Published in final edited form as:
Child Lang Teach Ther. 2003 October 1; 19(3): 311–337.

Syntactic development in fluent children, children who stutter, and children who have English as an additional language

Peter Howell, **Stephen Davis**, and **James Au-Yeung**
Department of Psychology, University College London, London, UK

Abstract

Children aged between two and 10 years were assessed on a new reception of syntax test (ROST). Validations of the test are reported for monolingual fluent control children under five (by examining the relationship with mean length of utterance and the Oxford Communication Development Inventory) and for over fives (relationship with a new judgement of grammaticality test using syntactic categories common to the two tests). Performance of these children was compared with performance of children who stutter and children with English as an additional language. In this study, the test was divided into under-five and over-five forms. Any young child progressing to the over-five syntactic categories, or any older child doing the under-five syntactic categories was dropped from the analysis. ROST scores prepared according to this scheme led to no differences between the control and either of the subject groups tested. However, compared to controls, the children with English as an additional language (but not children who stutter) had a significantly higher proportion of children above five who did the under-five categories (and were, therefore, excluded from the analyses). The higher proportion of children who did the under-five syntactic categories in the English as an additional language group indicates that group scores would have been lower if their syntax results had been included in the analysis. Further analyses provided some evidence that two groups with English as an additional language (Turkish and Cantonese speakers) did not perform any better on selected syntactic categories in their native language compared with their performance in English.

Introduction

This paper introduces a new test (the reception of syntax test, ROST) that is suitable for use with young children, for children who are suspected of having a syntactic deficit, and for testing syntax in languages other than English. The three aims of the paper are: 1) to validate ROST by establishing the relationship between performance on this test and other measures of syntax used in early language development and in second language learning; 2) to report results on English syntactic development over age for children who stutter and speakers who have English as an additional language (EAL); and 3) to compare syntactic performance across languages in EAL children (children who speak Turkish or Cantonese in addition to English). The next section of the introduction gives a brief description of ROST. Other measures of syntax that are used to validate ROST are then described. The grounds for wanting to test syntax development in children who stutter and EAL children are given in the third section. In the final section, differences between Turkish and English, and Cantonese and English for selected syntactic categories are described, forming the basis of the cross-language comparisons.

Address for correspondence: Peter Howell, Department of Psychology, University College London, Gower St., London WC1E 6BT, UK. E-mail: p.howell@ucl.ac.uk.

Reception of syntax test

ROST is a test of syntax reception. Currently, the most widely used such test is Bishop's (1983) test of the reception of grammar (TROG). This assesses young children's performance in English from about age four, somewhat above the ages to be tested here. Thus, a simpler receptive test is required. ROST meets this requirement. In ROST, a sentence is played from a laptop computer that displays two pictures. One of the pictures corresponds to the sentence and the other does not. A correct response is made when the image appropriate to the sentence is touched and an incorrect response occurs when the inappropriate image is touched (these responses are detected by using a touch screen). Touching the picture that corresponds with the sentence appears to be a natural response for a child to make, allowing the test to be used with young children.

The syntactic categories that are tested depend on a child's age (there are different forms for under- and over-fives). ROST checks that the chosen level of difficulty is appropriate for a child based on the results of an on-going analysis of syntactic categories. The selection of syntactic categories is changed if a child fails easy (over-five) or passes hard (under-five) syntactic categories appropriate to his or her age. Although there is this facility in ROST, in the current study, results are only reported on children who performed on their age-respective test form (under-fives on the under-five form and over-fives on the over-five form). The reason for omitting the other children is that the under-fives who progressed were not necessarily tested on all under-five syntactic categories and the over-fives that regressed were not tested on all the over-five categories. Although the under-fives who progressed may have had a problem on syntactic categories on which they were not tested, this would not be apparent. This is not a problem in monolingual children whose syntax is developing normally. However, EAL children, for example, may have a problem on simple syntactic categories but not complex ones because of carry-over effects from their native language. For instance, if the additional language has a different word order from English, this simple category might pose difficulties whereas more complex categories that correspond across the languages might not. All categories need to be tested to identify specific difficulties such as these. A similar argument applies to over-five EAL children. Such children may have difficulties on their age-respective test form with specific syntactic categories, which leads them to be tested on the under-five form, even though they may be able to process other difficult syntactic categories perfectly well.

Measures of syntactic performance used with preschool and EAL children

Mean length of utterance (MLU) (Brown, 1973) is probably the most widely used measure for assessing syntax in fluent monolingual children near language onset. The measure is based on the length of a child's spontaneous productions. Procedures for obtaining MLU scores are described fully in Brown (1973). MLUs are used here as a measure of syntactic development for children under four, and ROST scores are obtained with the same children. Although ROST is a receptive test and MLU a production measure, some relationship would be expected, because both assess general syntactic-processing ability.

A more recent assessment instrument, whose use is widespread in early language development in the UK, is the Oxford version of the Communicative Development Inventory (OCDI, Hamilton *et al.*, 2000). This measure is based on questionnaires filled out by a parent, who supplies information about the words a child uses and understands. The relationship between scores on OCDI and ROST is examined for children under four. A positive relationship would be expected again if both measures provide indexes of syntactic-processing ability.

As ROST is employed with EAL children in this study, a test was also included to assess syntactic development in such subjects. The test used is an automated form of Demuth *et*

al.'s (2000) judgement of grammaticality (JOG) test, which is probably less familiar to readers than MLU and OCDI. Grammatical judgement is a procedure for testing the level of syntactic skills a child has attained (McDaniel and Cairns, 1996). Most such judgement tasks ask a subject to give a 'Yes' or 'No' answer upon hearing a sentence, to indicate if the sentence is grammatical. In some forms of such tests, sentences constructed with different grammatical errors and error-free control sentences are presented to subjects and they are asked to judge if the sentences are grammatical or not (Johnson and Newport, 1989; Epstein *et al.*, 1996). Variants of the test procedure exist. For example, a scale of 'correctness' has been used instead of a straight 'Yes' or 'No' response (Montrul, 2001).

Such procedures are not suitable for testing young children because it is difficult to explain to them what is meant by 'grammatical'. The JOG grammaticality judgement test developed by Demuth *et al.* (2000) is suitable for young children (it has been tested on participants aged between three and 58 years). In the test, stimulus pairs were recorded and played that related to the animation of two hand puppets (a panda and a sheep). One of each pair of sentences was ungrammatical and the other error-free. Participants were asked to indicate which animal spoke best.

The JOG and ROST tests were developed for different reasons. The JOG test format is often used to assess progress when speakers are learning an additional language, whereas ROST is designed to assess syntactic development. However, both tests depend on understanding underlying grammatical characteristics. Pronoun gender and word order categories that appear in ROST were included in JOG, so performance on the two tests can be compared. A tendency would be expected that if a syntactic category is passed in one test, it would also be passed in the other, and similarly if a syntactic category is failed in one test it would also be failed in the other.

Syntactic processing in children who stutter and EAL children

Children start to use syntax (at least in a rudimentary form) when they progress beyond the one-word stage, usually at around two years of age. Speech disorders, like stuttering, also start at around this age (Yairi and Ambrose, 1992). A disorder like stuttering may be precipitated when speakers start to use syntactic constructions (Bernstein Ratner, 1997). A child who uses English as an additional language (EAL) may experience problems in language development due to having to deal with different syntactic forms in their two (or more) languages (as well as for other reasons). If so, they would be expected to be behind monolingual English children in their syntax development. There is also the possibility that stuttering and language acquisition in EAL children are related as stuttering has been reported to be more prevalent in EAL, compared to monolingual, children (Eisenson, 1986; Karniol, 1992; Mattes and Omark, 1991; Shames, 1989; Stern, 1948; Travis *et al.*, 1937).

Comparison of selected syntactic attributes between English, Turkish and Cantonese

An EAL child may have difficulty in English because of limited exposure to this language. Development in the additional language may, on the other hand, be unaffected. This is likely to depend, to some extent, on factors like age of acquisition of English and, in immigrant populations, age of arrival in the English-speaking country. In many cases it is not appropriate to obtain this information from a child or parent. An alternative approach is to assess syntactic performance in both languages. The problem then is that few tests are available that can be applied in a variety of different languages. An exception is ROST, which is convenient for this purpose as pre-recorded sentences are used. This also means that someone who is not a native speaker of the language tested can administer the test. Versions of ROST have been developed for a wide variety of languages (with the assistance of native speakers of these languages). In the current study, syntactic performance of English-Turkish and English-Cantonese speaking

children are assessed in both pairs of languages. Examination of selected syntactic categories across languages is made to ascertain whether children have a specific problem in one language.

The syntactic categories selected for comparison are word order, use of pronouns, pluralization of nouns, ways of negating and signifying comparatives that are used in different ways in the selected languages. For word order, English and Cantonese have subject verb object (SVO) structure while Turkish has subject object verb (SOV) structure. The English pronoun system distinguishes between both gender and number by use of different lexical items while both Cantonese and Turkish only distinguish between singular and plural pronouns with a morpheme. For the pluralization of nouns, English and Turkish use a plural morpheme (except for the irregular forms) while Cantonese lacks a plural morpheme and uses a system of singular and plural classifiers instead. The negation marker in Turkish is sentence-final (post-verbal), while both English and Cantonese have a pre-verbal marker, reflecting the differences in the SOV versus SVO basic structure. In comparative structures, Turkish has a morpheme to mark out the sentential object while English and Cantonese use a morpheme on the adjective to signify these structures. In summary, English and Turkish have similarities in use of the plural, but differ in word order, pronoun forms, negation and use of comparatives, whereas English and Cantonese have similarities in word order, negation and comparatives, but differ in pronoun forms and use of plurals. The differences in the way the syntactic contrasts are made between English and each of these languages could lead speakers to have difficulties in English on specific syntactic categories (word order, pronoun forms, negation and use of comparatives in Turkish; pronoun forms and use of plurals in Cantonese). Thus, studies have shown, for a variety of languages, that the syntax of a speaker's first language transfers to an additional language (Nagy *et al.*, 1997; Slabakova, 2000; Su, 2001).

If ROST is a valid measure of syntax, it should correlate with measures like MLU and JOG, which assess syntax performance in different ways, and OCDI, which measures lexical acquisition. In addition to these validations, results on application of ROST for tracking syntax development in children who stutter and EAL children are reported. Assessment of whether speaking two languages results in syntactic problems in just one or both languages is reported.

Method

Children

The total numbers of all monolingual English children tested on ROST are given in Table 1 for control children (fluent monolingual English), children who stutter (CWS) and EAL children. The criterion for admitting an under-five child to the test is that he or she performed the under-five version and the criterion for admitting an over-five child to the test is that he or she performed the over-five version The number of children who met these selection criteria (that are included in the analyses) for all language groups are given in parentheses. These are also broken down by age and gender.

Control children up to age four were tested in their own home (usually in the morning when the child was most alert). These children were brought up in a monolingual English environment. They were recruited from birth announcements in local papers and from National Childbirth Trust (NCT) groups. This resulted in a somewhat less representative sample than the population at large (some of the latter would be at nursery, and so on). Such considerations are seen as unavoidable in this type of research. For example, Gershkoff-Stowe and Smith (1997) reported that their sample consisted mainly of Caucasian, middle-class monolingual speakers.

Control children aged four and over were pupils from inner London schools and represent a wider cross-section of the population than the younger children. All children in each class were

tested. None of these children had special educational needs. These children were also reported to use no other languages at home, although this could not be verified with the children or their parents because the schools involved advised against it.

Children who stutter were located through clinics in London boroughs with a similar demographic makeup as the inner London schools. There were 24 children who stutter in the age range 5-10 years (there were no children who stutter aged eight). Twenty of the 24 children who stutter completed the over-five form of ROST.

The EAL children came from the same classes as the control children. The EAL children had a variety of additional languages. In total, 24 Turkish-speaking children and 18 Cantonese-speaking children were included in the EAL group. From each of these language-groups 18 children were also tested in their other respective language using alternative forms of ROST.

English, Turkish and Cantonese versions of the reception of syntax test, ROST

To give an overview of ROST, the child starts by looking at two pictures (or four for three of the advanced syntactic categories, X, Y and Z, indicated in Table 2) on the screen of a laptop computer. The child listens to a sentence played from the computer that corresponds with one of the pictures. The child should be able to link the picture with its sentence if he or she has acquired the target syntactic dimension. Responses are made by touching the picture appropriate to the sentence (detected automatically by the computer). Before testing, all subjects were given four training items consisting of two pictures of different objects and a recording of a single word utterance to ensure they knew how to make the responses.

The pictures in a syntactic-test pair contrast only on the target syntactic dimension. Each pair of pictures was tested twice, and a different picture was the correct answer on one of these two test occasions. The two pictures in each pair were displayed side by side on the computer screen, with a space of 18mm between them (in the four picture categories, two appeared at the top and two at the bottom of the display). The pictures occupied the same left/right position on the two occasions they appeared in order to counter any left/right response bias a child might have. The two pictures were surrounded initially by two identical pale blue frames.

The probe sentences or phrases were pre-recorded in WAV format and were played over an external loudspeaker connected to the computer (to produce good sound quality). A single native adult male speaker was used throughout the recordings and pronunciation was checked thoroughly. In some cases, a simple phrase instead of a full sentence was used for testing so as to minimize the amount of time and information that had to be processed. A loudspeaker icon was displayed bottom centre of the screen, below the two pictures. Whenever, the loudspeaker icon was touched (done by the experimenter for the young, but not the older children), the utterance for that particular test item was played back through the loudspeaker. Picture selection could not take place until the whole utterance had been played at least once, although the utterance could be played back as many times as desired.

For children under five, nine syntactic categories were tested. The nine categories and two sentences appropriate to one picture pair were (a total of four pairs of pictures was used for each syntactic category): (A) SVO word order (‘The boy kisses the girl’/‘The girl kisses the boy’), (B) article, a(n)/some (‘a fish’/‘some fish’), (C) plural morpheme (‘cat’/‘cats’), (D) pronoun he/she (‘He runs’/‘She runs’), (E) pronoun his/her (‘his cat’/‘her cat’), (F) pronoun he/they (‘he danced’/‘they danced’), (G) preposition in/on (‘in the box’/‘on the box’), (H) preposition (others, such as up/down), (‘going up’/‘going down’), (I) compound noun (‘car cake’/‘cake car’ where ‘car cake’ was meant to signify a cake made in the shape of a car and ‘cake car’ was meant to signify a car for delivering cakes). Simple words such as animal names, names of common objects and toys and words like ‘boy’ and ‘girl’ were used as subjects and

objects of the phrases. Action verbs were also kept simple (e.g. ‘kiss’, ‘run’, ‘look’). The selection of syntactic categories was developmentally motivated. However, practical considerations sometimes limited the test material. For example, it would be impossible to represent the contrast of ‘the’ versus ‘a’ pictorially. Therefore, ‘some’ and ‘a(n)’ were used. Special nouns were also needed so that the plural feature was not signalled by a plural suffix. The nouns used were ‘sheep’, ‘fish’, ‘deer’ and ‘aircraft’. Although the word ‘aircraft’ may be difficult for very young children, it was not possible to find an easier replacement. Note, however, that the information from the article alone was sufficient for picking the correct picture. On the two occasions on which a picture was presented (see above), each picture was the right response on one occasion. In total, there were 9 (syntactic categories) × 4 (picture pairs) × 2 (test presentations per picture pair) (=72) maximum possible test items in this form of the test.

The child responded by touching the picture they considered appropriate, and the response was sensed by a One Touch OT121PC touch-screen interface. The frame of the selected picture turned red to indicate which picture had been selected. At the same time, the frame of the other picture turned pale grey. After a picture had been selected, another button appeared at the bottom right-hand corner, which was used to move on to the next test item. Note that no indication was given to the child as to whether he or she had made the correct or incorrect choice so the child did not learn (over test sessions) the correct answer to a picture without using his/her syntactic knowledge. This allowed later re-testing with this version of ROST to track a child’s language development (not involved in the data reported).

On any trial a child performing at chance would be right 50% of the time in a two-picture task. Making seven out of eight equiprobable binary responses (four picture pairs × two test presentations per picture pair) correctly, would happen on less than 5% of occasions by chance. Thus, if the child makes seven correct responses on any syntactic category, the conventional 5% level for claiming statistical significance is achieved. If a child failed two test items in a category (which meant the child could not pass that category according to the statistical criterion just given), the child failed that category and no further test items in that category were tested. If a child passed the first seven items in a category, the eighth item was not tested as the child had met the statistical criterion. These restrictions reduced the testing time without compromising the power of the test.

Data on children under five in this article are only for performance on the first nine categories. Using performance on the first nine categories alone in the following analyses for the under-fives ensured that all these blocks were tested and scored as pass or fail. This form of the test for categories A-I took about 10 minutes.

Children over five were not tested on the first eight categories unless they failed some of the simpler remaining categories. Any child aged over five who had to perform the simpler categories was excluded from the current analyses to ensure results of children (pass or fail of each syntactic category) were based on categories actually tested rather than those deemed to be too complex for the child to be tested on given his or her performance on selected syntactic categories. The categories used for over-fives in the analyses are I, J, K, M, N, O, Q, R, S, T, U, W, X, Y, and Z (L, P and U are omitted as they are options that are tested). All the children (whatever their ages) are included in a normative study where options are used that allow level of difficulty of the syntax test to be tailored to a child whatever his or her age. Customizing the test can lead to certain categories being credited without being tested and these categories are those excluded in the current study. The test lasted for between 20 and 25 minutes. Examples of all syntactic categories are given in Table 1.

Similar tests to the English one just described have been developed for other languages with the assistance of native speakers of these languages (as linguistic informants and for the speech recordings). The languages reported on here are Turkish and Cantonese. For these EAL language groups, there are 18 children who have been tested in English and their other language. The syntactic categories tested are indicated in column 4 for Turkish (13 categories) and in column 5 (14 categories) for Cantonese. ROST results (for whatever language) were stored after each completed test and were analysed by another computer program later. All the children tested, so far, enjoyed and concentrated on the test.

Mean length of utterance (MLU)

Mean length of utterance (MLU) data were obtained from analysis of audio recordings made in the children's homes at the time of the ROST assessment. Following completion of the ROST assessment, the researcher and child interacted in an unstructured way for approximately 30 minutes. The session was recorded on a digital audio tape recorder and the recording was orthographically transcribed later. The first 100 utterances were used for analysis and the MLU was calculated using the criteria presented in Brown (1973). MLU scores were obtained for 32 children (17 boys and 15 girls, age range 24-57 months, mean 32.5, SD 0.62).

The Oxford communication development inventory (OCDI)

Productive vocabulary data were obtained from parental reports using the OCDI (Hamilton *et al.*, 2000). OCDI is a checklist composed of 416 words, which are among the first to appear in the vocabulary of young English children. The words are divided into the following categories: sounds, animals, vehicles, toys, eating/drinking, body parts, furniture/rooms, out, house objects, people, games/routines, action words, descriptive words, question words, time, pronouns, prepositions, and quantifiers. The OCDI was used to obtain indications of the child's productive vocabulary. Parents recorded productive vocabulary by filling in the form.

Judgement of grammaticality (JOG) test

The JOG test was run on 24 English-Turkish EAL children and 24 age-matched monolingual English children. JOG is a test where the subject had to choose which of a pair of sentences was grammatically correct. JOG implemented the procedure of Demuth *et al.* (2000) in a computerized form. The test was run on a laptop computer fitted with a touchscreen, as with ROST. A loudspeaker was also connected to the laptop for sound output. JOG used the recording of one male and one female native English speaker, and subjects had to decide whether the male or female spoke grammatically correct English.

In order to maintain attention level with young children, JOG was administered in four sessions, each lasting approximately 10 minutes. Before each session, four practice test items were used to familiarize or refamiliarize the child with the pictures and speakers used in the task. For these practice items, a picture of an object was displayed (e.g., a cat) top centre of a computer screen. A correct (e.g., 'cat' by the male) and an incorrect object name (e.g., 'fish' by the female) were spoken (these same speakers produced sentence material for the test proper). Throughout, the face that corresponded in gender to the first voice in the sequence appeared on the left-hand side of the screen and the other face on the right-hand side of the screen. The male/female position was fixed for each test session but was random for different sessions. The lower part of the screen displayed schematic faces (one of each gender) and control buttons for replaying the recording if required (as in ROST). When the test sentences were played, only the face of the speaker producing the utterance was displayed. After both sentences were produced, both faces were displayed for the child to select his or her response. Alternatively, after a sentence pair had been heard, the child could opt to hear the sentence pair again (by touching the loudspeaker button). To make a response, the child touched the picture of the speaker who spoke the correct name of the object displayed.

On a test trial, a picture that corresponded with the grammatically correct sentence was displayed at upper centre of the computer screen. A child then heard two sentences, one spoken by each of the speakers. A child indicated the speaker he or she thought was correct by touching the appropriate face on the touch screen. The grammaticality (correctness) of male/female recordings was randomized with the constraint that each speaker produced half the grammatical and half the ungrammatical sentences. The test was conducted in a quiet room. All the results were logged in a file for analysis.

Results

The results are organized according to the three aims outlined in the introduction: 1) ROST scores on fluent speakers and validation of these measures; 2) assessments on children who stutter and EAL children; and 3) assessment of syntax in two languages for EAL children.

Monolingual fluent English control baseline and validation of ROST on these children

The ROST scores of children under five are reported only for the first nine categories (A-I). Results for children under five who progressed and were not assessed on all these categories are not included. Data from children aged five and over are in the remaining categories (J onwards) plus category I but excluding L, P and U. Children who were not tested on all the selected categories were excluded from the analysis. Figure 1 shows the mean number of syntactic blocks (i.e., categories) passed at different ages (standard errors around the means are also indicated), separately for under-five and over-five versions of ROST. Even at age two, there is some evidence of syntactic ability (mean blocks passed 1.08, SD 1.73).

The percentage of four-year-old monolingual children who passed each of the syntactic categories (A-I) is shown in Figure 2. Results on the 15 categories for the seven-year olds are shown in Figure 3. The categories in Figures 2 and 3 are arranged in descending order of percentage of children passing in the categories tested. At four, only ‘compound noun’ appears to be difficult. ‘Compound noun’ continues to be a problem at age seven and ‘seems’, ‘S(sr) VO’ and ‘more relative clause’ also appear to be problematic for seven-year old children.

Results are next reported on validation of the form of ROST for children under five (up to 4 years 9 months). As stated in the introduction, the standard measures used to assess language in this age range are MLU and OCDI (or some related form of the latter). ROST was correlated with each of these measures. Scores on MLU and ROST were available for 32 children. A scatter plot using ROST score on the abscissa and MLU on the ordinate is given in Figure 4 with the best-fit line included. This shows that ROST scores increase as MLU increases. (There appear to be fewer than 32 points displayed in Figure 4 as some subjects fall on the same X, Y point.) The correlation between MLU and ROST blocks passed was significant ($r = 0.528$, $n = 32$, $P = 0.002$). The proportion of variance accounted for (r^2 represented as a percentage) was 28%.

OCDI scores for 13 children (eight boys and five girls) at 20-23 months (mean 21.84, SD 0.97) were compared to their ROST scores at around 30 months (range 29-32 months, mean 30.42, SD 1.29). The correlation between OCDI and ROST blocks passed was significant ($r = 0.638$, $n = 13$, $P = 0.019$), and percentage of variance accounted for was 41%. A second-degree polynomial fits the data somewhat better (proportion of variance accounted for was 50%). The scatter plot and best-fit line are shown in Figure 5. This shows that ROST scores increased as OCDI increased. Both the MLU and OCDI correlations with ROST show the test measures syntax development in young children. Note, though, that in the case of MLU in particular, a relatively low proportion of the variance is accounted for.

To validate ROST for older children, the performance on syntactic categories from this test was compared with performance on the equivalent syntactic categories tested in JOG. In total, 24 monolingual English children (aged 5-11 years) completed JOG and ROST. Pronoun gender and word order were tested in ROST and JOG. The results were examined to establish whether each child passed the category in both ROST and JOG, in neither, or in only one of the assessments. For the monolingual English children there were 48 comparisons (24 children × 2 syntactic categories). In 41 out of the 48 cases (85.4%) there was agreement between the assessments. Most of the discrepancies occurred with the younger children. Using data from children seven years and older ($n = 17$) the agreement rate rose to 94.1%, with only two out of 34 comparisons showing a discrepancy between assessments. With the monolingual English children, where there was disagreement between assessments, the children failed constructs in JOG that were passed in ROST

To check whether the validation applies to EAL as well as monolingual children, the performance of 24 EAL Turkish children (aged 5-11 years) who also completed JOG and ROST assessments in English were examined. Each of these children was selected to match in age and gender and to come from the same school class as the 24 monolingual English children. The same 48 comparisons involving pronoun gender and word order were made as with the English monolingual children. The results showed a similar pattern to the monolingual English children. In 40 out of the 48 cases (83%, compared with 85.4% for monolingual children) there was agreement between the assessments. Like the English children there was a high proportion who either passed the category in both assessments or failed in both assessments. All of the discrepancies between tests occurred in the younger children, similar to the data for the monolingual children (the corresponding figure for the monolingual English children was 94.1%). In the majority of the cases (six out of eight) where there was a discrepancy (in this case, always with young children), children who had passed a construct in ROST failed it in JOG. This (and the related finding with monolingual English children) may indicate that JOG is a more difficult test than ROST. The analyses of ROST and JOG for EAL Turkish children suggest that ROST is a valid measure for these children as well as for monolingual English children.

Performance on ROST of children who stutter and EAL children

The mean number of ROST blocks passed for the children who stutter (means and error bars) between ages five and 10 years are shown in Figure 6. Mean and standard errors of the monolinguals (presented earlier) are given for comparison. The two groups appear to perform very similarly.

An ANOVA was carried out including two factors: language group and age group. The language group factor had two levels (children who stutter and controls) and the age group had five levels (yearly age bands from five through ten, excluding age eight as there were no children who stuttered of this age). The analysis showed significant main effect of age, indicating syntax scores increased with age ($F(4,155) = 15.99$, $P < 0.001$). There was no effect of language group ($F(1,155) = 0.60$) and no interaction ($F(4,155) = 0.17$). Thus these results confirm that children who stutter and age-matched controls do not differ in terms of ROST blocks passed.

Counts were made of the number of children who stutter and their controls who did the under-five or over-five versions of ROST. A chi-square test on these data showed that there was no association between language group and numbers doing the under- or over-five syntactic categories [chi-square (df = 1) = 1.95]. The chi-square analysis suggests that the equivalent scores on ROST could not have been the result of proportionately more children who stutter than controls being unable to do the over-five categories. Note, though, that this conclusion needs to be taken with caution, given the low number of children in the stuttering group.

The mean number of ROST blocks passed and standard errors are shown for children aged between five and 10 years separately for EAL children and monolingual control children in Figure 7.

An ANOVA was carried out with two factors, language group and age group, with the same criteria for excluding children applied as earlier (the children had to perform all their age-appropriate ROST categories). The language group factor had two levels (monolingual and EAL children) and the age group had six levels (yearly age bands from five through 10). The ANOVA comparing performance on ROST of EAL children with controls showed significant main effects of age, as was expected ($F(5,260) = 20.04$, $P < 0.001$). There was no significant effect of language group ($F(1,260) = 0.56$). The interaction between age and language groups was also not significant ($F(5,260) = 0.50$). The comparisons show that for those children who passed the criterion for doing the over-five syntactic categories, EAL children performed the same as monolingual English children.

Fluent control and EAL children were classified as to whether they did the over-five or under-five syntactic categories. A chi-square test on these data was significant (chi-square = 13.7, df = 1, $P < 0.001$), indicating an association between language group and numbers doing the under- or over-five syntactic categories. It appears from this analysis that a significantly higher proportion of EAL children aged five and above had to do the under-five syntactic categories. This qualifies the conclusion that the ANOVA shows no difference as these children were dropped from this analysis and would have lower ROST scores.

There were enough data on EAL children to do a rank ordering of acquisition for four- and seven-year-old EAL children in the same way as for the monolingual English children. The data for four-year-olds are given in Figure 8 (these should be compared with the data in Figure 4) and the data for seven-year-olds are given in Figure 9 (these should be compared with the data in Figure 5). The rank order is very similar to the monolingual English children. This was supported by analyses separately on the four- and seven-year-old data using Kendall's tau. These indicated significant concordance between monolingual and EAL children in both age groups (Kendall's tau b was 0.551 for the four-year-olds, giving $P < 0.05$, and 0.683 for the seven-year olds, giving a $P < 0.001$). Thus, statistically, there is no difference in the order of acquisition of syntactic categories between language groups at ages four and seven years.

Influences across language with different syntaxes

The first question about the EAL children is how their performance in English compares with that of monolingual controls. The syntactic categories chosen for comparison were the five for which cross-language comparison is made later (word order, use of negative, pronoun, comparative and plural; see the introduction for a comparison of how these are used in Cantonese and Turkish compared with English). The number (out of five) of these categories that EAL speakers of Turkish (18 of the 24 used in the ROST-JOG validation reported above) and of Cantonese (18) was obtained and compared with monolingual English-speaking control children. Mean number of syntactic categories passed was lower for both EAL groups than their controls, although this was not significant by independent *t*-test for either language group.

Next, now it is known that syntactic performance in English is at about the same level as age-matched monolingual controls, what is their performance on their additional language on these syntactic categories? Without a monolingual control group for each of the additional languages (Turkish or Cantonese), it is not possible to establish whether learning English has affected the learning of the additional language. It is, however, worth checking whether performance on the additional language is superior to that in English. To make comparisons between the two languages, a score was obtained in each language (English and the additional language), separately for Turkish and Cantonese children, using the five attributes listed above (word

order, use of negative, pronoun, comparative and plural). The score again represented the number of categories (out of five) that each child passed. Related *t*-tests were then carried out to establish whether performance in one language was significantly different to that in the other for the 18 children in each language group aged 5-8 years. Results for Turkish-speaking and Cantonese-speaking children indicated that both groups passed more of these attributes in English than in their additional language. The differences were significant for both groups (Turkish-speaking $t(17) = 4.11$, $P < 0.001$, Cantonese-speaking $t(17) = 2.26$, $P < 0.05$).

Discussion and conclusions

The findings of this study are now discussed with respect to the three aims given in the introduction.

Validation of ROST on monolingual speakers who do not stutter

Validations have been reported for ROST. MLU and OCDI are measures of syntactic (Brown, 1973) and lexical development, respectively, that can be obtained with young children. Each of these measures correlated significantly with blocks passed in ROST in children under five, although proportion of variance accounted for was not high with MLU. Note that MLU is a productive measure of syntactic development and ROST a receptive measure. The significant correlations, bearing in mind the proviso about percentage variance accounted for with MLU, suggest that syntactic ability measured by the two techniques is shared between production and perception. The comparatively low proportion of variance accounted for by the regression may be due to mode-specific (production or perception) requirements. Validation of ROST with over-fives was made by employing another computerized test of syntactic development (JOG) that also tested pronoun gender and word order. When subjects passed on a syntactic category in ROST, they were likely to also pass on that dimension in JOG (there was some indication that the JOG forms may be more difficult than the ROST forms based on the asymmetry in categories passed in the two forms). From these results, it appears that ROST provides valid measures of a child's syntactic ability and allows such assessments to be attempted in children who stutter and EAL children.

Performance on ROST by children who stutter and EAL children compared with monolingual English controls

Performance of children who stutter was examined to establish whether they had difficulty in syntactic development. The first analysis compared ROST scores of children who stutter and the monolingual controls. Consistent with other examinations of the syntax of people who stutter (Kadi-Hanifi and Howell, 1992; Howell and Au-Yeung, 1995; Nippold, 1990, 2001), there were no marked differences between the groups. An additional analysis showed that there were not significantly more over-five children who stutter than controls who did the under-five syntactic categories of ROST.

Bernstein Ratner (1997) argues that the start of syntax use at around age two interferes with the lexical development already under way and may trigger the onset of stuttering around this age. Syntax itself appears to be developing normally in these children. It is, however, still possible to maintain that onset of syntax triggers stuttering. Assuming that at this age there is some independence of lexical and syntactic productive development, a speaker would need to decide how to distribute resources between the two processes (Starkweather and Gottwald, 2000). A child experiencing difficulty in dealing with syntax would need to channel more resources to controlling this aspect than on lexical output. The syntactic processing could then be performed appropriately, but individual words that are difficult to prepare for output (that need more than the allotted processing resources) would then lead to production difficulties. In this way, syntax could be a factor that precipitates stuttering that, at the same time, is not

revealed as a deficit in syntactic performance relative to fluent speakers. The appropriate way to test this would be to examine lexical as well as syntactic retrieval. Moreover, some temporal measure of these processes appears to be necessary as allocation of resources is affected by time pressure (including that which is self-inflicted by the speaker) (Howell, 2002; Howell and Au-Yeung, 2002). In summary, it is possible that time to make a syntactic decision is longer, but not less accurate, in children who stutter than in controls.

EAL children also showed no difference in ROST syntax scores compared with controls. With EAL children, order of acquisition was also examined and showed no difference. In this subject group, and in contrast to children who stutter, there were, however, significantly more children in the over-five test groups (EAL) who did the under-five syntactic categories of ROST. Inclusion of these subjects in the EAL group would have decreased ROST scores and potentially have led to a difference with controls.

Comparison of performance on selected syntactic categories across languages for EAL English-Turkish and English-Cantonese children

EAL children may not have developed syntax to the same level as monolingual children. When performance in English was compared for monolingual English and EAL Turkish and Cantonese children, the EAL children did not score as highly. Also, when the EAL children's English ROST scores and their ROST scores in their EAL language were compared, the EAL children showed that syntactic performance in their EAL language was not superior to that in English. As was warned in the results section, without a monolingual control for each of these additional languages, it is not possible to say that syntactic performance was affected in the additional language. However, it is of note that performance on the additional language was lower than that in English, bearing in mind that their performance in English was not at the same level as that of monolingual children. One other proviso is that, due to performing these tests in an educational setting, only children's age, not age of acquisition of English, was available.

Another comment about the current analysis of the EAL children is necessary. Separating out the over-five children who performed the under-five syntactic categories, taken in conjunction with the fact that children of the same age who performed the over-five categories were not significantly different from controls may suggest that only a subgroup of EAL children fall behind in syntax development. If so, from a practical perspective, it may be necessary to identify these children (using tests like ROST) and give them appropriate intervention.

In summary, ROST provides a valid measure of syntax development that is suitable for young children and for children who have English as an additional language. Although the current analyses showed no difference in ROST scores between controls and either speakers who stutter or EAL children when equivalent test forms had to be performed, a significantly higher proportion of EAL children could not perform at their age-appropriate level and were excluded in this analysis. As the EAL children who were excluded would have performed worse than those who were included, there is still the possibility that EAL children (or at least a subgroup of them) perform worse on this syntax tests than monolingual controls (Au-Yeung *et al.* in preparation). The test format involving touch screen technology and a portable laptop with test sentences pre-recorded appears to provide a valid test format for use with young and EAL children. Of particular note is the fact that pre-recording sentences allows testing of an EAL child by a worker who is not fluent in the test language.

Acknowledgement

This research was supported by the Wellcome Trust.

References

Bernstein Ratner, N. Stuttering: a psycholinguistic perspective. In: Curlee, R.; Siegel, G., editors. Nature and treatment of stuttering: new directions. 2nd edn. Needham, MA: Allyn & Bacon; 1997. p. 99-127.

Bishop, DVM. The test for reception of grammar. Age and Cognitive Performance Research Centre, University of Manchester, M13 9PL; 1983. Published by the author and available from

Brown, R. A first language: the early stages. Cambridge, MA: Harvard University Press; 1973.

Demuth K, Machobane M, Moloi F. Learning word-order constraints under conditions of object ellipsis. Linguistics 2000;38:545–68.

Eisenson, J. Language and speech disorders in children. New York: Pergamon Press; 1986.

Gershkoff-Stowe L, Smith LB. A curvilinear trend in naming errors as a function of early vocabulary growth. Cognitive Psychology 1997;34:37–71. [PubMed: 9325009]

Hamilton A, Plunkett K, Schafer G. Infant vocabulary development assessed with a British Communication Development Inventory: Lower scores on the UK then in the USA. Journal of Child Language 2000;27:309–43.

Howell, P. The EXPLAN theory of fluency control applied to the treatment of stuttering by altered feedback and operant procedures. Pathology and therapy of speech disorders. In: Fava, E., editor. Current issues in linguistic theory series. Amsterdam: John Benjamins; 2002. p. 95-118.

Howell P, Au-Yeung J. Syntactic determinants of stuttering in the spontaneous speech of normally fluent and stuttering children. Journal of Fluency Disorders 1995;20:317–30.

Howell, P.; Au-Yeung, J. The EXPLAN theory of fluency control and the diagnosis of stuttering. Pathology and therapy of speech disorders. In: Fava, E., editor. Current issues in linguistic theory series. Amsterdam: John Benjamins; 2002. p. 75-94.

Kadi-Hanifi K, Howell P. Syntactic analysis of the spontaneous speech of normally fluent and stuttering children. Journal of Fluency Disorders 1992;17:151–70.

Karniol R. Stuttering out of bilingualism. First Language 1992;12:255–283.

Mattes, LJ.; Omark, DR. Speech and language assessment for the bilingual handicapped. San Diego: College-Hill Press; 1991.

Nagy WE, McClure EF, Mir M. Linguistic transfer and the use of context by Spanish-English bilinguals. Applied Psycholinguistics 1997;18:431–52.

Nippold MA. Concomitant speech and language disorders in stuttering children: A critique of the literature. Journal of Speech and Hearing Disorders 1990;55:51–60. [PubMed: 2405212]

Nippold MA. Phonological disorders and stuttering in children: What is the frequency of co-occurrence? Clinical Linguistics and Phonetics 2001;15:219–28.

Shames GH. Stuttering: an RFP for a cultural perspective. Journal of Fluency Disorders 1989;14:67–77.

Slabakova R. L1 transfer revisited; the L2 acquisition of telicity marking in English by Spanish and Bulgarian speakers. Linguistics 2000;38:739–70.

Starkweather CW, Gottwald SR. The demands and capacities model: Response to Siegel. Journal of Fluency Disorders 2000;25:369–75.

Stern E. A preliminary study of bilingualism and stuttering in four Johannesburg schools. Journal of Logopaedics 1948;1:15–25.

Su IR. Transfer of sentence processing strategies: A comparison of L2 learners of Chinese and English. Applied Psycholinguistics 2001;22:83–112.

Travis LE, Johnson W, Shover J. The relation of bilingualism to stuttering. Journal of Speech Disorders 1937;2:185–89.

Yairi E, Ambrose NG. Onset of stuttering in pre-school children: selected factors. Journal of Speech and Hearing Research 1992;35:782–88. [PubMed: 1405533]

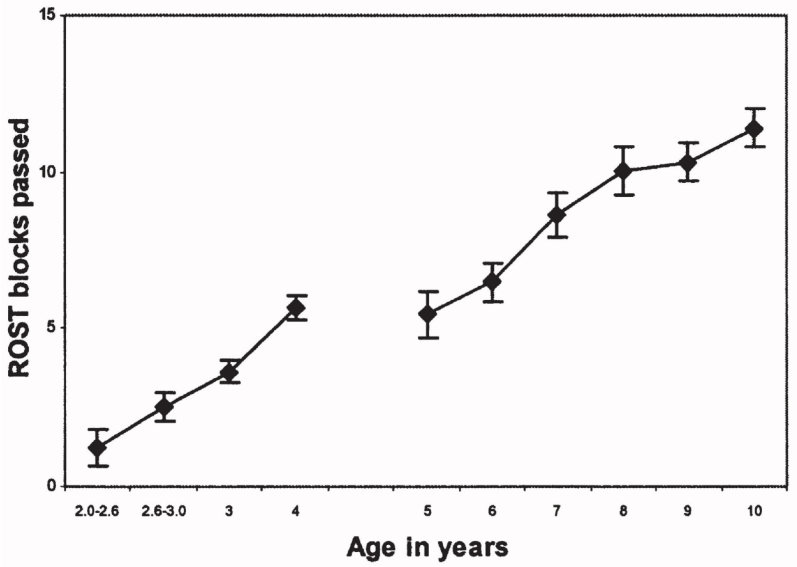

Figure 1.
Mean and standard error (SE) of ROST scores of control children (*y*-axis) versus age (*x*-axis) between two and four years and five and ten years. Note that the scores in the two age ranges are made on different syntactic categories (no overlap)

UKPMC Funders Group Author Manuscript

UKPMC Funders Group Author Manuscript

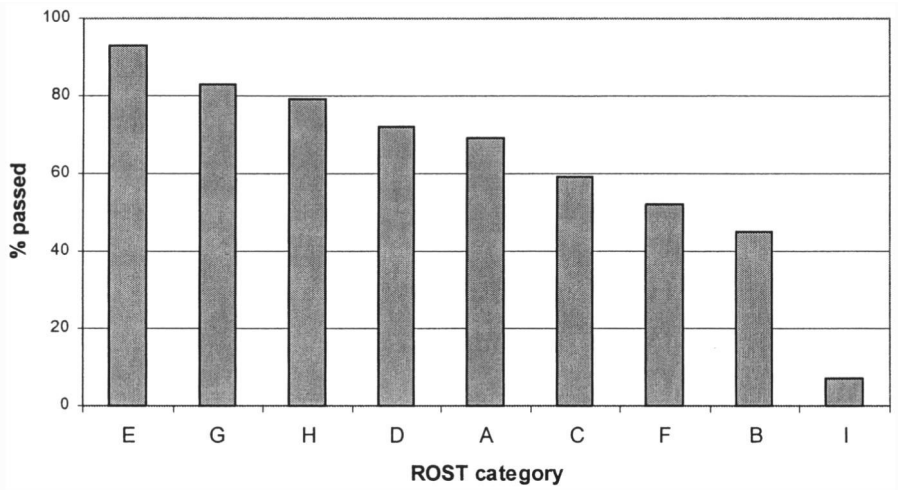

Figure 2.
Percentage of children under five (*y*-axis) passing each of the nine simple (A-I) syntactic categories (labelled on the abscissa). Syntactic categories are arranged in increasing order of difficulty (a lower percentage of children passing)

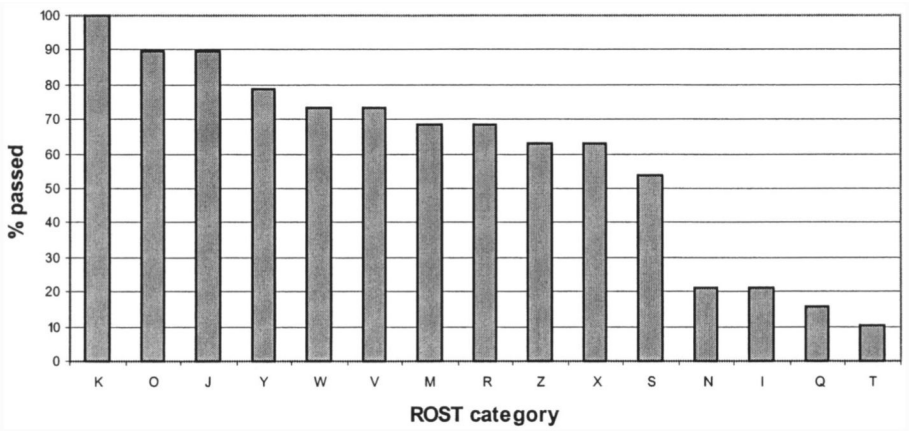

Figure 3.

Percentage of children over five (*y*-axis) passing each of the 15 complex syntactic categories (labelled on the abscissa). Syntactic categories are arranged in increasing order of difficulty (a lower percentage of children passing)

UKPMC Funders Group Author Manuscript

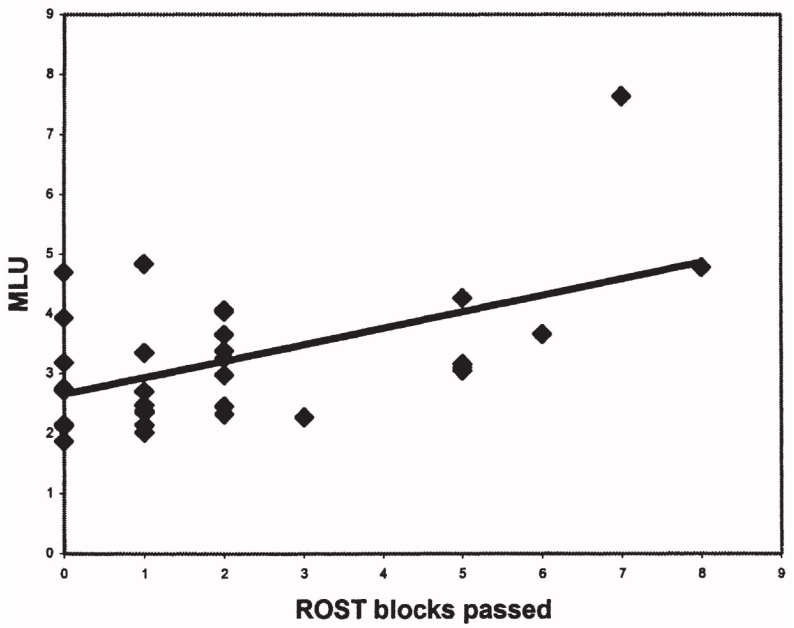

Figure 4.
Relationship between ROST scores (*x*-axis) and MLU scores for children between 24 and 57 months (*y*-axis) with best-fit line

UKPMC Funders Group Author Manuscript

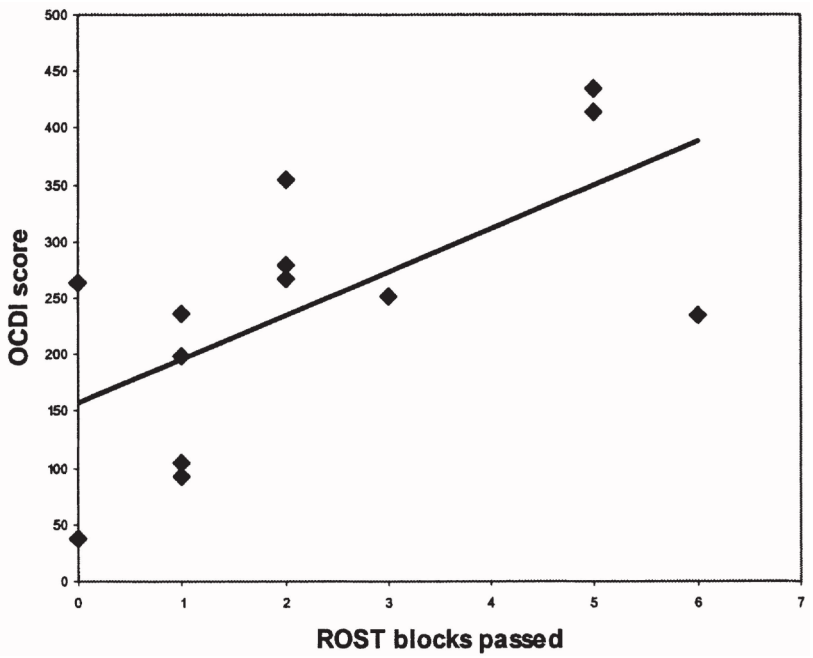

Figure 5.
Relationship between ROST scores (*x*-axis) and OCDI scores for children between 20 and 23 months (*y*-axis) with best-fit line

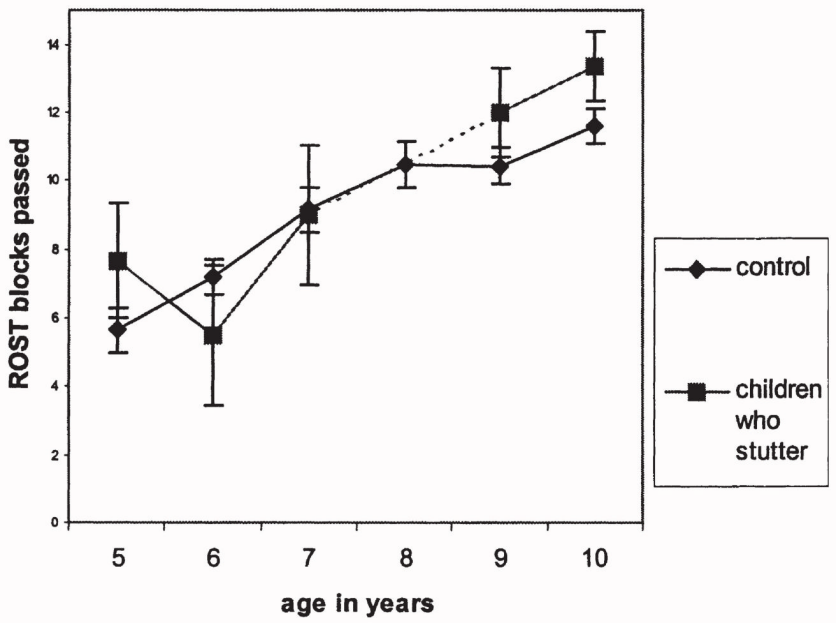

Figure 6.
Mean and SE of ROST scores of control children and children who stutter (*y*-axis) versus age (*x*-axis) between five and 10 years

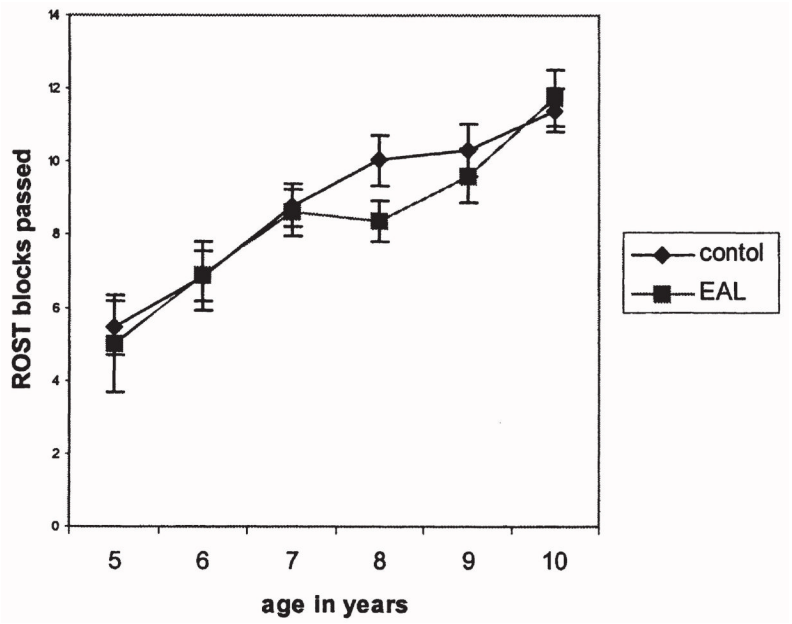

Figure 7.
Mean and SE of ROST scores of control children and EAL children (*y*-axis) versus age (*x*-axis) between five and ten years

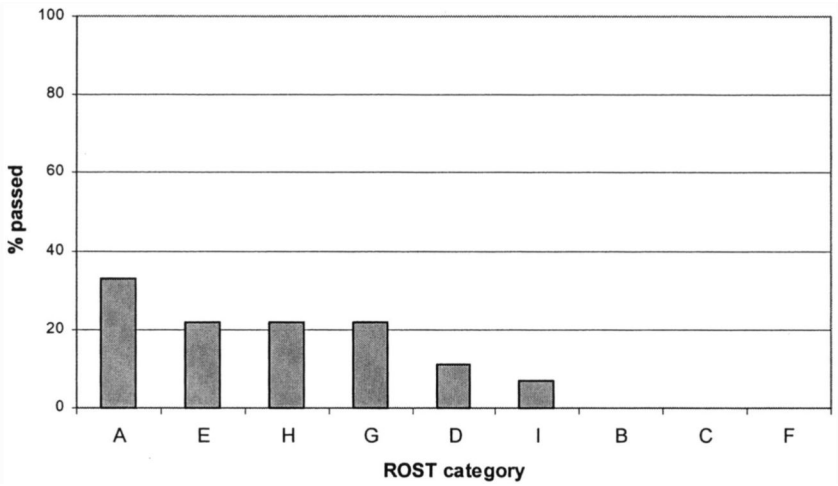

Figure 8.
Percentage of EAL children under five (*y*-axis) passing each of the nine simple (A-I) syntactic categories (labelled on the abscissa). Syntactic categories are arranged in increasing order of difficulty (a lower percentage of children passing)

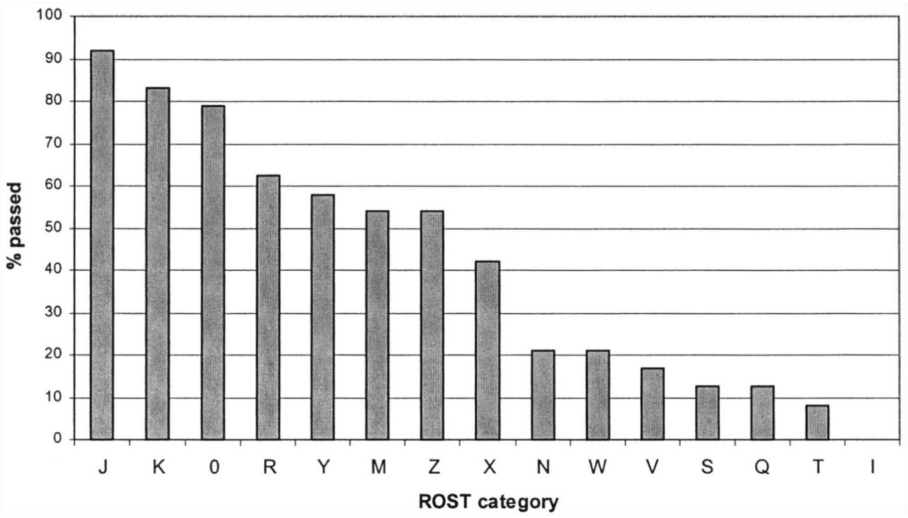

Figure 9.

Percentage of EAL children over five (*y*-axis) passing each of the 15 complex syntactic categories (labelled on the abscissa). Syntactic categories are arranged in increasing order of difficulty (a lower percentage of children passing)

UKPMC Funders Group Author Manuscript

Table 1

Summary of language groups tested (indicated at the top) for age groups (shown in the rows) between 2 and 10 years. Figures in brackets indicate the number of children who met the criterion for assessing the more difficult version of ROST

	Language group						
	Control		CWS		EAL		
Age group	Male	Female	Male	Female	Male	Female	Total
2-2.5 years	9 (9)	3 (3)	—	—	—	—	12 (12)
2.5-3 years	12 (12)	10 (10)	—	—	—	—	22 (22)
3 years	15 (15)	15 (15)	1 (1)	1 (1)	—	—	32 (32)
4 years	13 (13)	16 (16)	1 (1)	1 (1)	5 (5)	4 (4)	40 (40)
5 years	14 (9)	14 (11)	4 (3)	—	12 (6)	11 (5)	55 (34)
6 years	14 (13)	18 (17)	2 (2)	—	20 (10)	17 (9)	71 (61)
7 years	15 (12)	10 (9)	2 (2)	—	16 (12)	12 (10)	55 (46)
8 years	11 (11)	9 (8)	—	—	8 (8)	8 (7)	36 (34)
9 years	15 (15)	13 (13)	6 (5)	—	12 (10)	6 (6)	52 (49)
10 years	17 (17)	14 (14)	7 (6)	3 (2)	11 (10)	10 (10)	62 (58)
Total	135 (126)	122 (116)	23 (20)	5 (4)	84 (61)	68 (51)	437 (387)

UKPMC Funders Group Author Manuscript

Table 2

Syntactic categories indicated in column 2, tested in English (column 3), Turkish (column 4) and Cantonese (column 5) ROST tests

Category	Category/language	English (=26)	Turkish (=13)	Cantonese (=14)
A	Word order (SVO/SOV*)	X	X*	X
B	Article	X		
C	Plural (morpheme)	X	X	
D	Pronoun (he/she)	X	X	
E	Pronoun (his/her)	X		
F	Pronoun (number)	X	X	X
G	Preposition (in/on)	X	X	X
H	Preposition (others)	X		
I	Compound noun	X		
J	Negative	X	X	X
K	DO/IO	X		
L**	Passive	X	X	X
M	Long distance movement	X		
N	Seems	X		
O	Comparative	X	X	X
P**	As (adjective) as	X		
Q	S(sR)VO	X	X (R)SOV+	X
R	SVO(sR)+	X	X S(R)OV+	X SV(SR)+
S	S(oR)VO	X		X
	SVO(oR)+			X SV(OR)+
T	More relative clause	X		
U**	PRO	X		
V	Passive negative	X		X
W	Not as (adjective) as	X		X
X	X but not Y (4 pictures)	X	X	
Y	Not only X but also Y (4 pictures)	X	X	
Z	Neither X nor Y (4 pictures)	X	X	
	Classifier (number)			X
	Ba-structure (word order SOV-marker)			X

The labels used in the text for the syntactic categories are indicated in the left column. An X signifies that the category is tested in ROST for that language. Examples of sentence pairs for each category are given for English at the bottom.

Category	*Example*
A	Word order: The man chases the lady/The lady chases the man
B	Article: A fish/Some fish
C	Plural: Cat/Cats
D	Pronoun (he/she): He drinks/She drinks
E	Pronoun (his/her): His Cat/Her Cat
F	Pronoun (number): He sang/They sang
G	Preposition (in/on): In the house/On the house
H	Preposition (others): On the table/Under the table
I	Compound noun: A car cake/A cake car
J	Negative: The boy is not running

UKPMC Funders Group Author Manuscript

K	DO/IO: The boy gets the girl a cake/The girl gets the boy a cake
L	Passive: The girl is carried by the boy/The boy is carried by the girl
M	Long distance movement: The chair that the boy sees, the girl sits on/The chair that the girl sees, the boy sits on
O	Comparative: The boy is shorter than the girl/The girl is shorter than the boy
P	As (adj) as: The boy is as tall as the girl/The girl is as tall as the boy
Q	S(sR)VO: The boy chasing the girl wears a hat/The girl chasing the boy wears a hat
R	SVO(R): The boy chases the girl who is tall/The girl chases the boy who is tall
S	S(oR)VO: The girl the boy chases, is fat/The boy the girl chases, is fat
T	More relative clause: The boy the girl says is sitting/The girl the boy says is sitting
U	PRO: The girl wants to get the boy a cake/The boy wants to get the girl a cake
V	Passive negative: The girl is not carried by the boy/The boy is not carried by the girl
W	Not as (adj) as: The boy is not as short as the girl/The girl is not as short as the boy
X	X but not Y: The ship but not the car is red/The car but not the ship is red
Y	Not only X but also Y: Not only the car but also the ship is blue/Not only the ship but also the car is red
Z	Neither X nor Y: Neither the ship nor the car is blue/Neither the car nor the ship is red

** Categories dropped from statistical analysis.