them saw the early versions of Tasks 1 and 4 again, and they also saw one of the sentences included in Task 2. In Year 5, when the children were 9 to 10 years of age, 22 of them saw the early version of Task 1, the sentence from Task 2, and a slightly modified

version of Task 4. At no point did the children receive feedback on their performance on the early versions of the tasks.

[7] A reviewer suggested that we additionally compare Groups 1, 2 and 3 only, since these groups are most comparable in terms of sample size and language learning experience, all having been exposed to four years of instruction in Romance languages (Groups 2 and 3) or Esperanto and a Romance language (Group 1). A one-way ANOVA revealed no statistical differences between the three groups, although the p-value approaches significance ($F(2, 99) = 2.632$, $p = 0.08$). As Levene's test of homogeneity of variance was significant in this instance ($p = 0.03$), a Kruskal-Wallis test was also run. The result was the same, with the p-value approaching, but not quite reaching statistical significance ($p = 0.08$).

References

Anderson, J. R. (2005). *Cognitive psychology and its implications* (6th ed.). New York, NY: Worth Publishers.

Ashby, F. G., & Casale, M. B. (2005). Empirical dissociations between rule-based and similarity-based categorization. *Behavioral and Brain Sciences, 28*(1), 15-16.

Bailey, T. M. (2005). Rules work on one representation; similarity compares two representations. *Behavioral and Brain Sciences, 28*(1), 16.

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

Bialystok, E. (1994b). Representation and ways of knowing: Three issues in second language acquisition. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 549-570). London: Academic Press.

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

Bialystok, E., & Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science, 19*(1), 19-23.

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences, 16*(4), 240-250.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325-339.

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

Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. Berlin: Springer.

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.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency* (pp. 21-46). Amsterdam: John Benjamins.

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.

Collins, L., Trofimovich, P., White, J., Cardoso, W., & Horst, M. (2009). Some input on the easy/difficult grammar question: An empirical study. *Modern Language Journal, 93*(3), 336-353.

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. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, 499-533.

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.

Dietz, G. (2002). On rule complexity: A structural approach. *EuroSLA Yearbook, 2*, 263-286.

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

Ellis, N. C. (1994). Consciousness in second language learning: Psychological perspectives on the role of conscious processes in vocabulary acquisition. *AILA Review, 11*, 37-56.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning, 54*(2), 227-275.

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

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.

Formaggio, E. (1990). Lerneja eksperimento pri lernfacileco kaj transfero en la fremdlingvoinstruado. *Internacia Pedagogia Revuo, 20*, 1-9.

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.

Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning, 51*, 1-50.

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.

Housen, A., Pierrard, M., & Van Daele, S. (2005). Structure complexity and the efficacy of explicit grammar instruction. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 235-269). Berlin: Mouton de Gruyter.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 349-381). Malden, MA: Blackwell.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition, 27*(2), 129-140.

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.

Jessner, U. (1999). Metalinguistic awareness in multilinguals: Cognitive aspects of third language learning. *Language Awareness, 8*(3&4), 201-209.

Jessner, U. (2006). *Linguistic awareness in multilinguals: English as a third language*. Edinburgh: Edinburgh University Press.

Jessner, U. (2008). A DST model of multilingualism and the role of metalinguistic awareness. *Modern Language Journal, 92*(2), 270-283.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Karmiloff, K., & Karmiloff-Smith, A. (2002). *Pathways to language: From fetus to adolescent*. Cambridge, MA: Harvard University Press.

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.

Lodge, O. J. (2004/1905). *School teaching and school reform*. London: Elibron Classics.

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

Masson, H. (2006). Quel parti la language française peut-elle tirer de la défense et de l'illustration de l'espéranto? *La Informilo, 133*, 25-28.

Maxwell, D. (1988). On the acquisition of Esperanto. *Studies in Second Language Acquisition, 10,* 51-61.

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. (2008). Age-related differences in foreign language learning: Revisiting the empirical evidence. *IRAL*, *46*, 197-220.

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.

Muñoz, C. (Ed.). (2006). *Age and the rate of foreign language learning*. Clevedon: Multilingual Matters.

Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology, 113,* 535-551.

Reber, R. (2005). Rule versus similarity: Different in processing mode, not in representations. *Behavioral and Brain Sciences, 28*(1), 31-32.

Roehr, K. (2012a). The Springboard to Languages evaluation project: a summary report. In A. Tellier (Ed.), *Esperanto as a starter language for child second-language learners in the primary school* (pp. 23-34). Stoke-on-Trent: Esperanto UK.

Roehr, K. (2012b). Theoretical foundations. In A. Tellier (Ed.), *Esperanto as a starter language for child second-language learners in the primary school* (pp. 19-22). Stoke-on-Trent: Esperanto UK.

Roehr, K., & Gánem-Gutiérrez, G. A. (2009a). Metalinguistic knowledge: A stepping stone towards L2 proficiency? In A. Benati (Ed.), *Issues in second language proficiency* (pp. 79-94). London: Continuum.

Roehr, K., & Gánem-Gutiérrez, G. A. (2009b). The status of metalinguistic knowledge in instructed adult L2 learning. *Language Awareness, 18*(2), 165-181.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*(2), 263-308.

Sparks, R., Ganschow, L., Fluharty, K., & Little, S. (1995-6). An exploratory study on the effects of Latin on the native language skills and foreign language aptitude of students with and without learning disabilities. *The Classic Journal, 91*(2), 165-184.

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

Symoens, E. (1992). *Al nova internacia lingvopolitiko: La propedeŭtika valoro de Esperanto* (Vol. 28-29E Esperanto-Dokumento). Rotterdam: Universala Esperanto-Asocio.

Tellier, A. (2013). Developing a measure of metalinguistic awareness for children aged 8-11. In K. Roehr & G. A. Gánem-Gutiérrez (Eds.), *The metalinguistic dimension in instructed second language learning* (pp. 15-43). London: Bloomsbury.

Tellier, A., & Roehr-Brackin, K. (2013). The development of language learning aptitude and metalinguistic awareness in primary-school children: A classroom study. *Essex Research Reports in Linguistics, 62*(1), 1-28.

Thorndike, E. L., & Kennon, L. H. V. (1927). *Progress in learning an auxiliary language: A report*. New York, NY: Institute of Educational Research, Columbia University.

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.

Yelland, G. W., Pollard, J., & Mercuri, A. (1993). The metalinguistic benefits of limited contact with a second language. *Applied Psycholinguistics, 14*, 423-444.

Appendix A: Task 1

1

Which sentences mean the same?

1. El gato es blanco.
2. O elefante é grande.
3. La rana è verde.
4. De olifant is groot.
5. Tigrul este negru şi verde.
6. La rano estas verda.
7. The cat is white.

Pairs

☐ and ☐

☐ and ☐

☐ and ☐

The sentence which is left means:

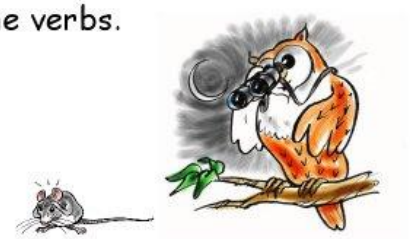
_____

Appendix B: Task 8

8

Look again at the two sentences in the boxes below this picture. Each one has three words. Two of the words in each sentence are nouns, and one is a verb.

Draw a circle around the verbs.



ogrtsi grarid omsu-ch	omsu-ch grarid ogrtsi

****

This sentence has two nouns and a verb, but no adjectives.

ogrtsi dul oplki-ch

Invent two adjectives, one for the owl and one for the ball. It doesn't matter how many letters they have or how you spell them. Write them here.

Here is the sentence again:  ogrtsi dul oplki-ch

Write it in the box below and put your two adjectives where you think they should go.

Your sentence will now have five words.
