Published in final edited form as:

J Psychol Afr. 2012 February 1; 22(2): 155–170.

Speech and Language Disorders in Kenyan Children: Adapting Tools For Regions With Few Assessment Resources

Julie Anne Carter

The Centre for Geographic Medicine Research (Coast), Kenya Medical Research Institute Centre for International Health and Development, Institute of Child Health

Grace Murira, Joseph Gona, and Judy Tumaini

The Centre for Geographic Medicine Research (Coast), Kenya Medical Research Institute

Janet Lees and Brian George Neville

Neurosciences Unit, Institute of Child Health

Charles Richard Newton

The Centre for Geographic Medicine Research (Coast), Kenya Medical Research Institute
Neurosciences Unit, Institute of Child Health University of Oxford Department of Psychiatry

Abstract

This study sought to adapt a battery of Western speech and language assessment tools to a rural Kenyan setting. The tool was developed for children whose first language was KiGiryama, a Bantu language. A total of 539 Kenyan children (males=271, females=268, ethnicity=100% Kigiryama. Data were collected from 303 children admitted to hospital with severe malaria and 206 age-matched children recruited from the village communities. The language assessments were based upon the Content, Form and Use (C/F/U) model. The assessment was based upon the adapted versions of the Peabody Picture Vocabulary Test, Test for the Reception of Grammar, Renfrew Action Picture Test, Pragmatics Profile of Everyday Communication Skills in Children, Test of Word Finding and language specific tests of lexical semantics, higher level language. Preliminary measures of construct validity suggested that the theoretical assumptions behind the construction of the assessments were appropriate and re-test and inter-rater reliability scores were acceptable. These findings illustrate the potential to adapt Western speech and language assessments in other languages and settings, particularly those in which there is a paucity of standardised tools.

Keywords

Child; Assessment; Kenya

The morbid consequences of infectious disease in resource-poor countries have received little attention in the literature, despite the fact that the mechanisms of many of these diseases indicate potential neuro-cognitive impairments in child survivors (Carter, Neville, & Newton, 2003a). For example, *P. falciparum* infection may be one of the major causes of acquired neurological impairment in children in the tropics and has been associated with speech and language deficits (Bondi, 1992; Holding, Stevenson, Peshu, & Marsh, 1999; Meremikwu, Asindi, & Ezedinachi, 1997; van Hensbroek, Palmer, Jaffar, Schneider, & Kwitkowski, 1997) but there have been few detailed descriptions of the speech and

language profile. The outcome of severe disease is further complicated by other health and development factors affecting children in resource-poor countries. Levels of functioning for similar levels of impairment may be influenced by other illnesses, nutrition, social disruption and poverty (Levine, 1999).

There is a lack of appropriate speech and language assessment tools for these settings. In view of the potential scale of such problems in malaria endemic regions (Carter et al., 2005a) and their human and economic costs, appropriate assessment tools are required. Although the use of translated but otherwise unmodified assessment tools is reported in the literature, there are problems associated with this approach (Howard & De Salazar, 1984; Miller, Onotera, & Deinard, 1984). Translation of assessments from the source language to a second, 'target' language is complex and often does not overcome the problem of bias. Practical problems include test directions, which are often technical and difficult to translate; ensuring equivalence of meaning, which necessitates several translations back and forth between languages; the fact that the psychological constructs underlying the assessment are not always universal across cultures and the variation in test-taking behaviour and orientation towards assessment procedures between cultures (Sperber, Devellis, & Boehlecke, 1994). However, the development of completely new tools is often impractical within the time and resource constraints of most research or clinical situations, as it necessitates ethnographically-based research on the language and culture of the intended study population to generate a developmental model to use for assessment (Ochs, 1983; Schieffelin & Ochs, 1986). Alternative approaches aim to adapt existing assessments and include renorming assessments for the target population, using dynamic assessment techniques to assess a child's 'modifiability' and modifying existing, standardised assessments.

The Development of Speech and Language Assessments in Resource-Poor Countries

Advances in the assessment of children with speech and language disorders are usually based on the accumulation of a knowledge base of clinical linguistic theory and normal language acquisition (Owens, 2007). Such a knowledge base does not yet exist in many resource poor countries, where there has often been no research into the developmental sequence of language acquisition and there is little information on the structure of the language. In consequence, it is necessary to adopt a broad approach in developing a theoretical framework that takes account of potential neurological damage and its hypothesised effects on language, while accommodating the constraints set by the context.

Study Context

The tools described in this paper were used to assess speech and language impairments associated with severe falciparum malaria in Kilifi District, a malaria endemic region on the Kenyan coast. The study area consists of rural and peri-urban communities with most of the population originating from the Mijikenda ethnic group. Agriculture is the most important economic activity in the area, engaging 70% of the labour force, mostly in the form of small-scale subsistence farms. The rate of food self-sufficiency is low, half of pre-school children are classified as chronically malnourished, and fewer than 70% of children receive formal primary education and even less progress to secondary education (Carter, 2002). Low levels of maternal literacy also limit children's informal education.

Kigiryama, the mother tongue of the Giryama people, is the largest of the Mijikenda group of central Bantu languages: it was the language of the majority of the target population and the medium of assessment. As a Bantu language, Kigiryama is agglutinative in verb phrase

construction: prefixes are attached to a root verb to specify the subject and object of the verb, tense/aspect and negation and can also be used to mark a relative clause. Kigiryama has 19 noun classes, which are essential for the correct construction of the noun phrase. A number of constituents of the noun phrase must agree with the noun: demonstratives, adjectives, numerals, subject and object prefixes, the relative prefix and the associative marker 'of'. All words in Kigiryama end with a vowel due to the open syllable structure of the language. There are 44 consonant phonemes, including semi-vowels, five vowel phonemes and two basic levels of tone for grammatical distinctions: high and low. Unlike many Bantu languages, Kigiryama does not use tone for lexical distinctions. Printed literature was produced in Kigiryama early in the 20th century but is now obsolete and out of print. A revised orthography has been developed but in practice, the language remains predominantly unwritten.

Constructing a Model for Assessment Development

The behavioural model originates in the linguistic and psychological sciences. A simple linguistic model of language disorders makes a primary distinction between language structure and language use. A detailed version of the behavioural or linguistic model – the Content, Form and Use (C/F/U) model – was proposed by Bloom & Lahey (1978) and further described by Lahey (1988). The C/F/U model is primarily a language production approach, in which language is seen as involving interactions between its meaning (content), shape (form) and function (use).

The mechanisms of action for cerebral malaria (CM) indicate global and diffuse involvement of the brain, suggesting that neurological impairments are likely to be widespread, diffuse and difficult to characterise (Idro, Marsh, John, & Newton, 2010). The mechanisms of other neurological complications of malaria are less clear. Previous studies of CM have cited learning difficulties (Meremikwu et al., 1997), memory impairment (Bondi, 1992) and developmental regression (van Hensbroek et al., 1997). Studies that have specifically investigated elements of cognitive and neuropsychological functioning following CM have variously reported deficits in attentional capacity (Boivin, 2002; Holding et al., 1999), visual-spatial processing (Boivin, 2002), aspects of memory (Boivin, 2002; Dugbartey, Spellacy, & Dugbartey, 1998), executive functioning (Boivin, 2002; Holding et al., 1999) and somatosensory discrimination deficits (Dugbartey et al., 1998). Aphasia is the most commonly reported neuro-cognitive deficit in short-term (Bondi, 1992; Brewster, Kwiatkowski, & White, 1990; Carme, Plassart, Senga, & Nzingoula, 1994; Meremikwu et al., 1997) and long-term studies (van Hensbroek et al., 1997). Hearing loss was reported by van Hensbroek and colleagues (1997), although they give no information on whether it was conductive or sensori-neural. There is no pathological evidence of damage to the cochlea or the VIII nerve as a result of malaria.

One of the original purposes of syndrome-based models was to assist in the determination of probable site of brain damage in the days before neuroimaging became available (Ellis & Young, 1996). In terms of language acquisition, different facets of language are highly interdependent and later language function is more integrated than traditional models would suggest (Bates, 1997). However, in many resource-poor settings, little is known about the behaviour being measured in either the ordered or disordered populations; in such situations, a 'coarse' or more basic approach is recommended.

The C/F/U model provides a broad classificatory framework, which encompasses the universal aspects of language content, form and use. As the original model was developed in the context of a large body of knowledge on normal language development in groups of English-speaking children, it is not possible to reproduce the same level of detail in

Kigiryama. Therefore, we constructed a framework that draws upon the major domains of the behavioural and C/F/U models but maintains a structure that reflects the current level of knowledge of the language (Figure 1).

Our Approach to the Instrument Development

Based on the theoretical framework we propose (as above), the process of assessment development followed procedures designed for the construction of novel assessment instruments (Clark & Watson, 1995; Haynes, Richard, & Kubany, 1995; Smith & McCarthy, 1995; Streiner & Norman, 1995). The current assessments were derived from existing tools or assessment methods; therefore we modified these processes to local conditions. Coastal Kenya has few language assessment resources, so the literature pertaining to relevant assessment models and techniques was drawn from the general African, European and North American literature.

Validity is most powerfully estimated through criterion-related validity (Anastasi & Urbina, 1997; Haynes, Nelson, & Blaine, 1999) but in the rural Kenyan context, standard measures were unavailable. Academic achievement or schoolteacher ratings were also considered an inappropriate means of validation because school attainment may be influenced by factors other than the child's cognitive abilities, such as the quality of teaching, the availability of teaching materials and the presence of a class environment conducive to learning (Connolly, 1998). In addition, many Kigiryama children do not attend school and class numbers are often large, so schoolteachers may not have the opportunity to observe and accurately rate individual children.

We chose an alternative approach to initial instrument development, using the principles of content validity. Content validity is the degree to which the elements of an assessment battery are relevant to and representative of the target construct (Haynes et al., 1995). The construct validity of an assessment instrument is limited by the degree to which its elements were initially designed to measure the targeted construct (Haynes et al., 1999), thus making the content validation or developmental phase of assessment design particularly important.

Aims of the Study

This study presents an approach to the development of speech and language assessment tools for children, using research on the morbid consequences of severe falciparum malaria as a model. We aimed to develop assessments appropriate to a rural region of Kenya and describe the process we followed, which we consider to be applicable to other settings with few assessment resources. Our specific research question was: could the C/F/U model be used to develop a culturally appropriate assessment of an African language?

Method

Participants and Setting

Our target population was children between 6 and 9 years of age, since this is the age by which most children living in a malaria endemic area will be exposed to severe falciparum malaria (see Table 1).

Permission for the study was granted the National Ethics Board of Kenya and the Institute of Child Health, London. Written consent was obtained from the parent or guardian, and verbal assent from the child to participate in the study.

Data Collection and Procedure

Initial tool development, including a pilot study, consisted of a battery of six assessments: receptive language (Holding et al., 1999), syntax (Armstrong & Ainley, 1989), lexical semantics (Crystal, 1992), higher level language, pragmatics (Andersen-Wood & Smith, 1997; Holding et al., 1999; Lees & Urwin, 1997) and phonology. The origins of and alterations to each assessment are described in more detail in Appendix 1. All of the target constructs illustrated in Figure 1 were included in the speech and language battery (Table 2). The aims of the pilot study were twofold. Firstly, to establish the cultural validity and suitability of the initial assessment battery for children in this context and to collect data to examine the battery's discriminatory ability. Second, to conduct a preliminary investigation into the presence of persisting speech and language impairments associated with severe malaria in children. Although long-term studies of impairments associated with malaria suggest that speech and language deficits persist, there are few reports and none that have employed detailed instruments. Fifty-two children (25 previously admitted to hospital with severe malaria and 27 unexposed to the disease) aged 8-9 years were recruited to the pilot study. Twelve were female and 13 were male. The 27 comparison children had a median age of 8 years 11 months (range 7 years 8 months to 10 years 4 months). There were 11 females and 16 males. Detailed descriptions of the results are reported elsewhere (Carter, Murira, Ross, Mung'ala-Odera, & Newton, 2003b), and the results of the study reported here are in terms of the data used for making refinements and alterations to the tools in stage 4.

Each assessment item was examined with reference to its appropriateness for each facet of the target construct. The interpretability of the items was also examined, focussing on their clarity and specificity, particularly looking for ambiguity, double-barrelled questions, jargon and lengthy questions. Each item was reviewed then tested on a sample of ten children of the same age as the target population. We measured the face validity of each assessment because the participants' perceptions of a test's difficulty or appropriateness will affect their level of co-operation and the acceptance of the instrument (Streiner & Norman, 1995). This is particularly important in a context in which there is a paucity of developmental assessments and unfamiliarity with testing procedures and situations. A local schoolteacher and four Giryama research colleagues working on unrelated projects were asked to rate battery items for face validity.

The assessments in this battery were adapted from existing UK assessment instruments and tools designed for a previous study examining cognitive impairments associated with malaria in the same geographical area (Holding et al., 1999). Although the content of the Kenyan assessments was maintained, the content of the UK assessments was altered extensively. The item pool was altered with the advice of Kigiryama-speaking members of the assessment team (GM, JG, JT) and local teachers. The response formats of assessments based upon existing instruments were maintained.

The response formats and scoring schemes of the original assessments were examined for their applicability to the local language and culture. Novel instructions were developed for each assessment with members of the assessment team and a local schoolteacher.

Assessments were developed in conjunction with a team of mother tongue speakers of the assessment language who had grown up in the area (GM, JG, JT). Further training in the theoretical and practical aspects of assessment was provided by the first author. The assessments were designed to measure spoken language only. Written language was not included for three reasons. First, Kigiryama is predominantly an unwritten language. Second, impairments in spoken language would be potentially more disabling in a community with an oral tradition and third, as many children had never attended school, they would be unlikely to know how to read and write.

Data Analysis

The internal consistency of the speech and language assessments was established using Cronbach's alpha. Twenty children were recruited to collect data on test-retest and inter-rater reliability. Retesting was carried out two weeks after the first administration of the assessment by the same assessor, at the same venue and where possible, at a similar time of day. Inter-rater reliability was carried out at the time of the first administration of the assessments. Test-retest and inter-rater reliability was measured using a method proposed by Bland & Altman (1986): the mean difference between the two measures, as an estimate of the overall bias and the 95% limits of agreement as an indication of the range of variability in individual scores.

Construct Validation

Two specific techniques – age differentiation and principal components analysis – were employed to provide initial evidence of the construct validity of the assessments. These procedures were performed on a group of 179 children unexposed to severe malaria who were part of a study examining the morbid consequences of severe falciparum malaria (Carter et al., 2005a, 2005b), in order that the large numbers would increase the reliability of the results.

Performance on speech and language assessments would be expected to exhibit a progressive increase with advancing age, although the increase in this group may be less marked in some facets of language because of the age of the children. Yearly age bands were used because it is difficult to ascertain month and day of age in this community. Means were calculated for normally distributed data and medians for skewed data (transformations could not be found that produced a normal distribution of these data).

Discriminatory Ability

The ability of the assessments to identify speech and language problems was tested in the study mentioned above, which investigated the morbid consequences of severe falciparum malaria. We recruited 152 children who had suffered cerebral malaria and 156 children who had suffered malaria and complicated seizures 2-9 years prior to the study. These children were compared with 179 children who were unexposed to either complication of malaria. The methods and results are reported in detail elsewhere (Carter et al., 2005a, 2005b; Carter et al., 2006) and are described here in terms of the discriminatory ability of the tools.

Results

Table 3 presents a summary of the scores for receptive language, syntax, higher level language, lexical semantics, pragmatics, phonology and word finding.

The final battery comprised seven assessments: receptive language (original assessment changed to an adaptation of the Test for the Reception of Grammar (Bishop, 1983); syntax (new scoring system adapted from the Renfrew Action Picture Test (Renfrew, 1988)); lexical semantics (minor alterations to original); higher level language (major alterations to reduce the number of different items and increase the number of questions per item); pragmatics (original assessment changed to an adaptation of the Pragmatics Profile of Everyday Communication Skills in Children (Dewart & Summers, 1995)); phonology (original assessment changed to a shorter, picture-based assessment) and word finding (a new assessment based on the Test of Word Finding [German, 1986]). The specific reasons underlying alterations made to each test are presented in Appendix 2.

Test-retest Reliability

The mean difference in scores indicates a high level of repeatability in all of the assessments (Table 4). The highest level of individual variability relative to the maximum assessment score was on the receptive grammar assessment, the direction suggestive of a practice effect. Ascertaining the clinical acceptability of the limits of agreement is difficult in the context of novel assessments but the figures are suggestive of an acceptable level of variability as they represent less than 10% of the mean score of the main study unexposed group.

Inter-rater Reliability

Statistical measurement of inter-rater reliability indicates an acceptable level of congruence between assessors: again, the mean difference in scores is less than 10% of the mean score for the main study unexposed group (Table 4).

Internal Consistency

Four of the assessments consisted of aggregated measures: the receptive grammar, syntax, higher level language and pragmatics assessments. The alpha coefficients (receptive grammar=0.8; syntax=0.83; higher level language=0.76; pragmatics=0.8) indicated a high level of internal consistency.

Further Validation: Age Differentiation

A progressive increase in scores with advancing age is shown in the assessments of receptive grammar, syntax, higher level language, word finding and pragmatics (tables 5 and 6). The largest increase in scores is seen in the higher level language assessment, reflecting the fact that this test targets advanced functions more likely to be present in older children. Scores remain almost static over the age range for the lexical semantics assessment, possibly reflecting the fact that many children do not attend school, thus do not exhibit increases in vocabulary as a function of education. Indeed, when non-schoolers were removed from the analysis, the content word component did exhibit a progressive increase with age, with the exception of the 9 year-olds category (6 years: mean=0.59 (SD=0.07); 7 years: 0.61 (0.05); 8 years: 0.63 (0.05); 9 years: 0.62 (0.06)). The function word scores remained the same. The median phonological assessment score and the inter-quartile range across the age groups is the maximum possible score, suggesting that the Kigiryama phonological system, as measured by these tools, is fully developed by 6 years of age.

Principal Components Analysis

Four components accounted for 81% of the total variance of the measured variables. All (except higher level language) are pure variables with loadings of 0.4 or greater on only one component (table 7: shown in bold font). Examination of the items on factor 1 indicates a 'general language ability' component with positive loadings on all variables. Factor 2 seems to specifically relate to 'vocabulary' or aspects of word meaning, with a high loading for expressive vocabulary and positive loadings for receptive vocabulary and word finding. Factor 3 relates to phonology and the fourth factor indicates a pragmatics component. The covariance on factor 1 provides evidence for the convergent validity of the language measures. The other three factors relate to distinct components of the assessment framework from the three major categories of content, form and use: although they formed distinct factors, these variables still had positive loadings on the 'general language ability' factor.

Discriminatory Ability

The speech and language results of the groups exposed to severe malaria compared to those unexposed are presented in detail elsewhere (Carter et al., 2006). Briefly, significant differences in scores between the CM and unexposed groups were found in the lexical

semantics, higher level language and pragmatics assessments and between the M/S and unexposed groups in the pragmatics and phonological assessments. Table 3 shows that the groups exposed to severe malaria had lower mean/median scores than the unexposed group in all the assessments except for phonology (also lexical semantics and word finding for the M/S group) and lower minimum scores in all assessments except for the higher level language assessment (also lexical semantics for the M/S group). The distribution of scores was broader in the exposed groups compared to those unexposed to severe malaria on most assessments.

Discussion

This study describes the development of speech and language assessments suitable for children living in a rural region of coastal Kenya. The validity of research and clinical findings is dependent on the reliability, validity and sensitivity of the assessment methods and tools employed. The development of these assessments was carried out following the principles of content validation and preliminary measures of construct validity, including age differentiation and principal components analysis. The results suggest that the theoretical assumptions behind the construction of the assessments were appropriate. Evidently, further use of the battery will depend on a continuing process of validation and a detailed evaluation of the current findings to identify and improve weaker test items and methods. The assessments attained an acceptable re-test and inter-rater reliability, although further studies will provide more evidence on the clinical acceptability of test-related variance and possibly allow the design of more complex frameworks for the study of language disorders in Kigiryama.

Using the tools to compare groups of children exposed to severe malaria and a group unexposed to the disease suggested that the language tools have some discriminatory ability. Poorer performance was found in the exposed groups on most assessments, although the data suggest that falciparum malaria is associated with language impairment in a subgroup of children, rather than affecting the majority of survivors. Due to the length of time since the presumed insult, these children are assumed to be past the acute stage of the disorder, therefore the expected pattern would be more subtle impairments in larger groups of children or impairments in a subgroup of children. Therefore these results make sense and suggest that these assessment tools are able to identify rural Kenyan children with language disorders.

More broadly, these findings illustrate a methodology for the adaptation of speech and language assessments in other languages and settings, particularly those in which there is a paucity of standardised tools. Some investigators argue that assessment tools reflect the values, knowledge and communication strategies of their culture of origin, which are so specific as to preclude adaptation to other cultures (Greenfield, 1997). However, valid assessment of these outcomes necessitates the demonstration of cultural equivalence (whether the underlying constructs have similar meanings within and across cultural groups) and functional equivalence (the extent to which cognitive/language ability operates in the same manner between cultures) of the psychological constructs underlying these tests (Pérez-Arce, 1999).

Implications for Neuropsychological Assessment in the African Context

There are few cultural specific tools for the neuro-cognitive assessment of children in East Africa, particularly assessments of speech and language. We have described a process that produces a cultural and language specific tool that has excellent measures of construct, inter-rater and re-test reliability, which could be used for the development of other language assessments within Africa. These tools will help in delineating the extent and modalities of

verbal cognitive impairment following brain insults. Recent evidence from functional magnetic resonance imaging suggests that literacy and education influence problem-solving. The development of language tools to assess illiterate populations may provide further insights into cognitive function.

Limitations of the Study and Suggestions for Further Research

Adopting this approach to the development of assessments is time and resource-consuming (although less so than other approaches) and is often underestimated at the proposal development stage. Assessment of language will necessarily reflect the values, knowledge and communication strategies of their culture of origin/adaptation (Greenfield, 1997). Thus the development of assessments using a similar process needs to recognise the influence of cultural beliefs on this process. We used children exposed to severe malaria to test the tools on children with language impairments. Other causes of language impairments may have features which are not easily accessible by this process of adaption, e.g., influences of social communication on the assessments of language.

Conclusion

This study has shown that tools can be modified for another setting in a way that does not forfeit content or construct validity. We believe that developing a critical mass of researchers/clinicians across resource-poor regions, who are following this or similar approaches to assessment, will facilitate the systematic measurement and comparison of language outcomes of diseases that affect millions of children each year and the development of interventions that can be adapted across different settings.

Acknowledgments

Thanks to Dr. Penny Holding for her useful comments on an earlier draft of this paper. J. A. Carter (059336) and C. R. J. C. Newton (050533) are supported by The Wellcome Trust, UK. This paper is published with the permission of the Director of KEMRI.

Appendix 1: Origins of the Initial Assessments in the Battery

Assessment	Developed from	Item development
Receptive language	Test of receptive vocabulary used in a previous study in Kilifi (Holding et al., 1999)	The original test consisted of 40 items, ranging from common nouns to abstract words. Each stimulus word is accompanied by a selection of four black and white drawings, from which the child selects the correct one by pointing. For the current study, items that had proved problematic during Holding and colleagues' study were removed to leave 23 items. In addition, more complex measures of receptive language were included. Based upon the facets of the theoretical framework, a measure of syntax - understanding of simple and complex commands - was also included in the final score. The response format was question and answer, using a simple 1/0 scoring scheme.
Expressive grammar (Syntax)	South Tyneside Assessment of Syntactic Structures (STASS) (Armstrong and Ainley, 1989)	The STASS consists of a series of pictures and accompanying questions, each one presented to the child to elicit one or more grammatical structures. The child's responses are recorded verbatim to be analysed later by the assessor for word, clause and phrase level structures and auxiliary verbs, pronouns and prepositions. To determine the grammatical elements that would be appropriate to local children in the target age group, a series of meetings was held with a local schoolteacher. His suggestions were supplemented by elements decided upon by rational deduction of the hypothesised level of syntactic development in Giryama children of 6-9 years of age, knowledge of the language and comments from the assessors. The elements targeted in the assessment were prepositions, plurals,

Assessment	Developed from	Item development
		conjunctions, comparatives, adverbs, adjectives, past tense, locatives, negatives, WH-questions ('where', 'who'), passives and future tense. New pictures were drawn according to the style of the STASS, one picture per question with a total of 20 questions. The response format was question and answer, using a simple 1/0 scoring scheme.
Expressive vocabulary (Lexical Semantics)	Profile in Semantics – Lexicon (PRISM-L) (Crystal, 1992)	The PRISM-L is a naturalistic method of assessing vocabulary size and breadth from the child's spontaneous language. Giryama children are unfamiliar with testing situations and test materials: therefore we tried to reduce the number of formal, picture-based tasks; expressive vocabulary was considered to be one of the assessments most amenable to a naturalistic style of testing. Utterances from a spontaneous language sample are transcribed and classified by semantic category or grammatical usage. The types - the total number of <i>different</i> words the child has said – are divided by the tokens - the total number of words the child has spoken – to calculate the 'type-token ratio' for minor and major words. The format of the profile was maintained for the current study, although the function words sections were altered to reflect the different grammatical categories in Kigiryama. Three methods were used to generate spontaneous language production: picture description, conversation and storytelling. A picture book depicting various animals was used for picture description and the spontaneous language session was recorded using a dictaphone for later transcription.
Higher level language	Subtests of the Clinical Evaluation of Language Fundamentals – Revised (CELF-R) and items from non-standardised assessments of higher level language	Higher level language can be described as the level at which the basic aspects of language integrate and interact. The persistence of impairments in this aspect of language is indicative of a poor prognosis in acquired childhood aphasia (Lees, 1997). In addition, assessments of more complex language usage can be useful in identifying subtle language deficits and have been used in the evaluation of residual deficits following acquired childhood aphasia (Crosson, 1996). As it was difficult to know the types of questions that would be sensitive and applicable to Giryama language and culture, a wide variety of question types was identified from a range of standardised and non-standardised assessments. Semantic tasks such as word associations, word classes and formulated sentences from the CELF-R were adapted for use in a Giryama context. In addition, discussions with a local schoolteacher and the assessors generated questions testing synonyms, homonyms, similarities and differences, antonyms, word definitions, figurative language, convergent and divergent semantic tasks and a range of inferential reasoning and problem solving questions. The response format was question and answer, using a simple 1/0 scoring scheme.
Pragmatics	An analysis of functional errors used previously in Kilifi (Holding et al., 1999); observational components and parental questionnaire (Andersen-Wood & Smith, 1997; Lees & Urwin, 1997)	The assessment of pragmatics was divided into three parts. The first part comprised an analysis of functional errors used in a previous study. The analysis comprised the identification of linguistic nonfluencies, revisions, delayed responses, nonspecific vocabulary, inappropriate responses, poor topic maintenance and need for repetition from a spontaneous language sample. An observational component formed the second part of the assessment. Discussions were held with a local schoolteacher on cultural norms: The child was observed during the assessment session for evidence of the following skills: appropriate physical proximity to the communicative partner; appropriate eye contact; appropriate facial expression; appropriate amount of presupposition; not easily distracted; use of language appropriate to the situation and people present and amount of language. The final part of the assessment comprised a parental questionnaire to investigate the child's range of communicative intents. The initial list was drawn from Lees & Urwin (1997) and Andersen-Wood and Smith (1997) and supplemented by ideas from a local schoolteacher

Assessment	Developed from	Item development
Phonology	Novel assessment	<p>and members of the assessment team.</p> <p>The test of phonology aimed to assess the integrity of the child's phonological system. Where possible, the use of formal, picture-based tasks was minimised: phonology was considered amenable to a more naturalistic style of testing. The first one hundred different words from the child's recorded spontaneous language sample were transcribed and profiled by the candidate using the International Phonetic Alphabet on a chart of all the consonant phonemes used in Kigiryama. The language has no word or syllable final consonant phonemes; word initial and word medial (syllable initial) phonemes were assessed separately. A phoneme was assessed as 'correct' if the child produced it correctly in all words in which it featured in either the initial or medial position.</p>

Appendix 2: Refinements, Alterations and Additions to the Original Assessment Battery

Original assessment	Altered to	Refinements, alterations and additions
Receptive language	Receptive grammar, based on the Test for Reception of Grammar (TROG) (Bishop, 1983)	<p>The distribution of both sections (receptive vocabulary and simple/complex commands) of the initial receptive language assessment suggested that the assessment was too easy. Fifty-seven percent of children from the groups exposed and unexposed to severe malaria achieved one of the top three scores (21–23) and none of the children scored less than 11 in receptive vocabulary. Sixty-seven percent of children from the unexposed group and 38% of children from the exposed group achieved the highest score (13) on the commands section. Consequently, it was decided that a more complex measure of receptive grammar was required for the main study.</p> <p>The Test for Reception of Grammar (TROG) (Bishop, 1983) provided the basis for a new receptive grammar assessment. To alter the TROG for use in this community, we first identified any items irrelevant to Kigiryama and conversely, any items relevant to the language not included in the English version, with members of the assessment team and a local schoolteacher. Items included in the original TROG and removed from the current receptive grammar assessment were the third person pronoun ('he' and 'she' are both represented by the pronoun 'a' in Kigiryama) and some of the complex sentence structures. The singular/plural item was expanded to take account of Kigiryama noun classes and the comparative/absolute item was altered to become a simple comparative. The content of all the items was altered for cultural validity, although the principles of using a restricted vocabulary, easily recognisable pictures and lexical and grammatical distractors were maintained.</p> <p>The assessment was designed primarily to provide a quantitative measure of this aspect of comprehension. However, a set of vocabulary cards was drawn to provide a basic indication of the nature of any impairment identified (grammatical or lexical impairments).</p> <p>A local artist drew the pictures: each one was piloted with children (n=22) from the target population (including children not attending school, who would have less exposure to picture material) for level of recognition. If 80% of children recognised the picture, it was accepted; otherwise it was redrawn or discarded. This process of recognition testing was repeated for each assessment that used picture stimuli.</p> <p>The final version of the assessment comprised an example picture, a practice picture and 34 items measuring 17 aspects of grammar of increasing complexity from single nouns, verbs and adjectives to post-modified subjects and relative clauses.</p>

Original assessment	Altered to	Refinements, alterations and additions
Syntax	New scoring system adapted from the Renfrew Action Picture Test (Renfrew, 1988)	<p>The scores on the syntax assessment were normally distributed and the pilot study participants generally responded well to the question-answer format. However, several of the items were unsuccessful and were thus removed or altered. First, the use of pictures to test prepositions was not effective. The prepositions 'infront' and 'behind' were particularly problematic: 24% of children responded 'on top' and 26% said 'under' for each one respectively, possibly due to difficulties understanding the manner in which depth is represented in two dimensions (Serpell, 1972). Consequently, real objects (a cup and banana) were used to test prepositions in the main study, by questioning the child on the relative positions of the two items. Pictures for past and future tense questions were also confusing and replaced with questions modelling the target, indirectly related to the pictures. For example, the picture for the 'passives' question showed two boys, one holding a mango and the other, a banana. After the child has given the answer 'this boy was given a banana', the assessor asks 'what will he do with it?' to elicit the future tense.</p> <p>The 1/0 scoring system was also altered to allow for more complex analysis. The scoring system from the Renfrew Action Picture Test (RAPT) (Renfrew, 1988) formed the basis of the new system, whereby answers that are not correct but show some evidence of the development of syntactic principles are awarded a score. For example, a child who says 'caught' instead of 'caught' in English knows the principle of using the past tense but not the irregular exceptions. To devise the new scoring system, all of the answers given by children in the pilot study were listed and the types and frequency of responses analysed to determine common response-types for each syntactic element.</p> <p>The final version of the assessment comprised five preposition questions and fifteen picture-based items ranging from negation and comparatives to grammatical constructs specific to Kigiryama, such as the impact of noun classes on plurals. There was an example item and a practice item for the preposition section and the picture-based section.</p>
Lexical semantics	Not altered	<p>The spontaneous language tasks resulted in large samples of conversation and storytelling. The scores for the function word and content word elements of the assessment were broadly distributed in both exposure groups, indicating that the assessment discriminates at different points along the continuum of the construct therefore the format of the assessment was maintained. In order to achieve a representative score, it was decided that each child should produce a minimum of 100 words (tokens) in the language sample for the score to be calculated.</p> <p>A booklet of guidelines outlining which words to include in each category was produced, using response data from children in the pilot study.</p>
Higher level language		<p>The initial higher level language assessment was designed to be over-inclusive because it was difficult to know the types of questions that would measure this construct in Kigiryama. An analysis of item endorsement suggested that some of the items did not discriminate between the two groups.</p> <p>In order to increase the reliability and homogeneity of the main study assessment, the number of different items was reduced and the number of questions per item increased. To maintain an assessment of reasonable length, five items were selected for the main study assessment, each of which comprised at least four questions. The homonyms, definitions, similarities/differences and sentence formulation tasks were selected on the basis that they were each chosen by a local schoolteacher as the items most familiar to schoolchildren of the same age as the target population, had broad response distributions and varying levels of difficulty. The categorisation task, which required children to name as many items within a group (for example, animals) as possible within one minute, was selected, as word fluency tasks are often sensitive to word finding difficulties (Crosson, 1996).</p> <p>The unexposed group attained a low mean score in the pilot study,</p>

Original assessment	Altered to	Refinements, alterations and additions
		<p>which suggested that in addition to the issue of individual item difficulty, lack of familiarity or understanding of test requirements may have contributed to poor scores. Therefore, a test booklet was devised to include detailed explanations of each task in addition to examples and in some instances, discussion of the task with the child to establish his/her understanding of its requirements. Pictures were included with the homonyms and similarities/differences tasks, to be used if the child found this useful.</p> <p>An analysis of item endorsement suggested that some of the items did not discriminate between the two groups. The figurative language and convergent semantic tasks drew the same answer from almost all of the children in each exposure group. The response distributions for the synonyms task were also very similar between the groups and the proportion of children with a score of '0' indicated that this task was too difficult. Over 70% of children attained a '0' score on the antonyms task, again indicating an inappropriate level of difficulty.</p>
Pragmatics	Pragmatics Profile of Everyday Communication Skills in Children (Dewart and Summers, 1995)	<p>In the first section of the pragmatics assessment there was no evidence of a difference between the groups, although the direction of the association suggested poorer performance among children from the exposed group.</p> <p>The other sections both proved to be problematic. The observational section was designed to assess the child's ability to adapt to the needs of the listener and his/her awareness and use of conversational devices and status relations. However, it was difficult to obtain a representative measure of these pragmatic behaviours, first because the length and setting of the observation was limited and secondly because several of the behaviours were affected by the situation or cultural norms. For example, traditional Giryama families teach their children that it is disrespectful to make eye contact with adults who are in an authoritative position. Also, the fact that the child knew there were strangers present immediately altered his/her behaviour. Consequently, the results of the observation were not deemed to be a representative measure of these aspects of pragmatics.</p> <p>The parental questionnaire, investigating the child's range of communicative intents, was also problematic to administer. The format was successful in that most parents were responsive and gave detailed and informative answers to questions they understood. However, some of the questions were very specific and despite the use of examples, many parents commented that they were obscure or confusing.</p> <p>As the parental questionnaire format was successful, a broad-based, accessible questionnaire was designed to replace all three sections, based on the Pragmatics Profile of Everyday Communication Skills in Children (Dewart and Summers, 1995).</p> <p>The Pragmatics Profile was adapted to the cultural context of the current study by the assessment team, with several consultations with a local schoolteacher. The current profile was designed to be quantifiable (thus referred to as a 'questionnaire') for ease of analysis. The questioner seeks to engage the parent in discussion about various aspects of the child's pragmatic skills. The parent's answers are recorded on a response sheet and later scored on a 0-5 scale.</p> <p>The fourth section of the original Pragmatics Profile – pragmatic skills in context – was omitted for several reasons. First, to reduce its length and second, because too little is known of the current context to make an assessment of appropriate contextual variation. The adapted profile comprises three sections with twenty questions, representing the major language use constructs described by Lahey (1988): social functions, conversational devices, adapting to the needs of the listener and linguistic contexts. Items are presented in the context of everyday skills, making questions more relevant to parents and providing information on possible service requirements. To ensure that the scores reflected cultural expectations of</p>

Original assessment	Altered to	Refinements, alterations and additions
Phonology	Shortened to 40-word picture-based assessment	<p>pragmatic abilities in the Giryama community, a scoring system was developed by the assessment team, based on the format of a version of the Profile quantified for use with children with Landau Kleffner syndrome (Hand, 1996).</p> <p>The pilot results indicated that the phonological assessment was time-consuming. A qualitative analysis of the results indicated a subset of phonemes that appeared frequently in words and another subset that was rarely used by any of the children. In order to reduce the transcription time and target commonly used phonemes, a 40-word picture-based assessment was devised. This assessment was also used as a word finding assessment (see below). To attain a score of '1' for a phoneme in a particular position, the child was required to produce it correctly in each word it appeared in that position. The methodology meant that less common phonemes may only be included in one of the 40 words, thus only assessed once. In consequence, if a child did not produce the correct word, it would not be possible to assess that particular phoneme. As a result, it was decided to use the percentage of all the phonemes used that were produced correctly to score the assessment.</p>
Word finding	Test of Word Finding (TWF) (German, 1986)	<p>The test procedure and scoring system was based on the TWF, measuring both response accuracy and delay. Each accurate response was awarded a score of '1'. A delayed response was categorised as a response >4 seconds after the presentation of the picture: an estimated time was used as it was thought that the use of a stopwatch would interfere with the fluidity of the assessment procedure. A comprehension test was not included due to time constraints: as the words had been selected from a frequency list generated from the target population, it was assumed that the majority of children would know them.</p> <p>The persistence of word finding difficulties has been found to be an indication of a poor prognosis in acquired childhood language disorders (van Hout, Evrard, & Lyon, 1985). A qualitative analysis of the pilot study language samples indicated that 46% of children with a history of cerebral malaria showed signs suggestive of anomia. Therefore, a word finding assessment was devised, based on the TWF (German, 1986). The word finding assessment developed for the current study was based on the principles of the TWF but had a simplified structure and format, since this assessment was only one part of a relatively long battery.</p> <p>The first stage of assessment development was to construct a vocabulary frequency list from the target population. Sixty children between the ages of 6 and 9 years were asked to name as many words as possible. Few verbs, adjectives, adverbs or abstract nouns were included in the list: any were discarded as they would be difficult to present pictorially. The first twenty and last twenty nouns were selected from the frequency list and analysed for phonemic content, so the assessment would also fulfil the requirements of the phonological assessment.</p> <p>Pictures representing the 40 nouns were drawn by a local artist. The process of recognition testing was carried out as described above, although the accepted recognition level was raised to 90% because the test requires children to name the word depicted quickly and without the cues included in other assessments, meaning that each picture had to be immediately recognisable. The test procedure and scoring system was based on the TWF, measuring both response accuracy and delay. The assessment was tape-recorded in order that transcription and analysis for the phonological assessment could be carried out after the administration of the battery.</p>

References

- Anastasi, A.; Urbina, S. Psychological testing. 7th ed. Upper Saddle River, NJ; Prentice-Hall: 1997.
- Andersen-Wood, L.; Smith, BR. Working with pragmatics. Winslow Press; Bicester: 1997.
- Armstrong, S.; Ainley, M. South Tyneside assessment of syntactic structures. STASS; Northumberland: 1989.

- Bates E. Origins of language disorders: A comparative approach. *Developmental Neuropsychology*. 1997; 13:447–476.
- Bishop, DVM. Test for reception of grammar. University of Manchester; 1983.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1:307–310. [PubMed: 2868172]
- Bloom, L.; Lahey, M. Language development and language disorders. John Wiley; New York, NY: 1978.
- Boivin MJ. The effects of early cerebral malaria on cognitive ability in Senegalese children. *Journal of Developmental and Behavioral Pediatrics*. 2002; 23(5):353–364. [PubMed: 12394524]
- Bondi FS. The incidence and outcome of neurological abnormalities in childhood cerebral malaria: A long-term follow-up of 62 survivors. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1992; 86:17–19. [PubMed: 1566292]
- Brewster DR, Kwiatkowski D, White NJ. Neurological sequelae of cerebral malaria in children. *Lancet*. 1990; 336:1039–1043. [PubMed: 1977027]
- Carne B, Plassart H, Senga P, Nzingoula S. Cerebral malaria in African children: Socioeconomic risk factors in Brazzaville, Congo. *American Journal of Tropical Medicine and Hygiene*. 1994; 50:131–136. [PubMed: 8116802]
- Carter, JA. Epilepsy and developmental impairments following severe malaria in Kenyan children: A study of their prevalence, relationships, clues to pathogenesis and service requirements (Unpublished doctoral thesis). University of London; United Kingdom: 2002.
- Carter JA, Neville BGR, Newton CRJC. Neuro-cognitive sequelae of acquired central nervous system infections in children: A systematic review. *Brain Research Reviews*. 2003a; 43:57–69. [PubMed: 14499462]
- Carter JA, Murira GM, Ross AJ, Mung'ala-Odera V, Newton CRJC. Speech and language sequelae of severe malaria in Kenyan children. *Brain Injury*. 2003b; 17(3):217–224. [PubMed: 12623498]
- Carter JA, Mung'ala-Odera V, Neville BGR, Murira G, Mturi N, Musumba C, Newton CRJC. Persistent neuro-cognitive impairments associated with severe falciparum malaria in Kenyan children. *Journal of Neurology, Neurosurgery and Psychiatry*. 2005a; 76:476–481.
- Carter JA, Ross AJ, Neville BGR, Obiero E, Katana K, Mung'ala-Odera V, Newton CRJC. Developmental impairments following severe falciparum malaria in children. *Tropical Medicine and International Health*. 2005b; 10(1):3–10. [PubMed: 15655008]
- Carter JA, Lees JA, Gona J, Murira G, Rimba K, Neville BGR, Newton CRJC. Severe falciparum malaria: A common cause of acquired childhood language disorder. *Developmental Medicine and Child Neurology*. 2006; 48:51–57. [PubMed: 16359594]
- Clark L, Watson D. Constructing validity: Basic issues in objective scale development. *Psychological Assessment*. 1997; 7:309–319.
- Connolly, KJ. Nutrition, Health and Child Development. Research Advances and Policy Recommendations. A joint publication: PAHO, The World Bank and the Tropical Metabolism Research Unit of the University of the West Indies; Washington, DC: 1998. Mental and behavioral effects of parasitic infection. PAHO
- Crosson, B. Assessment of subtle language deficits in neuropsychological batteries: Strategies and implications. In: Sbordone, RJ.; Long, CJ., editors. Ecological validity of neuropsychological testing. GR Press/St Lucie Press; City, FL: 1996. ??–??
- Crystal, D. Profiling linguistic disability. 2nd ed. Whurr; London, England: 1992.
- Dewart, H.; Summers, S. The pragmatics profile of everyday communication skills in children. NFER-Nelson; Windsor: 1995.
- Dugbartey AT, Spellacy FJ, Dugbartey MT. Somatosensory discrimination deficits following pediatric cerebral malaria. *Am J Trop Med Hyg*. 1998; 59:393–6. [PubMed: 9749631]
- Ellis, AW.; Young, AW. Human cognitive neuro-psychology. A textbook with readings. Psychology Press; Hove: 1996.
- German, DJ. Test of word finding. National College of Education, DLM Teaching Resources; Allen, TX: 1986.

- Greenfield PM. You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist*. 1997; 52(10):1115–1124.
- Hand, K. Can we measure response to treatment in children with epileptic aphasia (Landau-Kleffner Syndrome)? (Unpublished master's thesis). Department of Human Communication Science, University of London; UK: 1996.
- Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*. 1995; 7:238–247.
- Haynes, SN.; Nelson, K.; Blaine, DD. Psychometric issues in assessment research. In: Kendall, PC.; Butcher, JN.; Holmbeck, GN., editors. *Handbook of research methods in clinical psychology*. 2nd ed. John Wiley; New York, NY: 1999.
- Holding PA, Stevenson J, Peshu N, Marsh K. Cognitive sequelae of severe malaria with impaired consciousness. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1999; 93:529–554. [PubMed: 10696414]
- Howard DP, De Salazar MN. Language and cultural differences in the administration of the Denver Developmental Screening Test. *Child Study Journal*. 1984; 14:1–9.
- Idro R, Marsh K, John CC, Newton CR. Cerebral malaria: mechanisms of brain injury and strategies for improved neurocognitive outcome. *Pediatr Res*. 2010; 68:267–74. [PubMed: 20606600]
- Lahey, M. *Language disorders and language development*. Macmillan; New York, NY: 1988.
- Lees, J.; Urwin, S. *Children with language disorders*. 2nd ed. Whurr; London, England: 1997.
- Levine, MD. Neurodevelopmental variation and dysfunction among school-aged children. In: Levine, MD.; Carey, WB.; Crocker, AC., editors. *Developmental-behavioural pediatrics*. 3rd ed. W. B. Saunders; Philadelphia, PA: 1999.
- Meremikwu MM, Asindi AA, Ezedinachi E. The pattern of neurological sequelae of childhood cerebral malaria among survivors in Calabar, Nigeria. *Central Africa Journal of Medicine*. 1997; 43:231–234.
- Miller V, Onotera RT, Deinard AS. Denver Developmental Screening Test: Cultural variations in Southeast Asian children. *Journal of Pediatrics*. 1984; 104:481–482. [PubMed: 6707804]
- Ochs, E. Cultural dimensions of language acquisition. In: Ochs, E.; Schieffelin, BB., editors. *Acquiring conversational competence*. Routledge and Kegan Paul; London, England: 1983.
- Owens, RE. *Language development: An introduction*. 7th ed. Pearson; 2007.
- Perez-Arce P. The influence of culture on cognition. *Archives of Clinical Neuropsychology*. 1999; 14(7):581–92. [PubMed: 14590572]
- Renfrew, C. *The Action Picture Test*. Oxford: 1988.
- Schieffelin, BB.; Ochs, E. *Language socialisation across cultures*. Cambridge University Press; Cambridge: 1986.
- Serpell R. Intelligence tests in the Third World. *Race Today*. 1972; 4:46–49.
- Smith GT, McCarthy DM. Methodological considerations in the refinement of clinical instruments. *Psychological Assessment*. 1995; 7:300–308.
- Sperber AD, Devellis RF, Boehlecke B. Cross-cultural translation: Methodology and validation. *Journal of Cross-Cultural Psychology*. 1994; 25:505–524.
- Streiner, DL.; Norman, GR. *Health measurement scales: A practical guide to their development and use*. 2nd ed. Oxford University Press; Oxford: 1995.
- van Hout A, Evrard P, Lyon G. On the positive semiology of acquired aphasia in children. *Developmental Medicine and Child Neurology*. 1985; 27:231–241. [PubMed: 2581836]
- van Hensbroek MB, Palmer A, Jaffar S, Schneider G, Kwiatkowski D. Residual neurologic sequelae after childhood cerebral malaria. *Journal of Pediatrics*. 1997; 131:125–129. [PubMed: 9255203]

Figure 1.
Theoretical framework for the speech and language battery.

Table 1

Participant Ages and Exposure Groups

	Cerebral malaria		Nonsevere malaria		unexposed	
		%		%		%
Male	77	51	72	46	93	52
Age 6 years	43	28	45	29	42	23
7 years	43	28	45	29	37	21
8 years	34	22	28	18	48	27
9 years	32	21	38	24	52	29
Total	152		156		179	

Table 2**Items and Constructs in the Speech and Language Assessment Battery**

Domain	Construct	Assessment
Content	Topics (Vocabulary) Higher level aspects	Lexical semantics Higher level language
Form	Phonology Syntax Morphology	Phonological Syntax
Use	Social functions Contexts	Parental questionnaire Observation Analysis of functional errors

Table 3

Distribution of Raw Assessment Scores by Exposure Group (Mean, Standard Deviation (Sd) and Range For Normally-Distributed Data and Median, Interquartile Range (IQR) and Range for Skewed Data)

Assessment (Max possible score)	Cerebral Malaria		Malaria + complicated Seizures		Unexposed	
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Receptive language (34)	22.4 (5.2)	11-34	23 (5.3)	7-34	23.8 (5.2)	13-34
Syntax (40)	25.3 (7.5)	0-39	26.5 (6.3)	2-39	26.8 (7.4)	9-40
Higher level language (39)	16.3 (7.8)	0-35	18.6 (7.7)	0-35	19.6 (8.6)	0-38
Lexical semantics (1)	0.57 (0.06)	0.4-0.69	0.59 (0.06)	0.46-0.73	0.59 (0.06)	0.42-0.76
	Median (IQR)	Range	Median (IQR)	Range	Median (IQR)	Range
Pragmatics (100)	86 (79-91.5)	29-99	86 (80-91.5)	36-99	87 (82-91)	50-98
Phonology (100)	100 (100-100)	83.3-100	100 (100-100)	75.8-100	100 (100-100)	83.9-100
Word finding (40)	36 (32-39)	7-40	37 (33-39)	8-40	37 (34-39)	15-40

Table 4

Measures of Test-Retest and Inter-Rater Reliability of the Speech and Language Assessments

Assessment	Re-test reliability			Inter-rater reliability			Maximum assessment scores
	Mean difference (SD)	95% limits of agreement		Mean difference (SD)	95% limits of agreement		
Receptive grammar	1.2 (1.77)	-2.27	4.67	-0.2 (0.79)	-1.75	1.35	34
Receptive vocabulary	-0.62 (1.89)	-4.32	3.08	0.3 (0.55)	-1.44	1.25	25
Syntax	0.7 (1.56)	-2.36	3.76	0.4 (0.52)	-0.62	1.42	40
Higher level language	0.05 (1.7)	-3.28	3.38	0.0 (1.05)	-2.06	2.06	39
Lexical semantics (function words)	0.008 (0.04)	-0.09	0.07	0.004 (0.01)	-0.06	0.05	n/a
Lexical semantics (content words)	0.01 (0.03)	-0.07	0.05	0.01 (0.03)	-0.08	0.04	n/a
Pragmatics	1.0 (2.45)	-3.8	5.8	-0.5 (2.22)	-4.85	3.85	100
Phonology	0.15 (1.5)	-3.09	2.79	0.0 (0)	0	0	100
Word finding	-1.15 (1.04)	-0.89	3.19	0.0 (0.47)	-0.92	0.92	40

Table 5

Mean Scores Showing the Age Differentiation of Normally Distributed Tests

	Receptive Grammar	Syntax	Higher Level language	Lexical semantics (function)	Lexical semantics content
Age	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
6	20.0 (3.9)	22.2 (6.0)	14.9 (7.3)	0.23 (0.04)	0.57 (0.06)
7	21.7 (5.0)	23.0 (7.2)	17.0 (6.0)	0.23 (0.04)	0.57 (0.06)
8	25.2 (4.7)	29.3 (6.4)	20.7 (7.0)	0.22 (0.04)	0.58 (0.06)
9	25.9 (4.8)	30.7 (6.1)	23.9 (7.7)	0.23 (0.04)	0.58 (0.06)

Table 6

Median Scores Showing the Age Differentiation of Tests with Skewed Distributions

Age	Word finding Median (IQR)	Pragmatics Median (IQR)	Phonology Median (IQR)
6	36 (31-37)	85 (81-92)	100 (100-100)
7	37 (33-39)	86 (82-89)	100 (100-100)
8	38 (36-39)	87 (81-92)	100 (100-100)
9	39 (36-40)	88 (82-91)	100 (100-100)

Table 7**Principal Components of the Speech and Language Assessment Battery**

Assessment	Principal Components			
	1	2	3	4
Receptive grammar	0.45	-0.08	-0.14	-0.02
Receptive vocabulary	0.43	0.18	-0.01	-0.05
Syntax	0.42	-0.15	-0.28	-0.07
Higher Language	0.38	-0.26	-0.16	-0.24
Lexical semantics	0.18	<i>0.79</i>	0.21	0.29
Word Finding	0.43	0.21	-0.03	0.03
Pragmatics	0.21	-0.39	0.35	0.79
Phonology	0.17	-0.11	0.85	-0.37

Note. Items in bold are