The BSL Receptive Skills Test (WORKING VERSION)

This assessment measures comprehension of BSL grammar in children aged 3-13 years and enables professionals working with deaf children to make a baseline assessment, identify language difficulties, and evaluate the outcomes of therapy programs (Herman et al. 1998; Herman 1998). Children's test results can be compared on this test and the BSL Production Test (Herman et al. 2004).

The test has been standardised on 135 children and focuses on selected aspects of morphology and syntax in BSL. The assessment consists of the following: (1) a vocabulary check, (2) a video-based Receptive Skills Test.

(1) The vocabulary check is designed to ensure that children understand the vocabulary used in the Receptive Skills Test. Children complete a simple picture-naming task which identifies any signs in their lexicon which vary from those used in the test. Children are required to name pictures so that the assessor can check whether their version of the sign correspond to the one used in the test. This is particularly important for languages such as BSL where there is much regional variation of signs. After the vocabulary check, the tester can decide if he/she can administer either the northern UK or the southern UK version of the test.

(2) The Receptive Skills Test was originally presented on video. A DVD version will be available from 2009. The test consists of 40 items, organized in order of difficulty. There are two versions of this task; one for the north and one for the south of the UK. The items assess children’s receptive knowledge of BSL morphosyntax in the following areas: (1) negation, (2) number and distribution, (3) verb morphology, and (4) noun-verb distinction. The test procedure is explained by a deaf adult on the test DVD. The DVD has practice items to attune the child to the test format. Fade-outs between test items allow the child time to respond. Children respond by selecting the most appropriate picture from a choice of three or four in an accompanying colour picture booklet. The test takes about twelve minutes when not paused. Younger children sometimes need to take longer pauses between items, which extends the testing time to about 20 minutes. Scoring is on a pass/fail basis.

Materials: the pictures are easily recognizable and appealing to the age range of the subjects. Additionally, a range of distracter items are used to reduce the possibility of guessing, and the location of the target picture on the page is randomized.

The test is presented to the child on DVD, including test instructions, practice items, and individual test stimuli to guarantee a standardized presentation and reduce the demands on the tester. However, the vocabulary check is administered live and requires some BSL skills on the part of the tester.

The goal was to develop a norm-referenced test based on empirical data. Although a BSL test is needed for deaf children in deaf and hearing families, there are problems in standardizing assessments on the latter group because of the variability of language performance. So for standardization of the assessment, a more homogeneous group was used, including deaf children of deaf parents (DCDP), hearing children with native signing backgrounds, and selected deaf children from hearing families (DCHP) whose exposure to BSL was known (n=135).
Test development: Prior to the standardization of the BSL assessment, a pilot study was conducted. The criteria included for the subjects of the pilot was age and native signing background. All 41 children from the pilot were between 3;00 and 11;06 years old. This included 28 deaf children where only one parent was deaf, but BSL was used at home by all family members as well as 13 hearing children with a native signing background.

The children were assessed on the vocabulary checklist, the receptive task, and other BSL assessments in development. A deaf researcher with fluent BSL skills and a hearing researcher with good BSL skills administered the test.

The results of the deaf and hearing subjects indicated that they did not perform significantly differently on the receptive task, therefore both groups were included in the subsequent standardization phase. A preliminary analysis looked at the age effect by banding subjects into three broad age groups. Children who failed to attempt all test items the test were excluded from the analysis. This left 25 subjects divided into three groups: (1) eldest, age 9;06-11;06, n=8, (2) middle, age 6;00-9;05, n=10, and (3) youngest, age 3;00-5;11, n=7. A One-Way Analysis of Variance (ANOVA) revealed a highly significant relationship between age and raw score. The relationship was significantly different for age groups (1) and (3), (2) and (3), but not between (1) and (2). This was anticipated, since most grammatical development is completed by age eight. A statistical analysis of overall trends in the size of raw scores in relation to age revealed that the size of raw scores increases with age.

To investigate the difficulty of the items in the receptive task, an item facility analysis was used. Items that were either passed by all children (too easy) or failed by all children (too difficult) were excluded from the revised version. The items were ordered from easy to difficult based on the performance of the children in the pilot study.

The discrimination value was examined by looking at the subjects’ total test scores in comparison to their score for individual items. Items that failed to achieve a significant correlation of between 0.2 and 0.8 were discarded. The best 40 items of the original 68 were retained.

Inter-rater reliability was established by double-marking the performance of eleven children on all 68 items of the receptive task. The differences in the total raw score obtained by a different scorer was extremely low. Test reliability was investigated by retesting 10% of the subjects one month later. Test-retest reliability was found to be high.

Based on the findings of the pilot study, changes to the test and to administration were introduced. The DVD was amended to include practice items where feedback could be provided from the tester prior to starting the test proper. The option for the tester to present the practice items live was introduced, as this was felt to be especially important for the younger children. The revised version contained 40 items so that the test could be completed in one sitting. A few vocabulary terms were changed in light of children’s responses on the vocabulary check.

Some of the children from the pilot were included in the standardization study (their scores were analyzed separately). Additionally, selected deaf children from hearing families were included. There were 135 children included in the standardization procedure in total, located
in England, Scotland, and Northern Ireland. The age range was from 3;00 to 13;00. Children previously diagnosed as having additional handicaps were excluded from the sample.

Additionally, children completed two subtests of the Snijders-Oomen test of Non-Verbal Abilities to assess the non-verbal performance of children as a basis for inclusion in the project. Children below one standard deviation on this measure were excluded from the sample. Hearing children were assessed using a subtest from the Clinical Evaluation of Language Fundamentals (CELF) to measure their language development in English for subsequent comparison with their language level in BSL. All tests were administered by a deaf researcher and a hearing researcher with good BSL skills.

Six age groups were selected for the standardization of the receptive task. The age groups were: (1) age 3;00-3;11, n=10, (2) age 4;00-4;11, n=15, (3) age 5;00-5;11, n=18, (4) age 6;00-7;11, n=33, (5) age 8;00-9;11, n=32, and (6) age 10;00-13;00, n=27. It was considered important to maintain yearly intervals for the younger groups (age 3;00-5;11) since progress in language development is particularly marked in this age range. For the older children (6;00-13;00) two yearly intervals were chosen.

Table 1: Age groups and number of subjects for standardization of the receptive task

<table>
<thead>
<tr>
<th>Number of age group</th>
<th>Age range in each group</th>
<th>Number of subjects in each group (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>3;00-3;11</td>
<td>10</td>
</tr>
<tr>
<td>2.</td>
<td>4;00-4;11</td>
<td>15</td>
</tr>
<tr>
<td>3.</td>
<td>5;00-5;11</td>
<td>18</td>
</tr>
<tr>
<td>4.</td>
<td>6;00-7;11</td>
<td>33</td>
</tr>
<tr>
<td>5.</td>
<td>8;00-9;11</td>
<td>32</td>
</tr>
<tr>
<td>6.</td>
<td>10;00-13;00</td>
<td>27</td>
</tr>
</tbody>
</table>

The scores of the Receptive Skills Test were compared for children according to their exposure to BSL. Only in the youngest age groups did the children from deaf families perform better. For the older age groups, there was no significant difference between the native singers and deaf children from hearing families who were in bilingual programmes. There was no significant difference between children from deaf families and deaf children from hearing families with older deaf siblings.

In order to establish test-retest reliability of the receptive task, 10% of the children were retested. The test scores were better on the second testing, but the rank order of scores was
preserved. There was also a high correlation (.87) between the test and retest scores. Split-half reliability analysis for the internal consistency of the receptive test revealed a high correlation (.90), representing a high internal consistency. The scores of the children involved in the pilot were compared with those who had not previously been exposed to the test materials. There was a slight advantage in the pilot children, however the difference between the groups did not achieve statistical significance (p=0.7).

Data has more recently been collected from an unselected sample of deaf children who use BSL throughout the UK, for comparison with the standardisation sample (Herman & Roy 2006).

The BSL Receptive Skills Test is commercially available (Herman et al. 1999). Since the test is available to purchase, it can be assumed that the assessors do not need to be extensively trained. As the test is presented on DVD, test administration is standardized and scoring is on a simple pass/fail basis. It takes about 5 minutes to administer the vocabulary check. The Receptive Skills Test takes 20-30 minutes to administer. The time taken depends on the age of the child and the number of items fail. Discontinue rules apply after a number of consecutive fails. In addition, it may be necessary to pause the DVD to allow longer response time for younger children, whereas older children usually watch it straight through.

The BSL Receptive Skills Test has been adapted to LSF (French Sign Language; C. Courtin, personal communication, May 2, 2002), DSL (Danish Sign Language), LSI (Italian Sign Language) and Auslan (Australian Sign Language). Adaptations DGS (German Sign Language) and ASL (American Sign Language) are currently underway.

Among the strengths of the BSL assessment are that the tested items: (1) are based on empirical data, (2) have robust psychometric properties, (3) have standardized procedure and methods for conducting an assessment, (4) have a broad age range, 3-13 years old, (5) can be purchased, and (6) can be used in schools.

A weakness of the assessment is: (1) it assesses only certain linguistic structures on morphological and syntactic levels of BSL, not communicative competence.

A narrative skills test for BSL based on the same standardisation sample has also been published.

References:


