Assessing Language Skills in Preschool Children

Julie E. Dockrell

This review outlines some of the difficulties involved in assessing preschool children's language development. Language is a sensitive indicator of a range of developmental difficulties yet the accurate identification of children who are experiencing delays or disorders is problematic. A range of different approaches are used to identify and assess language problems including norm-referenced assessments, questionnaires and language samples. Each of these is critically evaluated. In particular the reliability and validity of the measures and their ability to discriminate atypical patterns of development is considered. It is concluded that there are no simple assessment tools that identify and assess language development. Moreover, the use of single measures is considered inadequate for determining whether a child at any age has typical or delayed language. Ways to circumvent the limitations of the assessment tools are considered.

Keywords: Preschool; language; assessment

Introduction

Language delays and difficulties comprise one of the most common problems in the preschool years. Such problems are initially identified through one of two routes—referral or screening. Following initial identification detailed assessments should be undertaken when a child is suspected of having a language delay or difficulty. Referrals frequently come from parents but are also initiated by other professionals such as teachers and general practitioners. The appropriateness of the referrals will depend on the referrer's sensitivity to problems with the developing language system. In contrast, language screening procedures are generally population measures that aim to identify children who are experiencing a problem with the language system. Their accuracy will depend on the reliability and validity of the measure used to screen the child's language performance. This paper addresses issues critical to the process of identifying and assessing preschool children with suspected language delays or difficulties. No attempt is made to identify the ideal test or assessment instrument; instead the focus is on the dimensions that professionals should consider when assessing language.

Julie E. Dockrell
Psychology & Special Needs, Institute of Education,
20 Bedford Way,
London WC1A 0AL

Delays or differences in patterns of language acquisition have been consistently viewed as indicators of developmental problems. Language is a sensitive indicator of neuromotor impairment, hearing loss, general learning disabilities and specific language and communication difficulties. Failure to follow typical trajectories in language and communication is both a risk factor for later language difficulties (Whitehurst & Fischel, 1994) and an indicator of potential difficulties with literacy (Catts & Kahmi, 1999; Catts et al., 1999), numeracy (Fazio, 1994, 1996; Grauberg, 1998) and socio/behavioural aspects of development (Benaisch, Curtiss, & Tallal, 1993; Botting & Conti-Ramsden, 2000; Lindsay & Dockrell, 2000). Prospective studies of children with speech and language difficulties have indicated that there are high rates of continued communication difficulties in this population. Children with a specific language problem appear to have a more favourable prognosis than those with language impairments secondary to sensory, structural, neurological or cognitive problems (Johnson et al., 1999). Thus the accurate identification and assessment of language problems is of central concern for the appropriate management of interventions, planning educational placements and to support children and their families. Achieving this objective is not straightforward. There are certain features of the language system that make it complicated to assess. It is important to bear these in mind when

considering what measures to use and how to interpret the results of the assessments.

Language is multidimensional and as such does not easily lend itself to single unitary measures. Comprehension and expression (production) must both be examined as well as the more subtle aspects of the language system, such as pragmatics. By implication this means that it will be necessary to profile a range of skills to achieve a valid picture of the child's language performance. Moreover, interpreting performance on the tasks that are used is complicated by the fact that there is much variation in normal patterns of development, especially in the early years. For example, expressive vocabulary size may range from fewer than 9 to well over 198 words in typically developing 16-month-old children and from fewer than 41 to well over 405 words in typically developing 20-month-olds (Fenson et al., 1993). Variability is typical in rate, style and profile of language development. Equally, performance on specific tasks can vary from one testing situation to another (Allen & Wightman, 1994, for a consideration of performance on auditory temporal processing tasks). These factors can make it hard to draw precise boundaries between typical and atypical patterns of development.

My aim in this review is (1) to consider, briefly, the key dimensions of the language system to provide a marker against which assessment processes can be evaluated, (2) to identify the three common assessment frameworks that provide evidence about children's language development and (3) to consider the limitations of current approaches to assessing language development. The primary focus of this paper is the assessment of the child's *language performance*. I take it as given that the assessments should be developmentally and culturally appropriate, taking into account both cognitive and social dimensions. Failure to consider these issues will result in invalid and unreliable assessments (Huang, Hopkins, & Nippold, 1997).

Key dimensions of the language system

The language system is composed of a number of subcomponents that are important for effective language use. These include semantics, the lexicon, grammar, phonology and pragmatics (see for a discussion Messer, 1994). These subcomponents work together in dynamic and developmental fashion. Thus, for example, delays in lexical acquisition can impact on the fluency of children's expressive language, while failures to process key syntactic components may impact on social interaction. Thinking of language as a dynamic system highlights the ways in which even minor problems can affect a child's ability to access and contribute to daily experiences. These difficulties can result in a variety of different developmental trajectories (Leonard, 1997). Some of the most detailed work on plotting patterns of language difficulties has occurred with children who have specific language problems.

We have known for some time that children with specific language impairment (SLI) form a heterogeneous population in terms of language difficulties. Analyses of children's language from both clinical (Rapin & Allen, 1987) and controlled experimental studies (Conti-Ramsden, Crutchley, & Botting, 1997; Conti-Ramsden & Botting, 1999) have shown that children designated as SLI can have varying patterns of difficulties with the subcomponents of the language system. Both the DSM-IV (American Psychiatric Association, 1994) and ICD 10 (World Health Organisation, 1990) now identify subgroups for specific language impairment. However, such classification approaches are not immune from criticism. An alternative way of analysing language skills is to describe the children's language in detail through using a linguistic framework (Bishop, 1987; Fletcher, 1991).

Thus, assessments of children's language skills require that the various subcomponents of the language system are considered in creating a profile of the children's skills. It must also be realised that this profile is likely to change. Children's language develops and by corollary the nature of their language problems are likely to change (Conti-Ramsden & Botting, 1999). As Whitehurst and Fischel (1994) point out: 'The frequency and patterning of language delays vary so greatly with age that data taken from one age group … may have little relevance to other age groups' (p. 639).

Identifying subgroups of children is particularly problematic for preschool children where the primary problem is the difficulty in distinguishing enduring from transient difficulties. In a longitudinal study by Silva, Williams, & McGee (1997) more children moved out of the category of language impaired between the ages of 3 to 5 than stayed within it. Yet at the same time other children who were not delayed at age 3 were delayed at age 5. The clear message for the practitioner is that to identify language problems children's current language performance must be sampled across a range of linguistic skills and children's progress must be monitored. Simply because a child was not identified at 3 years does not mean that a problem may not be evident at 5 years.

For purposes of assessment and subsequent intervention it is more helpful to consider a hierarchy of vulnerability in language functions (Bishop & Edmundson, 1987). Disorders of comprehension reflect the most severe form of language impairment, followed in order of decreasing severity by disorders of expressive semantics, disorders of morphology and syntax and disorders of phonology. As a group, children with expressive problems tend to progress well (Bishop & Adams 1990; Paul, 1996). When we consider academic performance there are also positive indicators. Whitehurst and Fischel (1994) reported that reading and mathematics scores obtained from school records of children who presented as late talkers were well within the normal range when these children were 7 years old. However, recent data from the original Bishop sample (Stothard, Snowling, & Bishop, 1998) suggest that even

the group of children whose language problems had resolved performed significantly less well than typical children on phonological processing and literacy skills at age 15 to 16. Prognosis for individual children will depend on the nature of the early language problems. From a clinical perspective the implications are clear: early and accurate profiling of children's language problems can play an important first step in supporting later learning and literacy development.

A long-standing concern for practitioners has been the discrepancy between children's verbal and non-verbal skills (Leonard, 1987). Discrepancy criteria have frequently been used in attempts to identify children with specific language impairment, where a contrast is drawn between language skills and non-verbal ability. However, allusions to discrepancies are problematic. There are both conceptual and practical issues to consider (Dockrell & McShane, 1993). At a practical level there are concerns about measurement and the determination of the appropriate formula for the discrepancy (Aram, Morris, & Hall, 1992; Plante, 1998). Problems include the methods for measuring the discrepancy, the meaningfulness of the distinction and the accuracy of the normative data being used to conclude that a discrepancy exists. In addition, the language problems the children have may impact on their ability to succeed on non-verbal tasks as well (see, for example, Johnston & Smith, 1989; Johnston, 1991, 1994). Practitioners may need to place less reliance on simplistic models of discrepancy between non-verbal and verbal skills and make greater attempts to characterise the child's performance on different tasks and situations, resulting in a profile of skills and needs.

The complexity of the language system poses many challenges for the practitioner. The first step for any practitioner is to establish the nature and extent of a child's difficulty and to consider this in relation to the child's age and developmental norms (where they exist). Interpretation of the results of any assessment will be constrained by how solid the measures are, the expertise in their administration and the assessor's skills of interpretation. It is to some of these issues I now turn.

Approaches to identifying and assessing language difficulties

Identification and assessment are not the same processes (Lahey, 1990). The purpose of identification is to distinguish between children who do and those who do not have a language or communication problem. The judgement should be based on the child's language performance at the point that the identification is being made. Furthermore, the child's performance should be compared with children of a similar age. In such situations expectations of normal development need to be explicit.

Language problems are often identified through screening procedures. A range of screening procedures exist to identify language problems (see Law, 1992). Some tests are designed to identify the presence or absence of the language problem at the time of testing, that is, they are concerned with *concurrent* difficulties. Other tests are designed to *predict* the likelihood of a child experiencing language difficulties in the future. In such cases children are thought to be 'at risk'. Screening measures designed to identify children who are at risk are problematic (Lindsay, 1995; Lindsay & Desforges, 1998). However, even screening tests that are designed to identify the presence of a current language problem are far from 100% accurate. In a recent comprehensive review of screening tests for early language delay Law and his colleagues (Law et al., 1998; Laing et al., 2000) conclude that while we know who is not language delayed there is continued disagreement about who is experiencing a language difficulty. The major problems surround the balance between specificity, that is accurate identification, and sensitivity, not identifying children who do not have a language problem. Nor is it clear what the implications are for children who are missed or over-referred. Interestingly, in their own British study, Laing et al. (2000) found that parents appear to be almost as accurate as screening tests although health visitors expressed a preference for the test. Law concludes that there is a need to move towards a system where initial intervention is seen as a means of dynamic assessment rather than intervention per se. This places a major responsibility on the clinician's ability to generate and test relevant hypotheses.

Identifying the existence of a problem is the first step in the language assessment process. The next is to characterise the nature and extent of the child's difficulties in terms of differing language skills. A broad range of information-gathering activities are available to meet this goal. The assessment process itself will be guided by the initial evaluation of the child, the theoretical orientation held by the assessor and practical constraints related to time and resources. Three broad approaches to the assessment of language problems can be identified – standardised tests, analysis of language samples, and parental or teacher questionnaires. These approaches are not mutually exclusive.

Standardised norm-referenced assessments

Published standardised tests for expressive and receptive language are commonly used both in clinical practice (Wilson et al., 1991, Dockrell et al., 1997) and for research studies (McCauley & Demetras, 1990). Such tests aim to provide objective information. Tests should be based on an appropriate standardisation sample, and therefore provide a reliable measure of a child's relative standing in comparison to developmental language norms. However, psychometric adequacy is not necessarily the determining factor in test choice. It is not surprising, given the demands in clinical practice, that practitioners report that test choice is based on whether they are quick and easy (Huang et al., 1997). Test choice varies but commonly used tests for clinical work include the CELF (Clinical Evaluation of Language Functions; Semel, Wiig, & Secord, 1987) in

Table 1. Commonly used tools in assessing children's language development[1]

Language dimension	Assessment tool	Specific aims	Age range	Total standardisation sample[2]	Reliability assessments reported in manual or research papers	Validity assessments reported	Discussions critiques	Gaps or limitations identified	Publisher
Global language measures	Reynell-III (1997)[D]	Language comprehension and expression within a framework of developmental progress + highlights for therapy	1;6–7;00 for comprehension 1;9–7;00 for expression	1,074	Yes	Yes – concurrent	International Journal of Language and Communication Disorders-clinical forum (1999, 34)	No direct assessment of vocabulary or pragmatics	Windsor-NFER Nelson
	CELF-R (1987)[D]	Identification, diagnosis and follow-up evaluation of language skill deficits in school age children. Sub-test standard scores are available	5;0-16;11	2,426 – USA standardisation	Yes	Yes – see manual	Summers et al. (1996); Spekman & Roth (1984)	See critique	Psychological Corporation
	CELF preschool	Broad range of expressive and receptive language skills	3;0-6;0	800 USA standardisation	Yes	Yes	www.tpcweb.com		
Measures of language subsystems	MacArthur Communicative Development Inventories[D]	1. gestures 2. lexical comprehension and production 3. early sentence	8 months – 30 months		Yes	Yes	www.tpcweb.com		Psychological Corporation
	British Picture Vocabulary Scale (BPVS-II) (1997)[D]	Receptive vocabulary for standard English	3;0-15;00	2,571	Yes see manual	Evidence for concurrent validity and predictive validity (Lewis, 1987; Howlin & Cross, 1994)			Windsor-NFER Nelson
	British abilities naming sub-test	Naming ability for single pictures described as verbal knowledge expressive	2;6-7;11	1,726 (from original sample)	Yes split-half correlation (see technical manual)	Yes Whippsi (see technical manual)		Limited number of items but standard errors provided	Windsor-NFER Nelson
	British abilities verbal comprehension	Comprehension of single words and simple sentences	2;6-7;11	1,726 (from original sample)	Yes split-half correlation (see technical manual)	Yes (see technical manual)		Includes single objects and actions and sentences used with concrete materials	Windsor-NFER Nelson
	Renfrew action picture test (1997)	Elicits descriptions of 10 action pictures which aims to distinguish between grammar and information	3;0-8;0	594	None reported	None reported			Bicester: Winslow
	Renfrew word-finding vocabulary test (1995)	Assesses object naming	3;6 to 8;5	741	None reported	None reported		No assessment of gap between comprehension and production	Bicester: Winslow
	Renfrew Bus story test (1996)	Children required to retell a story – provides information score and average sentence length	3;6 to 8;0	573	Concurrent validity in two studies ($n = 27$)	Indicates disorder & predictive of long-term disorder (Bishop & Edmundson, 1987		Borderlines may be particularly difficult to score	Bicester: Winslow
	Test of reception of grammar (Bishop, 1983)	Understanding of grammatical contrast	4:0 to 12:0	2000		None reported			University of Manchester

1[D] indicates a wider discussion in the text

both North America and Britain (Dockrell et al., 1997; Huang et al., 1997; Wilson et al., 1991) and the PPVT (Peabody Picture Vocabulary Test) in North America (Huang et al., 1997; Wilson et al., 1991). Test choice for research projects is somewhat different (see McCauley & Demetras, 1990 for a review of 70 journal articles) although again there is evidence that use of vocabulary measures is high. Table 1 provides descriptive information about a selection of common assessments.

Standardised language tests can be separated into those that provide general language measures and single tests

that provide measures about a child's performance on a specific component of the language system (see Wiig & Semel, 1980). Specific tests exist that cover virtually all dimensions of the language system in terms of both comprehension and expression. However, single measures of language are consistently inadequate for determining whether a child is developing typically or is experiencing a delay at any age and they become less and less reliable for younger and younger children (Thal & Katch, 1996). Of particular concern is the reporting of vocabulary scores as if they were indicators of general language ability. This may well result in inappropriate decisions about the child's language competence. For some children with language impairments vocabulary scores can be well within the norm despite wider problems with receptive and expressive language (Lahey & Edwards, 1995). Summarising a systematic assessment of four vocabulary tests with a group of preschool children, Gray et al. (1999) conclude that there was no support for the use of vocabulary tests in the identification of language problems. Specific tests are best conceived of as further evidence about the nature of a language problem subsequent to establishing that a language difficulty has already been identified. One way to identify more general difficulties is through the use of a global language measure.

A number of tests exist that purport to sample the breadth of children's language skills. In general these tests cover elements of vocabulary and grammar for both comprehension and production. They allow for areas of weakness to be identified for further testing and intervention. One of the most popular assessments in the United Kingdom, the Reynell Developmental Language Scales, has recently been revised (the Reynell Developmental Language Scales III, Edwards et al., 1997). The test taps language skills (both comprehension and production) in the age ranges 1:06–7:00 years (Edwards, 1998) and has a clear rationale based on recent linguistic research. The revised Reynell offers the benefits of assessments of concurrent validity in relation to the BPVS and the TROG as well as the inclusion of expressive language measures that can help analyses of classroom demands on the child's language skills (Lees, 1999). However, it is not possible to identify statistical differences between comprehension and expression and pragmatic skills are not assessed at all. There is currently no evidence to support the view that the test can monitor the efficacy of interventions.

The alternative measure commonly used in clinical practice is the Clinical Evaluation of Language Fundamentals (CELF). The CELF is designed for use with children from the preschool years to adolescence. It is designed to provide measures of both receptive and expressive language skills in the areas of phonology, syntax, semantics and memory and word finding and retrieval. Each subtest can be assessed independently. The test aims 1) to assist in the identification of children with language disabilities, 2) to provide a differential diagnosis of the areas of involvement and 3) to identify areas for follow-up for language intervention. Concerns have been raised about the lack of a theoretical framework as a basis for subtest selection and the limited data to support the test constructors' test aims. Of particular concern in clinical settings is the fact that subtests may be tapping a range of skills. As such, some subtest failure may be explained in different ways and it is important that test users consider a range of different explanations for subtest failure (Spekman & Roth, 1984).

This test now has a British standardisation (Peers, Lloyd, & Foster, 1999). The basis for the UK standardisation and the resultant changes raise important issues about use of tests on populations where the test has not already been standardised. Whilst the overall patterns of scores were found to be very similar in both the US and the UK, the US raw subtest score means tended to be consistently lower than the comparable UK scores, resulting in US norms overestimating language ability in the younger age groups. As Peers and Lloyd argue (personal communication) this is well illustrated with reference to one of the subtest core scales, 'Linguistic Concepts'. For the age group 3.0–3.5, the mean raw score for the UK normative sample was 10.3 (*SD* 4.2) compared with a mean of 8.1 (*SD* 4.3) in the US normative sample. This represents a statistically significant effect size indicating important differences in performance. Using the UK norms this mean raw score of 10.3 equates, as expected, to a standard score of 10. However, if the US norms are used this raw score equates to a standard score of 12. This equated score, therefore, represents an inflation of 2 standard score points for the UK standard score. A similar pattern was repeated for the other subtests in CELF and in some of the CELF 3 subtests. A list of detailed changes to the test is presented in Appendix 1. The CELF is one test where we now have empirical evidence about the importance of the standardisation sample. Such differences are likely to be equally important for other tests where comparative data is lacking and serve to reinforce the view that standardised tests *on their own* are not a satisfactory way of making decisions about individual children.

The use of age-equivalent scores from standardised tests is of particular concern with language measures (McCauley & Swisher, 1984b). Age-equivalent scores are frequently used to provide accessible information to parents or teachers or to indicate level of delay or impairment. Use of such measures are dangerous on a number of grounds. First, age-equivalent scores are often made by extrapolation from a child's score that lies midway between two age scores on the particular test. There is, however, no reason to assume that a score mid-way equates with age-equivalent midpoint score. Second, simply because a child scores several months lower than their equivalent chronological age does not necessarily mean they are significantly delayed. Much will depend on normal variation on the particular dimension assessed. As we saw earlier there is marked variation in the rate of development, variation

that does not necessarily mean delay. Finally, age-equivalent scores do not necessarily mean that a child is performing in an equivalent fashion to a peer of a younger age. A range of factors affect performance and additional language and social experience will be one such factor (McCauley & Swisher, 1984b). Often such scores are made up by a child's results on a number of subtests and there is no a priori reason to assume that the pattern of results will be the same across children of different chronological ages but with the same total score.

Summary

Standardised tests to assess children's language skills have a number of limitations. To characterise the nature and differences between tests a comparative performance of tests is needed, preferably initially on a typically developing sample (see Nation & Snowling, 1997 for this approach with reading). In general, such data are not available for language tests. As evidence to support this view, consider the work of Howlin and Cross (1994). They have shown that even with children developing apparently normally, there is much variation in their scores on different language measures. They tested children on six language measures and found that while on some measures the children's results appeared normal, on other measures they showed a marked discrepancy from their chronological age. They conclude, 'No test, however well designed, can ever be a substitute for careful observations and practical assessments of the child's communicative functioning'. Such results do not mean that assessing children's language should be stopped; rather it is necessary to consider a range of language skills in different contexts.

Two conclusions of practical import from recent critiques of standardised tests are relevant for practitioners. First, it is important to realise that such tests may have minimal value for individual programme planning or intervention strategies because they generally lack the number and variety of items necessary for planning and monitoring. Second, it is necessary to be clear what information can be acquired from standardised tests and what additional information is necessary for further evaluation. For example, Huang et al. (1997) take the position that: 'a test is a tool for observing behaviour and that its optimal use is the result of a complex interaction of the characteristics of the test, the client and the clinician' (p. 12). Thus there is increasing evidence that for the practitioner, standardised language tests should always be used in conjunction with other methods of assessment even when making only screening and placement decisions (Dale & Cole ,1991; Law, personal communication; Lund & Duchan, 1993).

Parental/teacher questionnaires

Difficulties in obtaining behavioural samples and the restricted context of both clinical and laboratory environments has led to a search for alternative reliable and valid ways of sampling a child's language development. One such method is the use of adult respondents to provide information about a child's language and communication skills. Parents, in particular, have the advantage of being able to provide representative information across a variety of settings. Such methods can be cost effective in terms of data collection, coding and analysis.

It has often been assumed that parent reports are prone to bias, forgetting and misinterpretation. However, there is increasing evidence that parents (particularly mothers) provide a reliable source of information on the communicative-linguistic development of their children (Klee et al., 1998). Further, parents are able to draw distinctions between verbal and non-verbal skills (Saudino et al., 1998). The validity and reliability of parental reporting is, however, dependent on three factors:

1. The information called for is current and not retrospective.
2. The skills are emergent.
3. Skills are identified by recognition and not recall (Dale et al., 1989).

One of the mostly widely used parental questionnaires is the MacArthur Communicative Development Inventories (CDI) (Fenson et al., 1993). These parent report inventories sample aspects of communicative development from 8 to 30 months. There are two separate forms. The CDI: Words and Gestures form designed for use with infants aged 8–16 months and the CDI: Words and Sentences form designed for use with toddlers from 16–30 months. The words and gestures form is divided into two parts: sampling receptive and expressive vocabulary, and the intentional and symbolic development of gestures. The words and sentences form also measures expressive vocabulary but includes utterance length and grammatical complexity. The CDI has proved to be a very powerful tool both for research and clinical work. The original measures are reliable and valid with typically developing children. Recent evidence indicates that the measures can also be used with delayed children functioning at appropriate level of the test (Thal, Tobias, & Morrison, 1991; Thal et al., 1999). However, in these situations parents' reports of productive vocabulary are more reliable than their assessments of comprehension. There are, however, particular difficulties with grammar portions of the CDI words and sentences where problems occur in differentiating between vocabulary and grammar for the language delayed children (Thal et al., 1999). This may reflect the test content or alternatively the two aspects of the linguistic system may not be as clearly differentiated in some language-delayed children. There is increasing evidence that the CDIs are valid more generally and can be profitably used for hypothesis generation to identify language factors that might be evaluated in greater detail.

The CDI has the additional benefit of being translated

and standardised in a wide variety of languages (see http://www.sci.sdsu.edu/cdi/foreign.html). More recently Philip Dale has developed a brief upward extension of the CDI approach that is suitable for assessing language skills in children between 30 and 42 months (see website for further information). The CDI III is a two-page questionnaire that includes a 100-item vocabulary list, 12 sentence pairs for assessing grammatical complexity, and 12 yes/no questions concerning semantics, pragmatics, and comprehension. Currently, however, there are only limited norming and validation data available. Information about test development and research projects related to the CDI can be found on their web site http://www.sci.sdsu.edu/cdi.

Despite the CDI's robustness at indexing the growth of language skills it has been criticised on a number of dimensions (Feldman et al., 2000). First, there is concern about the lack of specific data in terms of means and standard deviations at appropriate age points. Second, different subscales are subject to reporting biases, with those scales that require a greater degree of subjective interpretation being most vulnerable. Finally, the lack of predictive power, particularly at the younger ages, questions the role of the CDI as a screening device. However, Fenson et al. (2000) point out that the variability shown in the CDI is likely to be a true reflection of variability in the early ages of language development and that although prediction in the early years is low, prediction for the older children is better. They advise caution when interpreting scores at the youngest ages and for families from lower socio demographic backgrounds.

A range of other checklists exist for use with older children e.g., the Pragmatics profile (Dewart & Summers, 1995) which is completed by the parent or carer and considers communication up to the age of 10. More recently Bishop (1998) has developed the Children's Communicative Checklist (CCC) which has been evaluated on children aged 7 to 9 years. This checklist is designed to assess aspects of communication that are clinically important but 'not well covered by traditional language assessments'. Reliability and validity of the scale is, currently, limited to a particular age group of children who have already been identified as having language problems.

Summary

The use of parent and teacher measures holds significant potential for both practitioners and researchers alike. However, it is important to be aware of the boundary conditions of such tools. While these measures show high concurrent validity with behavioural measures for both vocabulary (Klee et al., 1998) and syntax (Dale, 1991), their general status as *predictive measures* has not been demonstrated. Nor do they provide information on the child's potential to benefit from intervention. Moreover, scales that place multiple demands on parents to observe and interpret various aspects are problematic (Stiles, 1994). In some cases it may be necessary to train parents (or groups of parents) to achieve reliable and valid results (Feldman et al., 2000).

Language samples

There is a long history of using language samples in research with typical and atypical populations. Language samples provide the practitioner with two important sources of information. For the practitioner they allow the assessment of language use *in vivo* and, unlike standardised tests, are not constrained by particular test items. Performance can be considered in relation to both context and speaker demands e.g., in classrooms, with peers, and so forth. Samples can be obtained either by recording spontaneous interactions or eliciting specific samples through verbal or visual prompts. Thus, the sampling and transcription process provides direct access to primary data. Crystal and colleagues (Crystal, 1982; Crystal, Fletcher, & Garman, 1989) place great stress on the use of spontaneous language to draw reliable conclusions about a child's language skills.

There is no clear consensus of the number of utterances required to draw valid inferences about the child's language levels. Early work suggested that a minimum of 50 utterances were required (Templin, 1957). More recently Miller (1996) has provided empirical data to support the view that 50 utterances provide sufficient data to sample speaking rate. Nonetheless, the ultimate goal must be to collect a sample that provides sufficient depth and breadth of both vocabulary and syntax to adequately reflect the child's language performance. The clinician or researcher must decide what measures to code to assess this performance. Mean length of utterance (MLU) is a commonly used metric but other measures are also possible, including measures of lexical diversity, syntactic complexity, pragmatic factors, rate, fluency, and error analyses. Such analyses can help clinicians identify and document specific language problems but they are dependent on the skills of the clinician both in terms of identifying the appropriate variables to code and interpreting the results appropriately. Nonetheless, time expended at initial assessment may well save time later on (Crystal et al., 1989).

Summary

Despite their many advantages, the sampling and transcription of language samples pose a number of problems. Elicited samples fall foul of the criticism that they are not likely to be representative of the child's actual abilities, interactions may be highly artificial and may be influenced by imbalanced relationships between the participants. Transcription is time consuming, with each 50 utterance sample taking from 1 to 3 hours to transcribe. Although a number of computer programmes (CHILDES, MacWhinney, 1996; Systematic Analysis of Language Transcripts (SALT) Miller & Chapman, 2000) exist to support

the analysis it can be complex and time consuming. Moreover, facilities for transcribing and charting language skills are not available in typical assessment situations where there is limited time to gauge a child's language competence. Nonetheless, when there are mismatches between test results and significant adults (teachers, parents, psychologists etc) language samples should be considered as an additional source of evidence.

New developments and conclusions

Reliable and valid assessments of a child's language development and communicative competence are of central importance for studying typical and atypical development. Trying to decide whether a child has a significant language problem or not can be daunting, even for experienced practitioners. The child, the particular tasks under investigation, and the context of learning will govern choice of the appropriate measures to address the specific questions of concern (Dockrell & Messer, 1999).

Standardised assessments and checklists focus on what a child can or cannot do on a limited set of test items (Miller, 1996). An overall score tells us neither about how a child approaches the task nor about which elements of the task they find difficult. Language learning is more complex than learning new vocabulary items and picking up new grammatical constructions. In an attempt to counteract this static approach to assessment new procedures are being tried in research contexts and some clinical settings. Two developments are currently under the focus of greater attention: process-based assessments and assessment of children's narratives. In process based approaches significance comes from what the child has learned during and after interaction with the assessor, who also provides pedagogical input. These assessments are designed to manipulate the context systematically to support learning and ultimately predict potential for change (Olswang & Bain, 1996; Bain & Olswang, 1995). Assessments of a child's learning potential offer a complementary and informative way to assess the child's language skills. As with other assessment tools it is important to establish the reliability and validity of such measures. Of particular concern will be issues related to inter-observer reliability.

By corollary, some researchers and clinicians are considering the use of narrative assessment as a means of understanding the child's developing communicative competence. Narratives include more elaborated units of texts than conversations and in addition contain an introduction, a logical and orderly presentation of results, and a conclusion. They appear to be an area of particular difficulty for children with language based problems and some authors would argue that an examination of narrative skills is an 'imperative in developing an overall profile of the student's communicative abilities' (Schoenbrot, Kumin, & Sloan, 1997, p. 271). Narrative assessments require the collection and analysis of language samples. However, the procedures are complex. The clinician must be able to assess both content and structure in a valid and reliable fashion. As yet there is not sufficient information to address these concerns.

Three key issues need to be considered as identification is attempted and assessment carried out. First, there is a need to consider a wide range of factors that impinge on the child's linguistic performance. These factors include appropriateness of materials, cultural factors, attention, and motivation. Assessment occurs in a context of relationships between the child and the task material. The setting in which the assessment occurs and what has been termed the 'social surround' (Messick, 1983) which includes the examiner, other children, and social expectancies such as sense of task orientation, apprehension at being evaluated, and the atmosphere of the testing session are central to the child's performance.

Second, when we consider the detailed assessment of language, the child's performance should be profiled across a range of language and cognitive dimensions. Sampling language skills in situations that stress the language system can provide valuable insights about the child's difficulties. This may well entail assessment in more complex situations, such as nursery classes or with peers. Underlying weaknesses in language may only emerge when a mismatch is provoked between the child's skills and strategies and demands of the learning environment. For example, a number of researchers have shown that children may have difficulties in accessing, organising and co-ordinating multiple mental activities simultaneously or in close succession. It is important not to dismiss clinical judgement or other potential sources of information. Multiple measures of early language skills and different professional expertise will enhance the reliability of the assessments.

Finally, the limitations of existing tools for assessing children's language and cognitive skills need to be considered. In particular, the practitioner must be aware of the reliability and validity of the chosen measures and their ability to discriminate atypical patterns of development. There is now strong evidence to show that there are psychometric inadequacies in many commonly used language measures (McCauley & Swisher, 1984a). Some measures lack concurrent validity (McCauley & Demetras, 1990). Moreover, reliance on any single measure to assess language difficulties is unreliable and invalid. Single measures are consistently inadequate for determining whether a child at any age has typical or delayed language (Miller, 1996). There are particular concerns when we consider the assessment of language in the early years. Standard tests are poor at predicting outcome for young language impaired children (Schery, 1985; Stark, Mellits, & Tallal, 1983) and even tests that pass relatively high numbers of psychometric criteria may not be precise discriminators of 4 and 5-year-old children (Plante & Vance, 1994).

These limitations should not mean that we stop using such measures altogether and use less reliable and poorly

understood measures instead. The fact that we are aware of the specific limitations of standardised measures should place us in a strong position to use them appropriately. Effectively, they provide a frame of reference to confirm or disconfirm our hypotheses. However, it is important to be aware that it is always possible to test a child and record some sort of numerical result. Thus, appropriately chosen standardised tests can provide current normative information about a child's language skills providing that the results are interpreted in an informed manner by an individual with the relevant expertise.

References

Allen, P., & Wightman, F. (1994). Psychometric functions for children's detection of tones in noise. *Journal of Speech and Hearing Research*, *37*, 205–215.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (3rd revised edition) (DSM-IV). Washington, DC: American Psychiatric Association.

Aram, D. M., Morris, R., & Hall, N. E. (1992). The validity of discrepancy criteria for identifying children with developmental language disorder. *Journal of Learning Disabilities*, *25*, 549–554.

Bain, B., & Olswang, L. (1995). Examining readiness for learning two-word utterances: Dynamic assessment validation. *Journal of Speech-Language Pathology*, *4*, 81–91.

Benaisch, A. A., Curtiss, S., & Tallal, P. (1993). Language, learning, and behavioural disturbances in childhood: A longitudinal perspective. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*, 585–594.

Bishop, D. V. M. (1983). *Test of reception of grammar*. Published by the author and available from Age and Cognitive Performance Research Centre, University of Manchester.

Bishop, D. V. M. (1987). The causes of specific developmental language disorder. *Journal of Child Psychology and Psychiatry*, *28*, 1–8.

Bishop, D. V. M. (1998). Development of the children's communication checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry*, *39*, 879–801.

Bishop, D. V. M., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry*, *31*, 1027–1050.

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

Botting, N., & Conti-Ramsden, G. (2000). Social and behavioural difficulties in children with language impairment. *Child Language Teaching and Therapy*, *16*, 105–120.

Catts, H. W., & Kamhi, A. G. (Eds.) (1999). *Language and reading disabilities*. Needham Heights, MA: Allyn & Bacon.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*, *3*, 331–361.

CHILDES: Child Language Data Exchange System. http://childes.pssy.cmu.edu

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language and Hearing Research*, *42*, 1195–1204.

Conti-Ramsden, G., Crutchley, A. C., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech and Hearing Research*, *4*, 765–777.

Crystal, D. (1982). *Profiling linguistic disability*. London: Edward Arnold.

Crystal, D., Fletcher, P., & Garman, M. (1989). *Grammatical analysis of language disability* (2nd ed.). London: Whurr.

Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech, Language and Hearing Research*, *34*, 565–571.

Dale, P., Bates, E., Reznick, J., & Morisset, C. (1989). The validity of a parent report instrument of child language at twenty months. *Journal of Child Language*, *16*, 239–249.

Dale, P. S., & Cole, K. N. (1991). What's normal? SLI in an individual differences perspective. *Language, Speech and Hearing Services in Schools*, *22*, 80–83.

Dewart, H., & Summers, S. (1995). *Pragmatic and early communication profile*. Windsor: NFER.

Dockrell, J. E., George, R., Lindsay, G., & Roux, J. (1997). Professionals' understanding of specific language impairments: Implications for assessment and identification. *Educational Psychology in Practice*, *13*, 27–35.

Dockrell, J. E., & McShane, J. (1993). *Children's learning difficulties: A cognitive approach*. Oxford: Basil Blackwell.

Dockrell, J. E., & Messer, D. (1999). *Children's language and communication difficulties: Understanding, identification and intervention*. London: Cassell.

Edwards, S. (1998). Testing children's language abilities: A description of a new language test: Reynell Developmental Language Scales. In W. Zeigler & K. Deger (Eds.), *Clinical linguistics and phonetics*. London: Whurr.

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

Fazio, B. (1994). The counting abilities of children with specific language impairment: A comparison of oral and gestural tasks. *Journal of Speech and Hearing Research*, *37*, 358–368.

Fazio, B. (1996). Mathematical abilities of children with Specific Language Impairment: A 2-year-follow-up. *Journal of Speech and Hearing Research*, *39*, 839–849.

Feldman, H., Dollaghan, C., Campbell, T., Kus-Lasky, M., Janosky, J., & Paradise, J. (2000). Measurement properties of the MacArthur Communicative Development Index at one and two years. *Child Development*, *71*, 310–322.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., Hartung, J., & Reilly, J. (1993). *Technical manual for the MacArthur communicative development inventory*. Developmental Psychology Laboratory, San Diego State University, San Diego, CA.

Fenson, L., Bates, E., Dale, P. S., Goodman, G., Reznick, J. S., & Thal, D. (2000). Measuring variability in early child language: Don't shoot the messenger. *Child Development*, *71*, 323–328.

Fletcher, P. (1991). Subgroups in school-age language-impaired children. In P. Fletcher & D. Hall (Eds.), *Specific speech and language disorders in children*. London: Whurr.

Grauberg, E. (1998). *Elementary mathematics and language difficulties*. London: Whurr.

Gray, S., Plante, E., Vance, R., & Henrichson, M. (1999). The diagnostic accuracy of four vocabulary tests administered to preschool children. *Language, Speech and Hearing Services in Schools*, *30*, 196–206.

Huang, R., Hopkins, J., & Nippold, M. (1997). Satisfaction with standardised language tests: A survey of speech language pathologists. *Language, Speech and Hearing Services in Schools*, *28*, 12–29.

Howlin, P., & Cross, P. (1994). The variability of language test scores in 3- and 4-year-old children of normal non-verbal intelligence: A brief research report. *European Journal of Disorders of Communication*, *29*, 279–288.

Johnson, C., Beitchman, J., Young, A., Escobar, M., Atkinson, L., Wilson, B., Brownlie, E. B., Douglas, L., Taback, N., Lam, I., & Wang, M. (1999). Fourteen-year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language and Hearing Research*, *42*, 744–760.

Johnston, J. R. (1991). Questions about cognition in children with specific language impairment. In J.F. Miller (Ed.), *Research on child language disorders: A decade of progress*. Austin, Tx: Pro-Ed.

Johnston, J. R. (1994). Cognitive abilities of children with language impairment. In R. Watkins & M. Rice (Eds.), *Specific language impairment in children*. Baltimore: Brookes.

Johnston, J. R., & Smith, L. (1989). Dimensional thinking in language-impaired children. *Journal of Speech and Hearing Research*, *26*, 33–38.

Klee, T., Carson, D., Gavin, W., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language and Hearing Research*, *41*, 627–641.

Lahey, M. (1990). Who shall be called language disordered? Some reflections and one perspective. *Journal of Speech and Language Disorders*, *55*, 612–620.

Lahey, M., & Edwards, J. (1995). Why do children with specific language impairment name pictures more slowly than their peers? *Journal of Speech and Hearing Research*, *39*, 1081–1098.

Law, J. (1992). The process of early identification. In J. Law (Ed.), *The early identification of language impairment in children*. London: Chapman Hall.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (1998). Screening for speech and language delay: A systematic review of the literature. *Health Technology Assessment*, *2*, 9.

Laing, G., Levin, A., Law, J., & Logan, S., (2000). *A comparison of two methods of identifying pre-school children with language problems*. Report submitted to the NHS Executive.

Lees, J. (1999). 'Which pig is not outside the field?' to 'which horse is not outside the field?': Commentary on the Reynell Developmental Language Scales III (RDLS III). *International Journal of Language and Communication Disorders*, *34*, 174–180.

Leonard, L. (1987). Is specific language impairment a useful construct? In S. Rosenberg (Ed.), *Advances in applied psycholinguistics* (*Vol. 1*). *Disorders of first-language development*. New York: Cambridge University Press.

Leonard, L. B. (1997). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Lewis, C. M. (1987). Vocabulary, sentences and words: Testing for agreement between two recent measures of language performance and receptive vocabulary. *Educational Psychology*, *7*, 129–132.

Lindsay, G. (1995). Early identification of special educational needs. In I. Lunt, B. Norwich & V. Varma (Eds.), *Psychology and education for special needs: Recent developments and future directions* (pp. 7–24). London: Arena, Ashgate Publishing.

Lindsay, G., & Desforges, M. (1998). *Baseline assessment*. London: David Fulton.

Lindsay, G., & Dockrell, J. E. (2000). The behaviour and self-esteem of children with specific speech and language difficulties. *British Journal of Educational Psychology*, *70*, 583–601.

Lund, N., & Duchan, J. (1993). *Assessing children's language in naturalistic contexts* (3rd ed.) Englewood Cliffs, NJ: Prentice Hall.

McCauley, R. J., & Demetras, M. J. (1990). The identification of language impairment in the selection of specifically language impaired subjects. *Journal of Speech and Hearing Disorders*, *55*, 468–475.

McCauley, R., & Swisher, L. (1984a). Psychometric review of language and articulation tests for children. *Journal of Speech and Hearing Disorders*, *49*, 34–42.

McCauley, R., & Swisher, L. (1984b). Use and misuse of norm-referenced tests in clinical assessment: A hypothetical case. *Journal of Speech and Hearing Disorders*, *49*, 338–348.

MacWhinney, B. (1996) The CHILDES System. *American Journal of Speech-Language Pathology*, *5*, 5–14.

Messer, D. (1994). *The development of communication: From social interaction to language*. Chichester, England: Wiley.

Messick, S. (1983). Assessment of children. In P. H. Mussen (Ed.), *Handbook of child psychology* (Vol. 1, 4th ed.) (pp. 475–526). New York: John Wiley & Sons.

Miller, J. (1996). Progress in assessing, describing and defining child language disorder. In K. Cole, P. Dale & P. Thal (Eds.), *Assessment and communication and language* (*Vol.* 6). Paul Baltimore, MA: H. Brookes.

Miller, J. F., & Chapman, R. S. (2000). Systematic Analysis of Language Transcripts (SALT) http://waisman.wisc.edu/salt

Olswang, L., & Bain, B. (1996). Assessment information for predicting upcoming change in language production. *Journal of Speech and Hearing Research*, *39*, 414–423.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, *67*, 359–370.

Paul, R. (1996). Clinical implications of the natural history of slow expressive language development. *American Journal of Speech and Language Pathology*, *5*, 5–21.

Peers I. S., Lloyd, P., & Foster, C. (1999). British Standardisation of the CELF. The Psychological Corporation's Speech and Language Assessment. http://www.tpcweb.com.

Plante, E. (1998). Criteria for SLI: The Stark and Tallal legacy and beyond. *Journal of Speech, Language and Hearing Research*, *41*, 951–957.

Plante, E., & Vance, R. (1994). Selection of preschool language tests: A data-based approach. *Language Speech and Hearing Services*, *25*, 15–24.

Rapin, I., & Allen, D. A. (1987). Developmental dyspraxia and autism in preschool children: Characteristics and subtypes. In J. Martin, P. Fletcher, P. Grunwell & D. Hall (Eds.), *Proceedings at the first international symposium on specific speech and language disorders in children* (pp. 20–35). London: Afasic.

Saudino, K. J., Dale, P. S., Oliver, B., Petrill, S. A., Richardson, V., Rutter, M., Simonoff, E., Stevenson, J., & Plomin, R. (1998). The validity of parent-based assessment of the cognitive abilities of 2-year-olds. *British Journal of Developmental Psychology*, *16*, 349–363.

Schery, T. K. (1985). Correlates of language development in language disordered children. *Journal of Speech and Hearing Disorders*, *50*, 73–83.

Schoenbrot, L., Kumin, L., & Sloan, J. (1997). Learning disabilities existing concomitantly with communication disorder. *Journal of Learning Disabilities*, *30*, 264–281.

Semel, E. M., Wiig, E., & Secord, W. (1987). *Clinical evaluation of language fundamentals-revised*. San Antonio TX: Psychological Corporation.

Silva, P. A., Williams, S. M., & McGee, R. (1987). A longitudinal study of children with developmental language delay at age three: Later intelligence, reading and behaviour problems. *Developmental Medicine and Child Neurology*, *29*, 630–640.

Spekman, N., & Roth, F. (1984). Clinical evaluation of language functions (CELF): An analysis and critique. *Journal of Speech and Hearing Disorders*, *49*, 97–100.

Stark, R. E., Mellits, E. D., & Tallal, P. (1983). *Behavioural attributes of speech and language disorders*. New York: Academic Press.

Stiles, J. (1994). On the nature of informant judgements in inventory measures: And so what is it you want to know? *Monographs of the Society for Research in Child Development*, *59*, 180–185.

Stothard, S. E., Snowling, M., & Bishop, D. V., (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech Language and Hearing Research*, *41*, 407–418.

Summers, P., Larson, G., Miguel, S., & Terrell, S. (1996). Test-retest comparisons using the CELF-RST and BLT-2S with kindergarteners. *Language*, *Speech and Hearing Services in Schools*, *27*, 324–329.

Templin, M. (1957) *Certain language skills in children: Their development and interrelationships*. Minneapolis: University of Minnesota Press.

Thal, D. J., & Katch, J. (1996). Predicaments in early identification of specific language impairment: Does the early bird always catch the worm? In K. Cole, P. Dale & P. Thal (Eds.), *Assessment and communication and language* (Vol. 6). Baltimore, MA: Paul H. Brookes.

Thal, D., O'Hanlon, K., Clemmons, M., & Fralin, L. (1999). Validity of parent report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech*, *Language and Hearing Research*, *42*, 482–496.

Thal, D., Tobias, S., & Morrison, D. (1991). Language and gesture in late talkers: A 1-year follow up. *Journal of Speech*, *Language and Hearing Research*, *34*, 604–612.

Whitehurst, G., & Fischel, J. (1994). Practitioner review: Early developmental language delay: What, if anything, should the clinician do about it? *Journal of Child Psychology and Psychiatry*, *35*, 613–648.

Wiig, E., & Semel, E. (1980). *Language assessment and intervention for the learning disabled*. Columbus, Ohio: Charles E. Merrill Publishing.

Wilson, K., Blackmon, R., Hall, R., & Elchotz, G. (1991). Methods of language assessment: A survey of California public school clinicians. *Language*, *Speech and Hearing Services in Schools*, *22*, 236–241.

World Health Organisation (1990). *International Classification of Diseases and Disorders* (10th ed.) Geneva: World Health Organisation.

Comparison of US and UK *CELF-Preschool subtests*

US subtest	UK subtest modifications
Linguistic concepts	Minor changes to vocabulary and syntax. e.g. 'turtle' changed to 'tortoise'. e.g. Item 17 'Point to either *one* of the monkeys and all of the tigers' modified to 'Point to either of the monkeys and all of the tigers'.
Recalling sentences in context	Narrative anglicised, and stimulus manual modified accordingly. e.g. 'baseball bat' to 'cricket bat', 'movers' to 'removal men', 'closet' to 'cupboard'.
Formulating labels	Target words and stimuli anglicised where necessary. e.g. 'sailboat' changed to 'sailing boat', 'parade' changed to 'band', US flag replaced with the Union Jack.
Basic concepts	No modifications.

Comparison of US and UK *CELF-Preschool subtests* (*cont*.)

US subtest	UK subtest modifications
Sentence structure	Minor changes to vocabulary and syntax. e.g. 'airplane' changed to 'aeroplane', 'baseball' changed to 'cricket'. e.g. Item 19 'The boy was followed by his cat' altered to 'The girl was followed by her cat'.
Word structure	Minor changes to vocabulary/examples of acceptable responses. e.g. 'bugs' changed to 'ladybirds/beetles'.

Comparison of US and UK *CELF* – 3 *subtests*

US subtest	UK subtest modifications
Sentence structure	Minor alterations to vocabulary and stimulus materials. e.g. 'line' changed to 'queue', 'cookies' changed to 'biscuits'. e.g. school buses modified to look more English.
Word structure	Minor alterations to acceptable responses and pictorial, stimuli. e.g. Item OD, acceptable response changed from 'does' to 'does/has' to allow for the frequently used English syntactic construction 'he has'. e.g. N28 American water hydrant removed from stimulus.
Concepts and directions	No modifications.
Formulated sentences	Stimulus materials anglicised where necessary. e.g. bank notes changed from dollars to pounds. e.g. baseball bat changed to tennis racket.
Word classes	Words anglicised where appropriate. e.g. 'state' changed to 'country', 'cookie' changed to 'biscuit'.
Recall sentences	Minor changes to vocabulary and syntax. e.g. Item 6 'The tall seventh grader made the field goal' modified to 'The tall year nine boy scored the goal'. e.g. Item 20 'Before the sophomores were dismissed for lunch, they were told to turn in their assignments' modified to 'Before the first years were dismissed for lunch, they told to hand in their assignments'.
Sentence assembly	Vocabulary anglicised. e.g. 'pickles' changed to 'biscuits', 'jar' changed to 'tin'.
Semantic relationships	Vocabulary anglicised. e.g. UK names substituted in place of US names: Kurt to Keith, Rosa to Rosie etc.
Word associations	Example responses modified where necessary. e.g. 'baby sitter' changed to 'child minder'.
Listening to paragraphs	Narratives anglicised. e.g. Ages 14.0 to 21.11: 'Dance Committee' changed to 'Disco Committee'.
Rapid automatic naming	No modifications.