INT. J. LANGUAGE & COMMUNICATION DISORDERS, 1999, VOL. 34, NO. 2, 151–191

Clinical Forum

Assessing the comprehension and production of language in young children: an account of the Reynell Developmental Language Scales III

Susan Edwards, Michael Garman, Arthur Hughes, Carolyn Letts and Indra Sinka

The University of Reading, Reading, UK

Keywords: testing, child language, RDLS III.

Introduction

Background

The Reynell Developmental Language Scales were first made available in an experimental version in 1969, with revisions in 1977 and 1985 (Reynell 1977, Reynell and Huntley 1985). Throughout this period and beyond, the Scales were amongst the most popular child language assessment tools to be used in the UK. There are a number of reasons for this: they were considered to be clinically practical in that even very young children readily co-operated with the test procedures, the Scales measured both comprehension and production of language, and they were standardized on a large sample of British children. In terms of standardized testing, there were no serious UK contenders. However, in recent years, it became clear that some revisions were needed, especially for the Expressive Scale, which had come to be used far less often than the Comprehension Scale. The Expressive Scale was difficult to administer objectively and yielded uninformative results. In a recent volume devoted to assessment procedures used by speech and language therapists (Beech *et al.* 1993), although the Comprehension Scale is referred to (Lees 1993), there is no mention of the Expressive Scale. In addition, some items and materials used in both scales had become dated and were no longer considered appropriate. A revision was clearly required. This paper describes the new scales, the Reynell Developmental Language Scales III: The University of Reading Edition (RDLS III: Edwards *et al.* 1997). The new Scales were constructed to reflect current knowledge of language impairment and to incorporate findings from up-to-date research on structural and lexical development in child language. As a result, the existing Expressive Scale has been completely revised and substantial changes have been made to the Comprehension Scale.

International Journal of Language & Communication Disorders
ISSN 1368-2822 print/ISSN 1460-6984 online
http://www.tandf.co.uk/JNLS/lcd.htm
http://www.taylorandfrancis.com/JNLS/lcd.htm

RDLS III is based on the original test in that some items and materials from the first and second revised editions were retained where they seemed appropriate. The changes made have been motivated by knowledge of relevant research and tempered by clinical experience. This paper aims to be explicit about the design and purpose of the Scales, as a lack of such an explanation may lead to the authors 'being pilloried in error' (Wirz 1993). Clearly, no test can do everything: this test has been designed to test language abilities of children between the ages of 1;06 and 7;00 years, concentrating on vocabulary and grammar, the 'nuts and bolts' of connected speech. The Scales follow a developmental perspective and yet at the same time afford the tester some insights into which aspects of language might be problematic for an individual child. For example, some of the sections within the Expressive Scale focus on the production of phrase construction, grammatical inflections, and clausal constituents. Although these domain-specific features have been included because of their potential diagnostic value, it is important not to lose sight of the integrated nature of language. Thus, while it is essential to include opportunities to test such features, it is also equally important to include checks on vocabulary and language use. Accordingly, although vocabulary *per se* is not tested (there are other standardized tests available, e.g. German 1986) the test vocabulary has been checked for age appropriateness (see below: Developmental rationale). Similarly, although this test does not claim to test pragmatic skills (again there are alternative ways of assessment, e.g. Dewart and Summers 1995), testing procedures are sensitive to pragmatic demands. For example, in the Expressive Scale, there is guidance for accepting pronominal substitution and the child's awareness of *new* versus *given information* and the Comprehension Scale includes a section on *inferencing* skills. The Scales do not test a child's phonological skills, although again, because phonological problems may affect vocabulary and grammar, clear guidance is given in the test manual regarding the acceptance of particular phonological forms. Finally, the Scales do not measure how effectively the child actually communicates within his or her environment but they will give the clinician some of the necessary information needed to make that assessment.

The Scales, then, have been designed to test *language* and therefore must minimally include tasks that probe grammatical and lexical skills. Neither of these domains is optional in language testing. While it is certainly not claimed that every child referred to a speech and language therapist will require the same assessment procedure, where language delay is suspected then it is absolutely essential that language is examined. To this end, the Scales for testing both comprehension and production of language are provided in line with professional standards (van der Gaag 1996). These Scales provide one way of making that examination. The results provide the tester with a description of the child's language skills compared with his or her peers, a comparison required by clinicians (Lees 1993). They also provide information which can be regarded as prescriptive, in the way in which the term is used by Wirz (1993), in that areas of weakness are highlighted for therapy or further testing. The Scales may be used as part of an extensive investigation of a child's abilities, where the clinician utilizes a number of tests; the manual provides additional guidance in selecting further appropriate procedures. Observation of the child, plus parental and/or teacher report will further contribute to the assessment process. It is also recommended that the clinician samples the child's connected speech in more naturalistic contexts; there are several published procedures which give guidance on segmentation and analysis (e.g. Garman and Edwards 1995).

Finally, when repeated at a later stage, the Scales can also be used to check progress and to measure the efficacy of therapy.

A brief review of some of the theoretical underpinnings of the test is given below (Developmental rationale). A description of the test (further details can be found in the manual that accompanies the Scales) is followed by a description of the methodology used in test construction and the outcome of data analyses. To begin, however, the new Scales are considered within the context of current tests. As will be demonstrated below, there are few tests available which give the same kind of information as the new RDLS III.

Comparison with other assessment procedures

In revising and re-standardizing this test, the authors wanted to produce a clinical tool that would be at least as useful as the previous version and which would reflect advances in knowledge in the field of child language acquisition, as well as changes in clinical practice. Advances in the former have resulted in the development of a range of sophisticated assessment tools, many of them capable of exploring quite specific areas of communicative difficulty. At the same time, changes in working conditions for clinicians have resulted in the need to monitor case loads systematically and to target remediation in order to achieve maximum efficacy. An initial question, then, was whether a test based on the original Reynell Developmental Language Scales would still be of use to clinicians.

The original RDLS

The original questions that the Scales set out to answer for any individual child, as outlined in the 1985 manual (Huntley 1985), were as follows:

(1) How far is language development behind what would be expected of children of their mental age? How far is this outside the range of normal variation?
(2) How can the child's progress be assessed in relation to his/her mental development?
(3) Is the difficulty a delay in receptive or expressive language, or both?

The answers to these questions would at the time have given the clinician a good idea as to whether any individual child had a language impairment (either delay or disorder), and some idea as to the nature of this impairment, for example whether there were problems with verbal comprehension as well as expression. In addition, the nature of some of the subtests gave insight into more specific problems that the child might have. For example, if the child failed on the section in the comprehension test involving coloured pencils of different length, this would suggest that s/he might have difficulty with the linguistic realization of concepts of colour and length. Alternatively there might be an auditory memory problem, the child having difficulty processing all the elements in these commands.

The original Scales served a purpose in screening for impairment, and also had a limited diagnostic power. The authors felt there was still a place for this role. The new Scales therefore aim to fulfil the following desiderata:

(1) to reflect advances in knowledge about child language development that have taken place since the original Scales were developed;

(2) to provide a screening tool, based on a robust standardization, which will enable clinicians to identify language impaired children; and
(3) to have a diagnostic value, by indicating to the clinician areas where the child has difficulty, and which further types of evaluation may be of use.

In addition, there were a number of practical aspects. The authors wished to retain the features of the original Scales that made them attractive to small children; also update materials so that they would have wider social and cultural acceptability in accordance with social changes that have taken place since the formulation of the original Scales. It was felt desirable to retain the original age range to which the Scales applied, since this reflects the time of greatest development of language, and continues to reflect the common age range for referrals to speech and language therapy in the UK.

Other relevant assessment procedures

There now follows a short review of some tests currently used to assess children in the UK, looked at in the light of the above criteria. It will be seen that there are few that fulfil the screening and diagnostic roles outlined above. It should be noted that this list is not intended to be comprehensive, but to include a selection of tests that appear to be used with some frequency by UK clinicians. Details of the tests discussed can be found in table 1.

The *Renfrew* battery of tests represents a practical set of assessments for looking at various aspects of language development in the age range 3–8 years, and are quick to carry out. However, despite recent updating of some of these tests, they were originally formulated at about the same time as the experimental version of the RDLS scales. They do not therefore reflect recent research. The *Action Picture Test* (fourth edition: Renfrew 1997a) is a quick screen for connected speech and can indicate delays in this area, but provides no more specific diagnostic indications. It could be argued that the grammatical part can be used to check the presence of specific structures, but it is not comprehensive and results would not constitute a valid basis on which to plan grammatical remediation. The *Bus Story* (fourth edition: Renfrew 1997b) has been found to provide a reliable indication of disorder, and in particular to be predictive of long-term disorder (Bishop and Edmondson 1987). Despite only minimal updating, the story and pictures appear to be universally attractive to children, regardless of time or cultural background. The assessment does not provide information on comprehension, however. It can provide a useful way of eliciting an expressive language sample, and of giving an indication of the child's narrative ability, but it remains less comprehensive than the RDLS III Scales.

The *Clinical Evaluation of Language Fundamentals—Revised* (CELF-R: Semel *et al.* 1987) and the CELF Pre-school Scales (Wiig *et al.* 1992) have become increasingly popular in the UK over recent years. CELF-R covers a wider upper age range than RDLS III, extending up to 16 years. In addition, these assessments are divided into comprehension and expressive subparts, each containing a number of subtests, and it is possible to calculate separate standardized scores for comprehension and expression, and for each subtest. In many ways, this assessment procedure performs some of the same functions as RDLS III, and its increasing popularity may well be a function of dissatisfaction with the earlier version of the Reynell Scales, as it became more dated.

Table 1. Characteristics of some child language tests commonly in use in the UK

Name of test	Date of publication†	Aspect of language assessed	Age range (years)	Standardization: Place, date‡	Standardization: No. in sample
BPVS	1997	Receptive vocabulary	3–15	UK	2571
CELF-R	USA, 1987 UK, 1997	Language comprehension and expression	5–16	USA	2426
CELF: Pre-school	1992	Language comprehension and expression	3;0–6;11	USA	1500
DLS	1982	Language comprehension, limited assessment of expression	N/A (for use in school setting)	–	–
McArthur CDIs	1993	Vocabulary comprehension and expression, use of gestures and phrases	(1) 0;8–1;4; (2) 1;4–2;6	USA	(1) 325; (2) 561
Pragmatics Profile	1995	Pragmatic skills	Birth–10	–	–
Renfrew Action Pictures	1997	Expressive language: grammar, information	3;6–8;0	(1988) UK	594
Renfrew Bus Story	1997	Expression, narrative	3–8	(1994) UK	573
STASS	1988	Expressive language: structure	3–5	UK: profile at 3, 4 +5 years	204
Token Test for Children	1978	Comprehension	3;0–12;6	USA	1306
TOWF	1986	Expressive vocabulary	Grades 1–6, 4–12	USA	1200
TROG	1989	Comprehension of grammar		(1983) UK	>2000
RDLS III	1997	Language comprehension and expression	1;3–7;6	UK	1076

†Latest edition.
‡Date is given where this is different from publication date.

There are two points to be made here. First, although a British version of CELF-R has recently been published, it has not been standardized on British children. Second, the theoretical motivation behind the organization of the CELF sections is rather different from that of RDLS III. The overall structure of the test is designed to reflect the form and content aspects of the Bloom and Lahey (1978) model of language, and the reader is referred to Wiig and Semel (1984) for a more complete account of the areas covered by the various subtests. Each area explored by the subtests stands completely alone, and norms are given for each throughout the age range (although some subtests apply to only a restricted age range). This is in contrast to RDLS III where there is a developmental progression through each section, and although each section has a different focus, successful completion of any one section requires the integration of a number of different earlier developing linguistic skills. One practical consequence of this is that the younger child does not have to go through the whole test once s/he has reached a ceiling; also, at the discretion of the clinician, the earliest sections of the test may be omitted when testing an older child. While clinicians may vary according to which approach they prefer, it is possible to get the impression that CELF is made up of a collection of disparate sections that do not relate to each other in any clear way. CELF may be of more value in adding further knowledge when specific areas of difficulty have already been located. At present, the main advantage may be that CELF offers an assessment procedure for the child whose age is above the RDLS III range.

There are other procedures that have the function of providing an efficient screening test with, in some cases, diagnostic indicators. The *McArthur Communicative Developmental Inventories* (CDI; Fensen *et al.* 1993) are based on parental report in terms of comprehension and expression of lexical items (Inventory 1) and of multiword utterances (Inventory 2). Again, the Inventories are not standardized on British children. However, they give a useful indication of early language development, especially up to 2;6 years. It may be considered that parental report gives a more reliable indication of this early development, since elicitation procedures are difficult to carry out effectively with very young children. For this reason, the RDLS III Expressive Scale has not been standardized on children younger than 1;9 years (while the Comprehension Scale retains the lower age limit of 1;6 years). The McArthur Scales therefore can be viewed as complementing RDLS III, in terms of being more useful with this very young age group.

A similar comment can be made concerning the *Pragmatics Profile* (Dewart and Summers 1995). Designed to elicit information on the child's pragmatic functioning, by means of parental (or other carer) report, it again can be very useful with the very young age group. The focus is on communication rather than vocabulary or structure. It is also useful with children who are developmentally at a young stage because of learning disability, and of course with any children felt to have a degree of pragmatic impairment. The profile is now extended up to 10 years, which makes it particularly useful with this last group.

Another procedure that is popular in the UK in terms of screening, diagnostic and remediation indicators, is the assessment that accompanies the *Derbyshire Language Scheme* (DLS; Knowles and Masidlover 1987). This provides an informal assessment, through toys and pictures, of a child's comprehension and production. The Rapid Screening Test, carried out first, gives an indication of where to begin with more detailed probing (Detailed Test of Comprehension: DTC). Limited information on expressive abilities can also be obtained while carrying out the DTC. Since the procedure was developed initially to provide the clinician (or teacher) with a level at which to begin a language programme, it is closely linked with remediation. It was originally designed for use in schools, with language-impaired or learning disabled children, and the training of therapists and clinicians on jointly run courses ensures that the scheme works very well in these situations. Disadvantages are that it is not a standardized procedure, and is most useful only for the earlier stages of language acquisition, up to about 3;6 years.

In addition to the assessment procedures outlined above, there is now a wealth of material available for assessing quite specific areas of language impairment, some of them standardized on a UK population. Examples include the *Test for the Reception of Grammar* (TROG; Bishop 1989) which, as the name suggests, focuses on grammatical understanding; the *British Picture Vocabulary Scales* (BPVS; Dunn *et al.* 1982, 1997) which focuses on lexical comprehension; and the *South Tyneside Assessment of Syntactic Structures* (STASS; Armstrong and Ainley 1988), a procedure for eliciting expressive language which can be analysed according to Linguistic Assessment, Remediation and Screening Procedure (LARSP) principles (Crystal *et al.* 1989). Procedures standardized elsewhere include the *Token Test for Children* (DiSimoni 1978), where comprehension of the linguistic realization of specific concepts is systematically investigated, and the *Test of Word Finding* (TOWF; German 1986), which looks at vocabulary production in a variety of elicitation contexts.

While it is invaluable to have such procedures as these for their specific purposes, it is not feasible to carry out large numbers of assessments on all children in a heavy case load. The clinician needs to be able to choose assessments that are going to help in the effective planning of remediation, and not waste time assessing areas that are not a cause for concern, or not a priority at any given moment. The authors have tried with RDLS III to help the clinician in this. Difficulty with particular aspects of the test should give insight into whether further assessment procedures will be helpful, and which procedures these should be. In this way, RDLS III is designed to fit into the modern context of detailed and sophisticated testing procedures.

Developmental rationale

Since the publication of the original Scales, intensive research activity in child language has provided much new information about the detail of the developments in both comprehension and production abilities in the pre-school period and beyond. Brown (1973) was an important landmark, as was the large-scale study described by Wells (1985). Now available are the extensive language production database in the Child Language Data Exchange System (CHILDES; MacWhinney 1995), and lexical development norms provided by Raban (1988) and by Fensen *et al.* (1994) and Dale and Fensen (1996). Recent reviews of comprehension development are provided by Milosky (1992) and Golinkoff and Hirsh-Pasek (1995). Much of this information has also been applied in the investigation of impaired language development: this research base is also reflected in a number of specific, clinically oriented receptive and expressive language assessment procedures (Miller 1981, Gutfreund *et al.* 1989, Miller and Leadholm 1992, in addition to others mentioned above).

Clearly, research findings from a range of studies employing different methodologies cannot simply be incorporated into any particular assessment procedure. Nevertheless, they provided detailed information on what sorts of items should be trialled. Beyond this, they afforded the opportunity to make two major changes to the original Scales: first, to bring the Expressive Scale more into line with the Comprehension Scale, and second, to provide for more detailed assessment of how language (as opposed to concept formation) develops over the age range covered by the Scales. The intention was for this research base to inform the new Scales while maintaining and enhancing their original breadth of coverage.

In realigning the Expressive and Comprehension Scales, the authors took account of the evidence, at different stages of language development, that comprehension precedes production (Bates *et al.* 1995); in confirmation of this, trialling of items in the revision suggested that, while the Comprehension Scale should start at 1;6, the Expressive Scale should start at 1;9. This is not to say that expressive language is not evident in a child below this age; it is, however, much more difficult to test systematically in the way that is possible with language comprehension, and different assessment procedures may be more effective (see Other relevant assessment procedures, above). The fact that comprehension and production each have, none the less, their own developmental logic is reflected in the differential sectioning of the two new scales. It was also noted that methodological advances in the field have confirmed, first, the impractical nature of attempting to elicit specific expressive targets in situations where the child has considerable freedom over what s/he chooses to say and, second, the impossibility of evaluating such data on the spot without listening to a recording at a later stage. This meant that the elicitation procedures had to be correspondingly more constrained, to permit the scoring of attempts against known targets, and a range of techniques had to be used, to avoid under-representing the child's abilities (see Description of the scales for details). It is, however, strongly recommended that wherever an expressive language difficulty is encountered on the new Expressive Scale, the assessment should be supplemented by a more spontaneous procedure (see above).

Incorporating the main trends of language development has required a broad-brush approach, in the spirit of Berman (1986), summarizing much child language work in English and other languages. While Berman's stages have not been formally recognized as sections in our revision, they serve as a useful framework here, to outline the developmental rationale that any modern assessment procedure in this area must take account of. The revised Comprehension and Expressive Scales will be considered together in what follows: see below for further detail on the particular sections and items in the scales. Table 2 illustrates how Berman's stages might correspond to relevant sections of RDLS III.

The early stage

This is characterized by comprehension and production abilities that are lexically restricted, pragmatically controlled, context-dependent and largely language-universal. The child knows many times more words than s/he produces, and comprehends simple multiword utterances in advance of their production.

Much of the child's awareness of what language forms mean at the very earliest stage is highly situation-specific, and not easily assessed outside the child's home environment (although this may be possible via a parental report protocol such as the CDI; Fensen *et al.* 1993). It is therefore appropriate to begin clinical assessment

Table 2. Comparison of Berman's stages with RDLS III sections

Stage†	RDLS Comprehension	RDLS Expressive
1: Context anchored; pragmatic; shared across languages	Sections Ai, Aii, B and C	Sections A and B
2: Structure-dependent; grammaticized; attentive to major structural properties of the native tongue	Sections D, E, F and G	Sections Bi, Bii, Ci, Cii, Ciii and D
3: Discourse sensitive; grammar and usage; attentive to specific norms and registers of the native tongue	Sections I and J	Sections Ei, Eii, Eiii and F

† Stages taken from Berman (1986).

as the child emerges from this earliest stage of comprehension. The emphasis in the early sections of the Comprehension Scale is on common nouns, since these are known to be dominant in early comprehension vocabularies (Benedict 1979, Gentner 1982). In addition, the authors were able to draw on frequency norms for both comprehension and expression, as recorded in the Dale and Fensen (1996) and Fensen *et al.* (1994) databases. The distinction made in the original Scales between real and miniature items is not maintained, however, as these recent norms suggest that size distinctions of this kind do not influence order of acquisition. The number of items in this part of the new Comprehension Scale (sections Ai and Aii) has been reduced to allow more scope for later sections. Vocabulary frequency was checked throughout all sections on the Scales, using both Dale and Fensen (1996) and Raban (1988).

The child's early comprehension of multiword utterances may be heavily dependent on lexical understanding, as assessed in the first two sections, supported by a strategy of putting the objects named in their normal functional relationships (e.g. the spoon in the cup: Chapman 1988). However, at this stage children can also use language to relate objects in specific ways (reflected in Brown's (1973) Prevalent Semantic Relations). Items have therefore been included that tap this transition from lexical to grammatical comprehension, as the child begins to process prepositions and simple verb-argument structures (clause constituents). The latter involve verbs with one and two arguments, with agent as subject, theme as object, and an intransitive construction with a locative expression in post-verbal position, which is the simplest type of adverbial, in canonical position (Comprehension Scale Sections B, C). In contrast to the original Scales, items where the child can relate objects in the 'normal' way, without necessarily employing grammatical understanding (as in the 'spoon in cup' example given above), have been avoided here.

As regards production, the single-word stage culminates in a vocabulary of around 50 words, around the middle of the second year. From this point on, the child's developing ability to produce verbs (after nouns have appeared: Benedict 1979) is the basis for later clausal developments. The vocabulary spurt that many children show at this point is largely a matter of increasing noun vocabulary. The pattern of early two-to-three-word utterances largely built on just a few verbs lasts until the syntactic spurt around 2;0, when functional categories, evidenced by the appearance of inflections and grammatical 'function' words, begin to emerge (Radford 1995). In the Expressive Scale, section A (nouns) taps into the middle of this period, and section B (verbs) towards the end, when more verbs are starting to appear.

The middle stage

This shows the start of language-specific control of morphology, and clausal and phrasal constructions, as well as the development of relational vocabulary, again with comprehension preceding production.

As they begin to develop control of clause constituents, children are unable initially to use clause constituent order alone to interpret *reversible* structures (see section H below). They should, however, be able to identify actions that are not derivable from the situation alone (so, for example in Comprehension Scale section D, Teddy has to be made to sit on the *box*, not the *bed*).

As well as integrating lexical with clausal information, children also show development of phrase structure, in terms of attributes, negation and further developing prepositions (Johnston and Slobin 1979, Miller and Paul 1995: Comprehension Scale sections E, F, G). This process, together with growing grammatical awareness, enables the child to map noun phrases in constructions on to appropriate grammatical roles, in situations where pragmatic strategies will not suffice (Slobin and Bever 1982, Chapman 1988). Examples are where either both arguments are equally plausible as agents or theme, or a pragmatic strategy will lead the child astray. This ability is explored in Comprehension Scale section H, with items such as 'the baby is pushing the mummy', where merely understanding the lexical items and relying on experience of what usually happens will not suffice.

With more verbs starting to appear in the child's multiword utterances and the start of the syntactic spurt (Radford 1995), phrase structure is the level of grammatical structuring that will eventually yield constituents of clauses (expanded subjects, objects, etc.), as well as independent phrasal utterances. Phrasal utterances are explored in Expressive Scale section Bi, Bii, and their incorporation into clausal utterances in Expressive Scale section D (see below).

Brown (1973) identified his Stage II of grammatical development arbitrarily as the period during which the child's mean utterance length increased from 2.25 to 2.75 morphemes, but more significantly as the point at which the most important changes in the appearance and growth of inflections takes place. The importance of this development is emphasized by Radford (1995), for whom it represents the shift from lexical-thematic to functional-relational syntax. The Expressive Scale sections Ci to Ciii monitor the emergence of the plural, past tense and agreement inflections; these are diagnostic of normal development, as well as of specific language impairment (SLI) (Leonard *et al.* 1987, Rice *et al.* 1995, Schelletter *et al.* 1998). Beyond this, the Expressive Scale section D taps the sort of information that is adumbrated in Crystal *et al.* (1989) and more extensively researched in Wells (1985): the increasing length and complexity of patterns of clause and phrase structure. The child becomes aware of the distinction between optional and obligatory verb arguments, something again that is relevant for the detection of SLI (Johnston and Kamhi 1984). In section D, all obligatory verb arguments must be expressed.

The late stage

This stage is marked by an increase in syntactic complexity, with further vocabulary development, as well as connected discourse and metalinguistic skills.

Comprehension Scale section I checks two areas that have been researched extensively in recent work (Maratsos *et al.* 1985, Crain 1991). First, passives are represented here, by the short or agentless form, which is more typical of spoken English than is the full form. Morphological, syntactic and lexical processing are involved, in items where the mapping of subject to theme has to be achieved for appropriate comprehension (e.g. 'the bull was chased'). Second, relative clauses represent the most complex form of noun phrases, and pose problems in processing in the earlier stages for language-normal children. Comprehension Scale section J looks at the processes that integrate linguistic meaning into knowledge of the world. This is not to claim that these processes do not occur at earlier stages of language development: they do, but at earlier stages they are more limited to the updating of information on the 'here and now', whereas in this section we are probing how far children can process linguistic forms in relation to their wider knowledge.

The Expressive Scale section E elicits more advanced structures involving complex clause formation of differing types by means of imitation, correction of errors via metalinguistic skills, and further advanced structures by means of utterance completion. The value of imitation as an index of developing grammatical ability was suggested early on by Rodd and Braine (1970), and taken up by Slobin and Welsh (1973), and the value of this procedure has been demonstrated over a number of years in clinical assessments that have been carried out at Reading. The use of imitation has also been established with autistic and Down syndrome children (Tager-Flusberg and Calkins 1990). Children's ability to provide explicit statements about aspects of language forms are well documented by Karmiloff-Smith (1993), and error correction has been used to identify the child's degree of metalinguistic awareness of verbs in the age range 6–8 years (Sutter and Johnson 1990).

Finally, auxiliary verbs in English are more than just phrasal elaborations of the verb group: they are central to the successful formation of interrogatives, negatives and tag questions. Their development and deployment in different types of sentences is a long process which begins before the third birthday and occupies the rest of the pre-school period in normally developing children (Fletcher 1985, Richards 1993). The auxiliary verb system has been identified as an area of difficulty for language-impaired children (Fletcher and Peters 1984). Although the structures exemplified in Expressive Scale section F appear in spontaneous speech at earlier ages, it was found that under the conditions of formal testing, this was the most demanding section, and hence is appropriately ordered last.

Description of the scales

Introduction

As well as being chosen to reflect social and cultural changes that have taken place in the British Isles, the pictures and toys used in the test were carefully trialled to ensure that they were attractive and familiar to children; materials that proved difficult for children to recognize at the trialling stage were removed from the final version of the test.

It is intended that the two new Scales be administered together, thus providing the tester with as complete a picture as possible of the child's verbal comprehension and expressive language development. The different sections in the Scales have been designed to represent development in terms of stages of language development (see above), with each section having an explicit linguistic focus. They also reflect the integrated nature of the language system: as the test progresses, increasing demands are made on the child's ability to utilize a range of linguistic and metalinguistic skills in order to understand the tasks set and give appropriate responses.

The Comprehension Scale

The Comprehension Scale has 62 items organized into 10 sections, A to J, and both pictures and toys are used. The sections cover a range of skills, from comprehension of single words through to more complex grammatical and inferencing skills. In response to the tasks set by the tester, the child is required either to manipulate a toy or point to a picture or a toy.

Single words

In section A, a combination of life-sized and miniature objects is used. These were originally ranked according to frequency norms (Dale and Fensen 1996), with adjustments made subsequent to trialling (see Trialling, below). The child is required to point to each item as it is named by the tester.

Building clausal constituents

Sections B to G make use of both toys and the Picture Book, with the child having to choose and manipulate the appropriate toys or select the correct picture. These sections correspond to the early and beginning middle stages of language acquisition described above. Within them, the child must show an understanding of two named objects (e.g. 'give me the apple and the teddy'), be able to select the correct agent to carry out an action (e.g. 'make teddy sit'), comprehend a variety of verbs (e.g. 'make rabbit push the bed'), understand nouns and their accompanying adjectives (e.g. 'show me the blue box') and respond to noun and prepositional phrases of increasing complexity (e.g. 'take two buttons out of the cup' and 'put teddy in front of the truck').

Complex grammar

Pictures are used in sections H and I to tap into more complex grammatical skills and correspond to the later middle and late stages given above; for each item, the child is asked to select one picture from an array of four. Section H explores noun phrases and their associated grammatical functions. In order to respond appropriately, the child must demonstrate an ability to apply rules for constituent order. A lack of knowledge in this area could result, for example, in the child selecting the wrong picture when presented with the instruction: 'show me *the boy's carrying an elephant*'. In such a case, the child may well choose the picture depicting *an elephant carrying a boy* instead. Thus 'world knowledge' would be applied inappropriately, with the child incorrectly assigning the agent role to the *elephant*. Section I incorporates the use of passives, relative clauses and more complex lexical items (e.g. 'all the girls except one are eating').

Inferencing

The Comprehension Scale is concluded with section J, which makes use of a single picture depicting a 'burger bar'. This particular picture was chosen because it crosses cultural boundaries and is a recognizable situation for most children. During this part of the test, the child is asked to point to one of the people in the picture, in response to questions asked. The focus of this section is on inferencing skills. For the child to respond appropriately, some degree of world knowledge and experience is needed. In order to answer the last question ('whose daughter is having a birthday party?'), for example, the child must, in addition to understanding the pronominal reference used ('whose'), successfully locate the 'party group' in the picture and then identify the parent sitting at the table.

The Expressive Scale

The Expressive Scale contains 62 items across six main sections, A to F. The distinction between single words and syntactic sequences incorporated in the Comprehension Scale is maintained in the Expressive Scale and nouns are ordered before verbs to reflect the observed sequence in normal development. Toys, pictures and finger puppets are employed to elicit responses using a range of procedures.

Single words

As in the Comprehension Scale, section A of the Expressive Scale deploys a combination of familiar life-sized and miniature objects, all ranked originally according to frequency norms (Dale and Fensen 1996), and then adjusted subsequent to trialling. In this section the child is required to name the objects, using recognizable, though not necessarily adult, forms of the target words (guidance is given in the manual regarding acceptable phonological forms).

Clausal, phrasal and morphological development

Both toys and pictures are used in sections B to D and the child is asked to respond appropriately either to actions carried out by the toys or to pictures s/he is shown in the Picture Book. These sections focus on different aspects of expressive language development, corresponding to the very end of the early stage, and the whole of the middle stage as outlined above, and include early verb vocabulary, phrasal and clausal structures, and inflections. Although these aspects are presented separately within the Scale for testing purposes, clearly the processes involved overlap, are interrelated and develop simultaneously to a certain degree.

In section B, the tester is given some indication of the child's use of verbs, including one item demanding a verb-object response ('Teddy's wash/ing (the) truck'), and of his/her ability to produce an expanded phrase (e.g. the prepositional phrase 'on the plate'). Section C focuses on inflections (see The middle stage, above, regarding the significance of inflection in both normal and impaired language acquisition). Here the child is asked to produce plural forms, third person inflections and past tense marking. The increasing syntactic complexity of utterances found as the child continues to develop is addressed in section D. Some items within this section are modelled for the child and others require spontaneous responses. The structures elicited contain three and four clausal elements and all the clause elements must be present for the child's response to be considered correct (e.g. 'Teddy's putting the knife under the bed'). Nevertheless a considerable range of responses is considered acceptable here in terms of how clause elements are realized, in that the child can use pronoun forms to signal given versus new information. Some verbs are also included that permit alternating argument structures (e.g. 'rabbit's giving teddy a red brick', 'rabbit's giving a red brick to teddy'), either of which is acceptable. There is considerable current research interest in the use of verbs permitting alternating structures (see for example King *et al.* 1995). The tester is referred to the manual for guidelines on acceptable responses here.

Complex structures

The sections from E onwards correspond to the later stage of language development (The late stage, above). The structures introduced in section E are more complex and are elicited by different procedures, reflected in a three-way division into parts. No equipment is required for part (i). In part (ii), however, finger puppets are introduced for the first time (worn by the tester). In part (iii) the child is asked to retell a story based on one or two pictures from the Picture Book.

The first procedure (Ei) focuses on *imitation*. Here the tester models multiclausal sentences and the child has to repeat these accurately. In Eii, the procedure employed is *error correction* in which the child must listen to an 'incorrect' sentence (e.g. 'the boy eating apples') and produce a corrected version (e.g. 'the boy's eating apples'). As discussed above, with this procedure, the child demonstrates an awareness of grammatical structure and must use metalinguistic skills to manipulate the structure appropriately. Finally, Eiii uses an *utterance completion* task, in which the child is invited to retell a story using a main and subordinate clause (the prompt given by the tester restricts the child to use language in this way).

Auxiliaries

In section F of the Expressive Scale, finger puppets are used by both the tester and the child. This final section concentrates on the child's use of auxiliary verbs in negation and question formation. The focus is an important one: the auxiliary system is seen as developing throughout the pre-school period, and can present a particular problem for language-impaired children. The tasks set in this section are very demanding and tap into the child's ability to manipulate language, to role-play, and to switch tasks.

Development and standardization

Preliminary stages

The development of the new Scales began with a study of the existing Scales, based on critical reviews which had been commissioned by the publisher, the views of local clinicians, the team's own experience and on what is now known about the way in which children's language develops. The decision was taken to replace the existing Expressive Scale and to modify considerably the Comprehension Scale. Once areas of language and methods of elicitation were chosen, items were written, pre-trialled, and modified where necessary. These items and a number of items retained from the existing Scales were then used to construct two versions of the test, with a number of shared 'anchor' items in each version to allow comparisons of difficulty to be made. The two versions of the test were modified in the light of comments of clinicians to whom they were shown.

Trialling

Approximately 400 children aged 1;06 to 7;00 took part in the trialling, taking each version in approximately equal numbers. None of the children had been referred to speech and language therapy. All were judged to be developing normally and to have hearing within normal limits, and all had English as their first language (these same criteria applied to children who participated in the standardization). On the basis of item response theory (Rasch) analysis, traditional item analysis (to ensure that items discriminated well between children and that there was a good range of difficulty throughout the test), and taking into account the experience of the field workers in administering the two versions of the test, a final version of the test was put together, with 62 items in the Comprehension Scale and 62 items in the Expressive Scale.

Standardization

Data collection

The test was administered to a further 1074 children (519 boys 555 girls) aged between 1;3 and 7;6 in England, Scotland, Wales, Northern Ireland and the Republic of Ireland. For concurrent validation purposes, a subset of these children (84 boys and 107 girls) in Reading and Edinburgh were given the TROG (Bishop 1989) and the 1997 edition of the BPVS (Dunn *et al.* 1997). Although not published at the time the authors were able to use the standardization version of the 1997 BPVS.

Analysis

Statistical analysis was carried out using SAS. Presented below are summary statistics, reliability coefficients, the results of an investigation into factors other than age, the results of the concurrent validation study, and an explanation of how age equivalences and percentiles were calculated.

Summary statistics

Summary statistics are given in table 3. The Pearson product–moment correlation between scores on the two scales ($n = 1039$)[1] is 0.85, in line with what one would expect when two different but closely related abilities are measured.

Reliability

Kuder–Richardson Formula 20 reliability coefficients are shown in table 4. For subsets of children by age, reliability coefficients (K–R) and standard errors of measurement (SEM) are shown in table 5.

Table 3. Results of basic analysis on total sample

	Number	Mean	Standard deviation
Comprehension	1074	48.32	13.45
Expressive	1039	36.66	13.10

Table 4. Reliability coefficients based on total sample

Comprehension	0.97	Expressive	0.96

[1] $n = 1039$ rather than the total number of children participating, because a small number of children below the age of 1;9 years did not do the Expressive Scale (see the manual, Edwards *et al.* 1997: p. 10) for discussion of the lower age limit for the Expressive Scale.

Table 5. Reliability coefficients, standard errors of measurement, means and standard deviations by age

	Comprehension					Expressive				
Age group (years)	*N*	K–R	SEM	Mean	SD	*N*	K–R	SEM	Mean	SD
Up to 3†	182	0.96	2.75	24.11	13.75	147	0.92	2.03	14.24	7.18
3–4	153	0.87	2.50	45.54	6.92	153	0.86	2.30	28.00	6.16
4–5	237	0.75	2.17	51.55	4.34	236	0.86	2.32	36.22	6.21
5–6	216	0.66	1.95	54.79	3.34	217	0.89	2.45	43.00	7.40
Over 6	286	0.44	1.65	57.64	2.21	286	0.85	2.37	48.35	6.11

†Up to and including third birthday; next age group begins at 3 years and 1 day, and so on.

Factors other than age

General linear modelling revealed that for both the Comprehension and the Expressive Scales there was a statistically significant but very small effect associated with the school where the test was taken and with the person who conducted the test. However, the size of their contribution is so small in relation to the effect of age that they do not affect the validity of the standardization. There was no effect associated with the county in which the testing took place or with whether the test was taken in a rural or urban area.

There are, however, differences between the scores of boys and those of girls on the two scales (see table 6). It is in the earlier years that differences between boys and girls are most noticeable. Up to the age of 3;02, girls score between 2 and 7 points higher on the Comprehension Scale, and between 1 and 5 points higher on the Expressive Scale [*t*-tests confirm the statistical significance between means of boys and girls up to age 3: $p = 0.002$ for comprehension ($n = 188$, $t = -3.17$, d.f. $= 183$) and $p = 0.004$ for expression ($n = 155$, $t = -2.91$, d.f. $= 153$)]. The 95% confidence intervals are 2.4–10.4 and 1.10–5.78 respectively. It is not possible to say whether this represents a real difference between boys and girls in the British Isles (the standardization of the previous version of the Reynell test also revealed slightly higher scores for girls) or is the result of the composition of the particular sample of children used in the standardization (the sample sizes are small where the differences are greatest). Further research in this area might well be profitable, since other recent work in first language assessment has also revealed differences favouring girls. Such differences are reported in the McArthur scales (Fensen *et al.* 1993), and were also noted by Wong *et al.* (1992) using the Cantonese version of the original RDLS Scales on a large sample of children (431 boys, 424 girls) attending nursery schools in Hong Kong. By contrast, Wells' (1985) study, using naturalistic sampling techniques with two groups of children aged 42 months ($n = 60$) and 60

Table 6. Mean scores of boys and girls compared

	Boys		Girls	
	Mean	SD	Mean	SD
Comprehension	47.42	14.57	49.16	12.26
Expressive	35.70	13.48	37.55	12.68

months ($n = 65$), found no significant gender-related differences in development of 'a wide variety of language systems' (p. 345). It should be noted that the number of subjects is comparatively small in this study.

Concurrent validity

The results of the concurrent validity study are shown in tables 7 and 8. Correlation coefficients between the tests were calculated using raw scores for the Reynell Scales and the BPVS II and number of blocks passed for the TROG. The strength of the correlations, as shown in table 8, is as expected between tests that are measuring related but different skills and where the age range, 4;00 to 7;00, is somewhat limited (this age range was selected because the TROG is not standardized below 4;00 years).

Calculating age equivalence and percentiles

In order to smooth out irregularities in the data, the mean scores and standard deviations (SDs) for all ages were first calculated by a weighted, moving 6 month average. The weighting scheme gives a weight of a half (0.5) to children 3 months on either side of the centre, and for intermediate ages an exponential weight was used, with the exponent proportional to the distance from the centre.[2] The intervals used for obtaining the means and SDs were still slightly too noisy to give a coherent table. One problem was that for the youngest and the oldest children one does not have a centred group (e.g. for 1;06 years there are almost none below); another is that with samples varying between 45 and 100, the SD estimates can move up and down sufficiently to make a smooth standardization difficult. For this reason, further smoothing of the standard deviations was necessary. Six-monthly steps, 2;00, 2;06

Table 7. Means and standard deviations obtained in the concurrent validity study

	N	Mean	SD	Min.	Max.
Reynell Comprehension	191	54.67	4.12	37	61
Reynell Expressive	190	42.96	7.65	27	58
BPVS II (raw score)	191	61.69	17.56	26	109
TROG (blocks passed)	190	12.09	3.85	3	19

Table 8. Correlations† between Reynell, BPVS II and TROG

	Reynell Exp.	BPVS II	TROG
Reynell Comprehension	0.62	0.68	0.70
Reynell Expressive		0.75	0.67
BPVS II			0.75

†Correlations calculated on raw scores.

[2] Formally this may be expressed as:

$$\text{weight} = 2^{(\text{centre} - \text{age})/3}$$

where 'centre' is the target age and 'age' the actual age of a child, both in months.

and so on were selected as centres and the SDs were smoothed around these points by using a 6 month moving average (e.g. using 1;09–2;03 to obtain a smoothed SD for 2;00). The SD for intermediate age points was calculated by linear interpolation between the adjacent centres. The percentile points and standard scores were then calculated.

Tables in the manual allow the tester to convert scores to age equivalents, standard scores (with an average of 50), and percentile scores. This is in line with most modern standardized test procedures, and should permit the tester to give a clear indication of performance when reporting on a child's language.

Summary

The Reynell Developmental Language Scales III have been designed to test the language abilities of children from 18 months to 7 years. They have been standardized on a large, representative sample of children in the United Kingdom and thereby provide a robust measure of language achievement. These Scales provide clinicians with the means of comparing the performance of children who have been referred for therapy, or who are in therapy, with the performance of a large group of normally developing children. The Scales reflect the developmental progression of normal child language in the early years, focusing on key features of child language acquisition at different stages as well as on features which are known to distinguish language impaired-children from language-normal children. While not claiming that the tests are fully comprehensive, the authors maintain that they afford the clinician with a means of looking at components of language while also gaining some insights into the child's ability to integrate vocabulary and grammar. It is recognized that no test can fully assess every aspect of language; essentially a test must be practical in terms of design and, for screening purposes, sufficiently comprehensive to be used with a large section of children referred. In addition to fulfilling these criteria, the Scales have been designed to provide some diagnostic indicators of areas of difficulty.

The aims of the original Scales were to provide a measure of verbal ability that could be compared with mental age, and to provide separate scores for production and comprehension. These aims have also been realized in the new version. Additionally a test has been provided that is consonant with information gained from the intensive research in the field of child language that has taken place since the original Scales were devised. The Scales can reveal information about a child's language abilities not readily available from observation and serve a useful function in both identifying disorder and evaluating therapy. Clinicians are increasingly being asked to demonstrate effectiveness of therapy but are hampered by inappropriate outcome measures (Law 1997). It is hoped that the RDLS III provides a test that is valid and reliable and one that can contribute to both the characterization of speech and language disorders and the measurement of treatment efficacy.

Acknowledgements

We acknowledge the contribution of Paul Fletcher who, as a member of the development team, contributed to the design, development and construction of the new Scales and manual on which this paper is based. We also acknowledge the

contribution of our field workers without whom the project would not have been completed so swiftly.

References

Armstrong, S. and Ainley, M., 1988, *South Tyneside Assessment of Syntactic Structures* (Northumberland: STASS Publications).

Bates, E., Dale, P. S. and Thal, D., 1995, Individual differences and their implications for theories of language development. In P. Fletcher and B. MacWhinney (Eds) *The Handbook of Child Language* (Oxford: Blackwell), pp. 96–151.

Beech, J., Harding, L. and Hilton-Jones, D. (Eds), 1993, *Assessment in Speech and Language Therapy* (London: Routledge), pp. 35–48.

Benedict, H., 1979, Early lexical development: comprehension and production. *Journal of Child Language*, **6**, 183–200.

Berman, R. A., 1986, A crosslinguistic perspective: morphology and syntax. In P. Fletcher and M. Garman (Eds) *Language Acquisition: Studies in First Language Development* 2nd edn (London: CUP), pp. 429–454.

Bishop, D. M. V., 1989, *Test for the Reception of Grammar* 2nd edn (Cambridge: MRC).

Bishop, D. M. V. and Edmondson, A., 1987, Language impaired 4-year olds: distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders*, **52**, 156–173.

Bloom, L. and Lahey, M., 1978, *Language Development and Language Disorders* (New York: Wiley).

Brown, R., 1973, *A First Language: the Early Stages* (London: Allen & Unwin).

Chapman, R. S., 1988, Language acquisition in the child. In N. Lass, J. Northern and D. Yoder (Eds) *Handbook of Speech-language Pathology and Audiology* (Philadelphia, BC: Decker), pp. 309–353.

Crain, S., 1991, Language acquisition in the absence of experience. *Behavioural and Brain Sciences*, **14**, 597–611.

Crystal, D., Fletcher, P. and Garman, M., 1989, *The Grammatical Analysis of Language Disability* 2nd edn (London: Whurr).

Dale, P. and Fensen, L., 1996, Lexical development norms for young children. *Behaviour Research Methods, Instruments and Computer*, **28**, 125–127.

Dewart, H. and Summers, S., 1995, *The Pragmatics Profile of Early Communication Skills* 2nd edn (Windsor: NFER-Nelson).

DiSimoni, F. G., 1978, *Token Test for Children* (Allen, TX: DLM Teaching Resources).

Dunn, L. M., Dunn, L. M., Whetton, C. and Burley, J., 1997, *The British Picture Vocabulary Scale* 2nd edn (Windsor: NFER-Nelson).

Dunn, L. M., Dunn, L. M., Whetton, C. and Pintillie, D., 1982, *The British Picture Vocabulary Scale* (Windsor: NFER-Nelson).

Edwards, S., Fletcher, P., Garman, M., Hughes, A., Letts, C. and Sinka, I., 1997, *The Reynell Developmental Language Scales III: The University of Reading Edition* (Windsor: NFER-Nelson).

Fensen, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J. and Hartnung, J., 1994, Variability in early communicative development. *Monographs of the Society for Research in Child Development*, **59**, 242.

Fensen, L., Dale, P. S., Reznik, S., Thal, D., Bates, E., Hartnung, J. P., Pethick, S. and Reilly, J. S., 1993, *The McArthur Communicative Developmental Inventories* (San Diego: Singular Publishing Group).

Fletcher, P., 1985, *A Child's Learning of English* (Oxford: Blackwell).

Fletcher, P. and Peters, J., 1984, Characterising language impairment in children: an exploratory study. *Language Testing*, **1**, 33–49.

Garman, M. and Edwards, S., 1995, The syntactic assessment of expressive language. In K. Grundy (Ed.) *Linguistics in Clinical Practice* 2nd edn (London: Whurr), pp. 134–166.

Gentner, D., 1982, Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.) *Language Development, Vol. 2: Language, Thought and Culture* (Hillsdale, NJ: Erlbaum), pp. 301–334.

German, D. J., 1986, *Test of Word Finding* (Allen, TX: DLM Teaching Resources).

Golinkoff, R. M. and Hirsh-Pasek, K., 1995, Re-interpreting children's sentence comprehension: toward a new framework. In P. Fletcher and B. MacWhinney (Eds) *The Handbook of Child Language* (Oxford: Blackwell), pp. 429–461.

Gutfreund, M., Harrison, M. and Wells, G., 1989, *The Bristol Language Development Scales* (Windsor: NFER-Nelson).

Huntley, M., 1985, *Reynell Developmental Language Scales Manual* (2nd revision) (Windsor: NFER-Nelson).

Johnston, J. R. and Kamhi, A., 1984, Syntactic and semantic aspects of the utterances of language-impaired children: the same can be less. *Merrill–Palmer Quarterly*, **30**, 65–86.

Johnston, J. R. and Slobin, D. I., 1979, The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language*, **6**, 529–545.

Karmiloff-Smith, A., 1993, Innate constraints and developmental change. In P. Bloom (Ed.) *Language Acquisition: Core Readings* (London: Harvester Wheatsheaf).

King, G., Schelletter, C., Sinka, I., Fletcher, P. and Ingham, R., 1995, Are English-speaking SLI children with morpho-syntactic deficits impaired in their use of locative-contact and causative alternating verbs? In P. Kerswill, R. Ingham, Y. Huang and L. Shockey (Eds) *Reading Working Papers in Linguistics 2* (Reading: The University of Reading), pp. 45–65.

Knowles, W. and Masidlover, M., 1987, *The Derbyshire Language Scheme* (revised) (Derbyshire: Educational Testing Service).

Law, J., 1997, Evaluating intervention for language impaired children: a review of the literature. *European Journal of Disorders of Communication*, **32**, 1–14.

Lees, J., 1993, Assessment of receptive language. In J. Beech, L. Harding and D. Hilton-Jones (Eds) *Assessment in Speech and Language Therapy* (London: Routledge), pp. 16–34.

Leonard, L., Sabbadini, L., Leonard, J. and Volterra, V., 1987, Specific language impairment in children: a cross-linguistic study. *Brain and Language*, **32**, 233–252.

MacWhinney, B., 1995, Computational analysis of interactions. In P. Fletcher and B. MacWhinney (Eds) *The Handbook of Child Language* (Oxford: Blackwell), pp. 152–178.

Maratsos, M. P., Fox, D. E., Becker, J. A. and Chalkley, M. A., 1985, Semantic restrictions on children's passives. *Cognition*, **19**, 167–191.

Miller, J. F., 1981, *Assessing Language Production in Children: Experimental Procedures* (Baltimore, MD: University Park Press).

Miller, J. F. and Leadholm, B. J., 1992, *Language Sample Analysis: the Wisconsin Guide* (Madison, WI: Bureau for Exceptional Children, Wisconsin Department of Public Instruction).

Miller, J. F. and Paul, R., 1995, *The Clinical Assessment of Language Comprehension* (Baltimore: Paul Brookes Publishing).

Milosky, L., 1992, Children listening: the role of world knowledge in language comprehension. In R. S. Chapman (Ed.) *Processes in Language Acquisition and Disorders* (St Louis, MO: Mosby Year Book), pp. 20–44.

Raban, B., 1988, *The Spoken Vocabulary of Five-Year Old Children* (Reading: Reading and Language Information Centre, The University of Reading).

Radford, A., 1995, Phrase structure and functional categories. In P. Fletcher and B. MacWhinney (Eds) *The Handbook of Child Language* (Oxford: Blackwell), pp. 483–507.

Renfrew, C. E., 1997a, *Action Picture Test* 4th edn (Bicester, UK: Winslow Press).

Renfrew, C. E., 1997b, *Bus Story Test* 4th edn (Bicester, UK: Winslow Press).

Reynell, J., 1977, *Reynell Developmental Language Scales* (revised) (Windsor: NFER).

Reynell, J. and Huntley, M., 1985, *Reynell Developmental Language Scales* (2nd revision) (Windsor: NFER-Nelson).

Rice, M., Wexler, K. and Cleave, P., 1995, Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research*, **38**, 850–863.

Richards, B., 1993, *Language Development and Individual Differences: a Study of Auxiliary Verb Growth* (Cambridge: Cambridge University Press).

Rodd, L. and Braine, M. D. S., 1970, Children's imitations of syntactic constructions as a measure of linguistic competence. *Journal of Verbal Learning and Verbal Behaviour*, **10**, 430–443.

Schelletter, C., Sinka, I., Fletcher, P. and Ingham, R., 1998, English-speaking SLI children's use of locative/contact and causative alternations. In M .Garman, C. Letts, B. Richards, C. Schelletter and S. Edwards (Eds) *Proceedings of the Child Language Seminar 1996* (Bulmershe Papers: The University of Reading) (In press).

Semel, E., Wiig, E. H. and Secord, W., 1987, *Clinical Evaluation of Language Fundamentals—Revised* (NY: Psychological Corporation).

Slobin, D. I. and Bever, T. G., 1982, Children use canonical sentence schemas: a cross-linguistic study of word order and inflections. *Cognition*, **12**, 229–265.

Slobin, D. I. and Welsh, C. A., 1973, Elicited imitation as a research tool in developmental psycholinguistics. In C. A. Ferguson and D. I. Slobin (Eds) *Studies in Child Language Development* (New York: Holt, Rinehart & Winston), pp. 485–497.

Sutter, J. and Johnson, C., 1990, School-aged children's metalinguistic awareness of grammaticality in verb forms. *Journal of Speech and Hearing Research*, **33**, 84–95.

Tager-Flusberg, H. and Calkins, S., 1990, Does imitation facilitate the acquisition of grammar? Evidence from a study of autistic, Down syndrome and normal children. *Journal of Child Language*, **17**, 591–606.

van der Gaag, A., 1996, *Communicating Quality* (London: Royal College of Speech and Language Therapists).

Wells, G., 1985, *Language Development in the Pre-school Years* (Cambridge: Cambridge University Press).

Wiig, E. H. and Semel, E., 1984, *Language Assessment and Intervention for the Learning Disabled* 2nd edn (Columbus, OH: Merrill Publishing Company).

Wiig, E. H., Secord, W., and Semel, E., 1992, *Pre-school CELF* (NY: Psychological Corporation).

Wirz, S., 1993, Introduction: historical considerations in assessment. In J. Beech, L. Harding and D. Hilton-Jones (Eds) *Assessment in Speech and Language Therapy* (London: Routledge), pp. 1–15.

Wong, V., Lee, P., Lieh-Mak, F., Yeung, C., Leung, P., Luk, S. and Yiu, E., 1992, Language screening in pre-school Chinese children. *European Journal of Disorders of Communication*, **27**, 247–264.

Address correspondence to: Carolyn Letts, Department of Linguistic Science, The University of Reading, Reading RG6 6AA, UK.

Commentaries upon this paper follow below and the response to these commentaries appears on pp. 184–191.

Reynell Developmental Language Scales III: a quick and easy LARSP?

Martin J. Ball

University of Ulster, County Antrim, N. Ireland

It is 12 years since we last had a revision of the Reynell Developmental Language Scales, and so version III (RDLS III) has been awaited with interest. This is particularly so in that, to a large extent, the previous version was considered more as a test of comprehension than production: and RDLS III promises to bring more balance to these two aspects of assessment. It is also of interest, of course, that RDLS III comes from a department well known for its impact in the area of linguistic assessments in the form of profiles; and we will want to examine how this work informs the authors in their major restructuring of the Reynell scales. In this commentary, I will concentrate mainly on the Expressive Scale.

Indeed, we may wonder why the Reading department undertook to revise the Reynell, when we consider some of the criticisms levelled at earlier versions. For example, Crystal *et al.* (1989) attack the 1969 version of the Reynell for its selective approach to syntactic analysis. Further, the same authors are not much more encouraging in their reaction to assessments of syntax that produce overall scores. For example, they state that 'syntax scores have often been tried,... but we have found them unilluminating, and at times counter-intuitive' (Crystal *et al.* 1989, p. 18). It is the problems of the previous neglect of syntax in the existing assessments that lead to the development of profiles: initially the grammatical profile LARSP.