Published in Journal of Deaf Studies and Deaf Education

Adapting the Assessing British Sign Language Development: Receptive Skills Test into American Sign Language

Charlotte J. Enns, Rosalind C. Herman

The education of deaf students has included some significant changes and controversies over the past century. In particular, the use of signed languages in schools has been an issue. Although signed languages were regularly used to educate deaf students at the end of the 19th century (Lane, 1984), this changed to an emphasis on oral and auditory skills until about the 1970s (Marschark, Lang, & Albertini, 2002). Over the past few decades, signed languages have again become part of deaf education programs that incorporate a bilingual approach to teaching and learning (Knight & Swanwick, 2002; Mahshie, 1995; Prinz & Strong, 1996).

In order to monitor the success of bilingual deaf education programs, and in particular to monitor the progress of children acquiring signed language, it is essential to develop an assessment tool of signed language skills. Although researchers have developed some checklists and experimental tests related to American Sign Language (ASL) assessment, at this time a standardized measure of ASL does not exist (Singleton & Supalla, 2005). There have been tests developed in other signed languages, for example British Sign Language, that can serve as models in this area. The purpose of this study was to adapt the *Assessing British Sign Language Development: Receptive Skills Test* (Herman, Holmes & Woll, 1999) for use in ASL in order to begin the process of developing a standardized measure of ASL skills.

Background:

The key premise upon which all bilingual deaf education programs are based is establishing a first language foundation in a natural signed language. Bilingual programs emphasize first language acquisition in signed language because these languages are considered the most natural and accessible languages for deaf children (Johnson, Liddell & Erting, 1989; Klima & Bellugi, 1979). Without an established first language the entire program is brought into question. The primary objective of bilingual deaf education programs is to facilitate the normal acquisition of language, cognition, and social structures through an accessible first language and then build the skills of academic learning and literacy upon this foundation. Therefore, if deaf students enter school without an established language base, developing their signed language skills must be the focus of education before proceeding with other curricular areas. For this reason, the need for reliable and valid assessments of children's signed language proficiency are essential in furthering the educational objectives of bilingual programs for deaf students.

Determining children's level of signed language proficiency as they begin formal schooling is a major purpose of assessment. The need to monitor children's progress is another purpose of assessment. Deaf children who are having difficulty developing signed language skills are often identified by professionals through assessment. Therefore, identification of acquisition difficulties and strengths is yet another purpose of assessment. Assessment is also required for reporting purposes so that parents are aware of their child's level of functioning and rate of progress. Clearly, accurate assessment can serve a variety of purposes, and clearly identifies the need for signed language assessment tools. Unfortunately, in the area of signed language acquisition very few

commercially available assessment measures exist (Singleton & Supalla, 2005). As a result, teachers often rely on informal descriptive measures to develop teaching goals and monitor progress (Herman 1998).

There are legitimate reasons for the paucity of tests in the area of signed language acquisition. Identifying developmental problems in the acquisition of minority languages, whether signed or spoken, is challenging (Johnston, 2004). Norms for these populations often do not exist. There is a lack of controlled elicited data from representative samples of native users of various natural signed languages upon which norms for competency could be established (Schembri, Wigglesworth, Johnston, Leigh, Adam, & Barker, 2002). The number of studies of signing deaf children's language development is limited and in the studies that do exist, the number of subjects is small. This is because only a minority of deaf children (less than 10%, Mitchell and Karchmer, 2004) can be considered native signers, with a normal experience of language acquisition from exposure to deaf parents who sign. Despite these limitations, language researchers have defined some key developmental milestones and acquisition patterns in the signed language development of young deaf children (French, 1999; Lillo-Martin, 1999; Newport & Meier, 1985; Schick, 2003). Considerable information is also available regarding the linguistic features of ASL and their relative grammatical complexity and this can be used to develop guidelines regarding the sequence of acquisition (Neidle, Kegl, MacLaughlin, Bahan, & Lee, 2001; Valli & Lucas, 1992). Although the connection between ASL acquisition research and the development of practical assessment tools needs to be extended, there are several experimental measures that are worthy of review.

The American Sign Language Assessment Instrument (ASLAI) (Hoffmeister,

1994, 2000) consists of eight different measures, each developed to assess a particular ASL structure. It provides an in-depth investigation of both language production and comprehension. The ASLAI was used within the framework of a larger research project investigating the relationship of ASL as the first language and English (literacy) as the second language in deaf children. It is not currently available (Haug, 2005).

The Test of ASL (TASL) was also developed as part of a larger study examining the relationship between ASL and English literacy skills (Strong & Prinz, 1997; 2000). The TASL consists of two production measures (Classifier Production Task and Sign Narrative) and four comprehension measures (Story Comprehension, Classifier Comprehension Test, Time Marker Test, and Map Marker Test). The TASL has been pilot tested and used for research purposes on 155 deaf students aged 8 to 15 years, but it is not commercially available (Haug, 2005).

The ASL Proficiency Assessment (ASL-PI) (Maller, Singleton, Supalla, & Wix, 1999) is a screening tool developed to determine the level of ASL skills of non-native deaf children. It is based on studies of ASL acquisition that identified the development of eight morpho-syntactic structures, including 1) one- and two-sign utterances, 2) non-manual markers, 3) deictic pointing, 4) referential shifting, 5) verbs of motion, 6) aspects and number, 7) verb agreement, and 8) noun-verb pairs. The test procedure involves eliciting a videotaped language sample through interview with the examiner, peer interaction, and story retelling. The children are scored globally according to three levels of proficiency; Level 3 (16 targets or more produced), Level 2 (11 – 16 targets) and Level 1 (less than 11 targets). Initial testing involved 80 deaf children, aged 6 – 12 years, and although some psychometric testing for reliability and validity has been conducted, the

measure has not been standardized with large sample norms. As this test is still under development it is not available to the public (Haug, 2005).

The MacArthur Communicative Development Inventory for ASL (ASL-CDI) (Anderson & Reilly, 2002) is a measure of early vocabulary development in deaf children acquiring ASL. This assessment focuses on infants aged 8 – 36 months and is completed through parental or caregiver report. The assessment is based on the English version of the CDI (Fenson, Dale, Reznick, Thal, Bates, Hartnung, Pethick, & Reilly, 1993). Although the ASL-CDI has been shown to be reliable and valid, and it is commercially available, its scope is limited to assessing productive lexical development at the preschool level.

Each of the measures of ASL currently under development has strengths and weaknesses, particularly with regard to availability and purpose of assessment. Clearly the need for an assessment tool that can easily, reliably, and efficiently be administered and scored by teachers, as well as be used to monitor progress and provide guidelines for instruction is needed. This need lead to the exploration of tests developed in other signed languages that were standardized and commercially available. Namely, the *Assessing British Sign Language Development: Receptive Skills Test (BSL RST).*

The goal of the *Assessing British Sign Language Development: Receptive Skills Test (BSL RST)* (Herman, Holmes, & Woll, 1999) is to assess understanding of syntactic and morphological aspects of BSL in children aged 3-11 years. The test has both a vocabulary check and a video-based receptive test. The vocabulary check is a simple picture-naming task of 24 items, and is used to confirm knowledge of the test vocabulary and to identify any sign variations children may have that differ from those used in the test. If necessary,

the test administrator shows the child the test sign and ensures they can accept this version. If children do not know (unable to name or recognize) more than 5 of the vocabulary items, the test is discontinued at this point.

The receptive test includes 3 practice items, followed by 40 test items, organized in order of difficulty, and presented by video (tape or DVD format). Test items assess children's knowledge of BSL grammatical structures, including negation, number and distribution, verb morphology, and noun-verb distinction. Administering the test involves the child watching the video of a deaf adult explaining the test procedure and then presenting each test item. There are fade-outs between items that allow the child time to respond. The child responds by pointing to the appropriate picture represented by the signed item from a choice of three or four pictures in a colourful picture booklet. For children who require longer response time, the video can be paused for this purpose. Testing time varies from 12 minutes to 20 minutes, depending on children's response times. Scoring includes a quantitative raw score (number of items passed) that is converted to a standard score; as well as a qualitative error analysis to describe the pattern of errors made in relation to grammatical structures. The normative data is based on 138 children tested in England, Scotland and Northern Ireland. The sample included 76 girls, 62 boys; 20 hearing, 118 deaf; 78 from deaf signing families; 23 from established bilingual programs; and, 37 from Total Communication programs. Through the standardization it was determined that there was no difference in test performance between hearing and deaf children from deaf families, that there was no difference between children from deaf families and children in established bilingual programs, or from children in TC programs with deaf family members who sign. Children in TC

programs with no access to BSL outside school performed significantly below the other groups. Representation from all these groups is included within the standardization sample.

The *BSL RST* is the first standardized test of any signed language in the world that has been normed on a population and tested for reliability (Johnson, 2004). For this reason, researchers from several difference countries have chosen to adapt it into other signed languages. The advantage of adapting an existing test rather than developing an original test is that important considerations and decisions have already been evaluated. For example, the *BSL RST* is based on what is known about signed language acquisition and highlights grammatical features identified in the research as important indicators of proficiency, such as verb morphology and use of space (Herman, Holmes, & Woll, 1998). Considering that many signed languages share these important grammatical features it is likely that test items will be relevant in signed languages other than BSL.

Another important consideration is the composition of the standardization sample given the inconsistent exposure to signed language that occurs for most deaf children. Decisions regarding including hearing children with deaf parents, or deaf children with hearing parents attending bilingual or TC programs have already been made and substantiated with research evidence for the *BSL RST.* In addition, clear guidelines for the assessment format have also been validated. These decisions include: using pictures vs. toys to keep attention but not be distracting; ensuring familiarity with vocabulary through a pre-test; keeping items to an appropriate length to avoid excessive memory load; reducing fatigue effects due to length of total test items; and incorporating a video of target structures to standardize presentation and minimize influence by test administrator.

The advantages of adapting an existing test account for the fact that the *BSL RST* has currently been translated into five other signed languages, including French Sign Language, Danish Sign Language, Italian Sign Language, German Sign Language and Australian Sign Language (Haug & Mann, 2008). At this point, detailed information regarding the effectiveness of these translated tests is only available for Auslan, or Australian Sign Language (Johnston, 2004). It should be noted that the adaptation necessary for using the *BSL RST* in Auslan was minimal, in that only two signs included in the test (DOG, PENCIL) are signed differently from BSL. For this reason, the signed stimuli were re-videotaped but the BSL picture book was used and the Auslan scores were compared to the BSL norms. The findings suggested "inflated" scores for the Auslan users when compared to the BSL norms. One reason for this is related to the fact that one test item (WRITE-PENCIL) was not assessing the same construction (noun-verb distinction) in Auslan as it was in BSL; however, the item was kept in the test and it may have inflated some scores. Specifically, five students would have reached ceiling if they had not got this item right. Discussion of the adapted test results also considered factors such as age of exposure and length of exposure to Auslan, as well as hearing status. The Auslan sample included hearing children who were participants in a reverse sign bilingualism program and some of these hearing children scored better than the deaf students. This raised issues regarding the influence of overall language skills (were the hearing children benefiting more directly from the "mouthing" of English words in conjunction with the signing?) and non-linguistic factors, such as cognition and age. Clearly, further evaluation is needed regarding the equivalency of the BSL and Auslan versions of the test and whether one set of norms can be applied to both languages.

Overall, researchers concluded that early exposure is an important requirement in developing a first language, whether it is signed or spoken (Johnston, 2004).

The interest in adapting existing signed language tests into other languages was examined by Haug & Mann (2008). They begin their discussion by clarifying the distinction between "translation", defined as a one-to-one transfer without consideration of linguistic differences, and "adaptation", which involves developing a parallel test that "acknowledges the linguistic, cultural, and social conditions of those taking the adapted test while retaining the measurement of the constructs found in the original" (Oakland & Lane, 2004, p. 239). If the goal is to develop a test that closely resembles the existing test, but incorporates the specific needs of the target language, then adaptation is the appropriate term to use to describe the process.

The process of adapting tests from one signed language to another requires careful consideration of the linguistic differences that exist between the two languages; however limited cross-linguistic research related to signed languages can make this a challenging task (Mason, 2005). These challenges are illustrated by Haug & Mann (2008) through examples involving differences in the categorization of linguistic features (classifiers) between ASL and Swedish Sign Language; lexical differences (no distinction between "boy" and "child") in Italian Sign Language and BSL; and morpho-syntactic issues, such as more devices for negation (Italian Sign Language) or less variety of devices for negation (French Sign Language) compared to BSL. Cultural issues also play a part in test adaptation. This can be as simple as pictures depicting the size, colour, and shape of a British mailbox that is in contrast to a German mailbox; or as complex as a story

involving the experience of obtaining a driver's license, which is common in America but not in Switzerland (Haug & Mann, 2008).

The decision of whether it is advantageous to adapt an existing instrument that has already been tested and standardized must be considered within the framework of evaluating the linguistic and cultural differences between the original and target languages. The current study has worked within such a framework and therefore, provides valuable insights into the similarities and differences between assessing the receptive skills of children learning BSL and children learning ASL. Some of these differences were easily resolved through the modification of test stimuli, but others required more significant changes to the test. The study also reinforces the benefit of collaboration among researchers in advancing better understandings of natural signed language acquisition and measurement.

Method:

The process of adapting the *BSL RST* into ASL included the following phases:

1. Consultation with ASL linguists and adult native ASL signers to determine the following:
    a) the suitability of the test vocabulary, in particular, the presence of regional alternatives,
    b) the suitability of a direct translation of each existing BSL sentence into ASL,
    c) whether the existing distracter pictures were viable alternatives for ASL users,

d) the need to add test sentences to reflect ASL linguistic constructions that were not currently included and to replace current BSL structures not represented in ASL.

2. Development of new test items identified in 1d) above (note: this required developing more items than necessary in case not all proved to be equally effective).

3. Redrawing of any culturally inappropriate images (e.g., mail box, steering wheels on right side, etc.).

4. Recording of new test video in ASL, to include the test instructions in a child-friendly register.

5. Piloting of translated sentences and new sentences on a sample of typically developing native signers within the recommended age range 3 – 11 years (probably at least 40 children) to determine:

a) effectiveness of new items, and

b) developmental order of difficulty.

6. Item analysis on newly developed items to determine:

a) any items that are too easy (passed by all) or too hard (failed by all),

b) ability to discriminate based on age,

c) developmental order of difficulty.

7. Standardization on a larger sample to develop norms by age.

The current study describes the implementation of the first six phases of the test adaptation process, as the seventh phase, standardization on a larger sample, has not been completed at this time.

The first phase was completed by assembling a panel of ASL consultants consisting of two university researchers (both hearing; one in Education and the other in ASL Linguistics) and five teachers (all deaf and native-ASL users; two college instructors, two school teachers, and one early childhood educator). Over a series of sessions, the consultants reviewed each of the *BSL RST* test items and determined whether they were suitable for use in ASL. The results indicated that 28 of the 40 test items did not require changes – direct translation into ASL was possible and would appropriately assess the parallel ASL grammatical structure. For example, negation is assessed in item #3 of the *BSL RST* with the stimulus sentence "ICE-CREAM NOTHING", and response pictures of a boy with a single cone, a boy with a double cone, a girl with a cone, and a boy without any ice-cream. This item translated into ASL and using the same response pictures would effectively assess children's understanding of the negative term "NOTHING" in ASL. Similarly, item #10 of the *BSL RST*, "TWO-PEOPLE-MEET" (signed with index finger classifiers on each hand moving towards each other), was designed to measure spatial verb morphology, and achieved the same goal in ASL.

Of the twelve remaining test items that did require modification, eight items required changes to target sentences (and development of new items), and four items required changes only to the pictures. In addition, the panel recommended adding a fourth picture to the six *BSL RST* items that only had three picture response choices, in order to keep all test items consistent with four choices. Some of the test items that required modification were due to content. For example, an item in the *BSL RST* assessing role shift involved a boy hitting a girl, and this was modified to the more socially appropriate behaviour of the boy tapping (for attention) the girl instead. The item still measured children's

understanding of role shift in ASL. Similarly, the noun-verb distinction item in the *BSL RST* using the signs for "PENCIL" and "WRITE", was not appropriate in ASL as the signs are different, and was therefore replaced with "CHAIR" and "SIT". The modifications required to the pictures included changing the steering wheels to the left side of the vehicles and altering mailboxes and train logos.

In phase three, the re-drawing of test items, it was decided that all the test pictures should be re-drawn to ensure that the style of the pictures was consistent throughout the test. The drawings were digitized and printed into a book format. Phase four involved recording the ASL test sentences on video. This was completed in an appropriate studio space using high quality video recording equipment. The person signing the ASL sentences was a deaf teacher, with deaf parents, and a very fluent ASL user. She was also involved in the adaptation process, so was familiar with the test items and testing procedure. She was able to present the target sentences very clearly and in a child-friendly manner.

The first adapted version of the ASL Receptive Skills Test was a similar format to the *BSL RST*, in that the children watched video recorded ASL sentences and selected pictures to match. In the ASL version, the child always selected from a choice of 4 pictures, whereas in the *BSL RST*, six items only had a choice of 3 pictures. The adapted test included a vocabulary check of 20 items (2 less than the *BSL RST*), 3 practice items (same as the *BSL RST*), and 41 test items (one more than the *BSL RST*). The adapted ASL test also assessed the same grammatical categories as the *BSL RST*; number/distribution, negation, noun/verb distinctions, spatial verbs (location and action), size/shape specifiers, and handling classifiers.

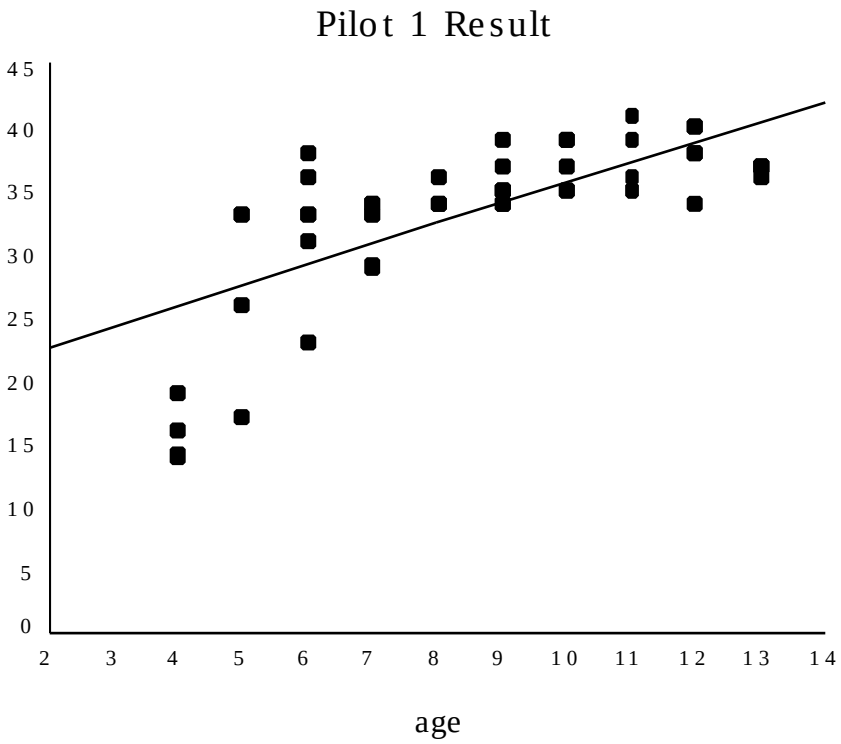
It was determined that the children participating in the pilot testing should be native ASL users, or more specifically that they be deaf and have deaf parents that have exposed them to ASL from birth. The purpose for limiting the pilot sample in this way was to ensure that the test reflected the appropriate developmental sequence of ASL acquisition based on fully accessible exposure to the language. Given the size of typical Canadian Deaf communities, it was not possible to recruit enough children in one location that met this criterion. For this reason, testing was conducted in Manitoba, Alberta, Ontario, and Minnesota. All testing was administered by a deaf research assistant (native ASL user) and videotaped to confirm the accuracy of scoring.

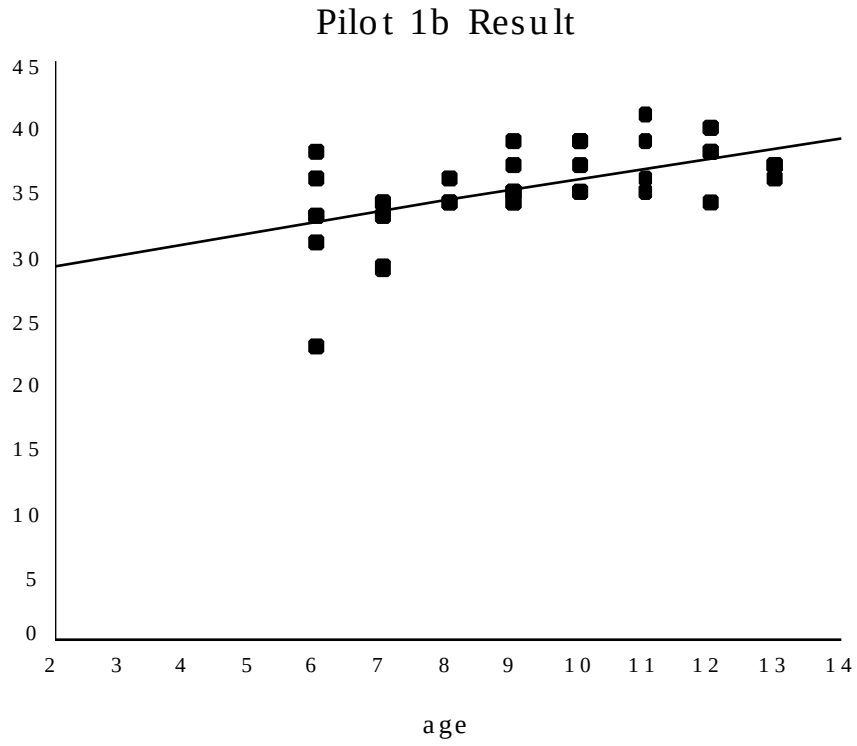
Two rounds of phase five (pilot testing) and phase six (item analysis) were necessary to finalize the adapted ASL version of the *BSL RST*. The results of the first round of pilot testing revealed a significant correlation between children's age and raw score; however, this correlation weakened as the children got older. Essentially, it was felt that the adapted ASL test was "too easy" and did not distinguish children's receptive skills beyond the age of 10 years. Further analysis of test items and consultation with the *BSL RST* author, indicated that modifications to stimulus sentences and distracter pictures were necessary in order to more effectively assess the intended ASL grammatical features. The specific procedures and results of the two rounds of pilot testing and item analysis are discussed in the next section.

Results:

The first round of pilot testing included 47 children, from both Canada and the United States, between the ages of 4 and 13 years. The test was easy to administer and took approximately 15 minutes or less for each student to complete. Similar to the

administration of the *BSL RST* no repetition of items was allowed except with the youngest children (4 years of age). The inter-scorer reliability was very high; all 47 tests were re-scored and only two differed by one point each. The relationship between age and raw score was compared using a Pearson correlation coefficient and revealed a strong correlation ($r = .711$, $p < .001$). The correlation between age and raw score was also analyzed with only the children 6 years of age and older ($n = 40$), and although this correlation was still significant ($r = .589$, $p < .001$) it was weaker. The results of the first round of pilot testing are displayed in Figures 1 and 1b.





Item analysis indicated that five test items were passed by all children; however, it should be noted that no data were collected with 3 year olds, and perhaps these items would have been more challenging for that age group. It was also noted that no items were failed by all children. These results contributed to the high scores for children of 8 years and older, and the limited ability of the test to distinguish ASL abilities of children over 10 years of age. Please refer to Table 1 for a summary of the first round of pilot test results.

Table 1: First Round Pilot Testing Results

Age	*# of Children*	*Mean Raw Score (max 41) and Range*
4 yrs	3	16.3 (14 – 19)
5 yrs	4	27.3 (17 – 33)

6 yrs	6	32.3 (23 – 36)
7 yrs	5	32.6 (29 – 34)
8 yrs	3	34.7 (34 – 36)
9 yrs	7	35.6 (34 – 39)
10 yrs	5	37.0 (35 – 39)
11 yrs	5	37.2 (35 – 41)
12 yrs	5	38.0 (34 – 40)
13 yrs	4	36.8 (36 – 37)

A more thorough analysis of error patterns and consultation with signed language test developers revealed ways to modify test items to make them more challenging and effective in assessing the target ASL structures. Revisions were needed for 23 of the original 41 test items, including changes to distracter drawings (11 items), signed stimulus sentences (4 items), and changes to both drawings and signed sentences (8 items). The revisions made to the distracter pictures ensured that the children needed to use their understanding of ASL grammar to answer the questions, rather than being able to guess from the pictures. It was necessary to replace some pictures with items that resembled the target ASL sign more closely than what was used in the *BSL RST*. For example, when the target was an open book lying on the bed (CL: B palm up), the *BSL RST* distracter item, a brush, was replaced with a shirt to more closely resemble the ASL CL: B handshape but with the palm down. Similarly, for the target two rows of beds (CL: N straight fingers), distracter pictures were changed from a bunk bed and a single bed to two rows of chairs (CL: N bent fingers) and two rows of pencils (CL: INDEX), to again

more closely resemble the ASL target structure. Revisions were also made to some of the signed stimulus sentences. These included shortening the hold or length of the action on noun-verb distinction items, for example “DRINKING” and “DRIVING”, and ensuring that the signers’ hands returned to a neutral position at the end of each item. In some cases it was necessary to revise both the pictures and the signed stimulus. For example, the item “EAT BIG-SANDWICH”, involved replacing pictures of eating an apple and eating chips, with eating a watermelon and eating a hamburger - foods that more closely resembled holding a sandwich. In addition, the handling classifier needed to be signed more clearly in the stimulus sentence. The item, “(NO)-REACH” was modified so that the sign was not held as long, and one of the distracter pictures was modified to increase the reach of a boy struggling to climb onto a chair, where previously his action was not as similar as the target item of the boy reaching for a teddy bear.

In addition to the modifications made to existing test items, four new items were added to assess understanding of role shift and conditional clauses. These constructions are considered to reflect more complex ASL grammar (Emmorey & Reilly, 1995) and the intention was to add items that the older children would find more challenging. The final modification to the adapted ASL test was to re-order the items to more accurately reflect the developmental level of difficulty according to the number of children that passed each item.

A second round of pilot testing was required for the revised version of the adapted ASL test. The revised version was a similar format to the initial adapted ASL test, in that the children were required to watch the video recorded ASL sentences and then select the appropriate picture from a choice of four. It also included a vocabulary check of 20 words

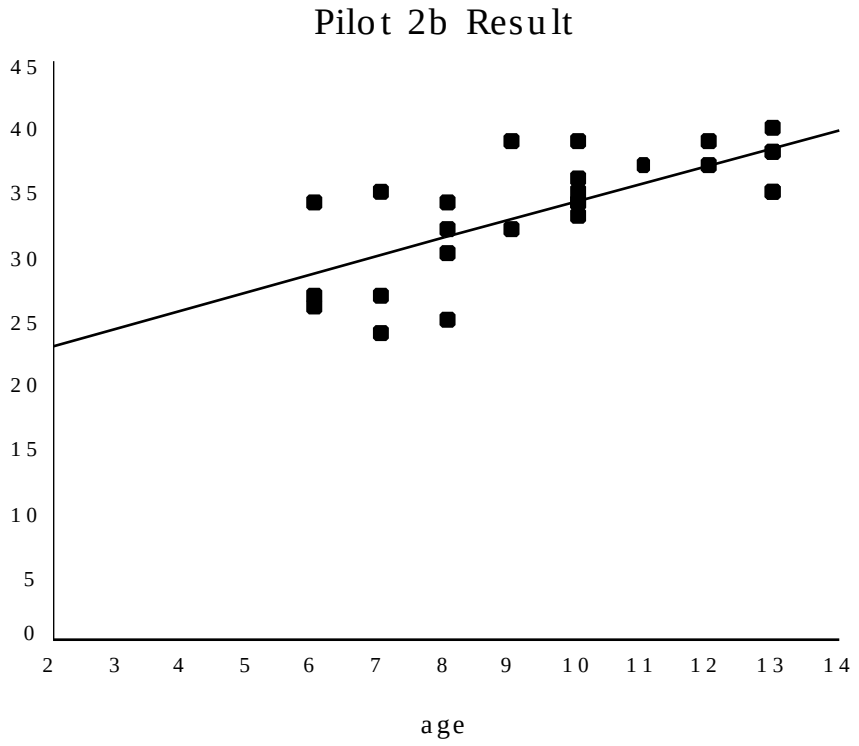
and three practice items, but the total test items were increased from 41 to 45 test items. The grammatical categories assessed through these items included the six previous structures (number/distribution, negation, noun/verb, spatial verbs, size/shape specifiers, handling classifiers), as well as the two additional categories of role shift and conditionals.

The second round of pilot testing included 34 children between the ages of 3 and 13 years. These children were again recruited from the same schools in Minnesota and Ontario, and therefore included both Canadian and American children. The results of the re-testing indicated that the modifications made to previous test items, and the new test items had effectively made the test more challenging and more clearly distinguished children's skills at different ages. Specifically, the analysis comparing age and raw score showed a strong correlation when all 34 children were included ($r = .821$, $p < .001$), as well as when only the children 6 years and older ($n = 20$) were included ($r = .719$, $p < .001$). Please refer to Table 2, and Figures 2 and 2b for more specific results of the second round of pilot testing.

Table 2: Second Round Pilot Testing Results

Age	*# of Children*	*Mean Raw Score (max 45) and Range*
3 yrs	2	3.0 (2 – 4)
4 yrs	3	7.7 (5 – 13)
5 yrs	3	26.0 (23 – 31)
6 yrs	3	29.0 (26 – 34)
7 yrs	3	28.7 (24 – 35)
8 yrs	4	31.5 (30 – 34)

9 yrs	2	35.5 (32 – 39)
10 yrs	6	35.6 (33 – 39)
11 yrs	1	37.0 (37)
12 yrs	2	38.0 (37 – 39)
13 yrs	5	37.2 (35 – 40)



A finalized version of the ASL Receptive Skills Test was developed based on the results from the second round of pilot testing. Modifications included deleting three test items, and re-ordering test items to more appropriately reflect the developmental sequence of language acquisition, as indicated by the number of children that correctly answered each test item. The deleted items were considered redundant (several other items were measuring the same grammatical structures) or in one case, culturally inappropriate (the item included an escalator and children from rural areas were often not familiar with this concept). In its finalized version the ASL Receptive Skills Test includes a vocabulary check of 20 words, three practice items, and a total of 42 test items. Eight grammatical categories are assessed through these items, including number/distribution,

negation, noun/verb distinction, spatial verbs (location and movement), size/shape specifiers, handling classifiers, role shift, and conditionals.

Conclusion:

The primary goal of this study was the same as any study involving test adaptation – "to have two tests that measure the same trait in fair, equitable, and somewhat equivalent fashion" (Oakland & Lane, 2004, p. 239). This goal has been accomplished. In addition, the work of developing effective measures of signed language acquisition contributes to understanding signed language grammatical systems and patterns of sociolinguistic variation. Inconsistencies and gaps continue to exist in the knowledge of signed language acquisition, even for well-documented languages like ASL, therefore descriptive investigation, such as in the current study, contributes to our confidence in designing valid and reliable measures of signed language grammatical proficiency (Schembri, et. al., 2002).

The finalized version of the ASL Receptive Skills Test now requires standardization. This will involve testing at least 20 children at each age level to get a standard score for what is expected at that age. When the test is used with other children they can be compared to the normal score for their age and it can be determined if they are developing an understanding of ASL age-appropriately. This research project of ASL test development begins to fill an enormous gap regarding the credibility of ASL as a language of instruction in schools. It also provides a contribution for teachers and researchers to help them deliver appropriate educational programming, monitoring and reporting.

References

Anderson, D. & Reilly, J. (2002). The MacArthur Communicative Development Inventory: Normative Data for American Sign Language. *Journal of Deaf Studies and Deaf Education, 7*, 2, 83-119.

Emmorey, K. & Reilly, J. (Eds.) (1995). *Language, gesture, and space*. Hillsdale, NJ: Lawrence Erlbaum.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual.* San Diego, CA: Singular Publishing Group.

French, M. (1999). *Starting with Assessment Toolkit.* Washington, DC: Gallaudet University Press.

Haug, T. (2005). Review of sign language assessment instruments. *Sign Language & Linguistics, 8,* 61-98.

Haug, T. & Mann, W. (2008). Adapting tests of sign language assessment for other sign languages – a review of linguistic, cultural, and psychometric problems. *Journal of Deaf Studies and Deaf Education, 13,* 1, 138-147.

Herman R (1998) The need for an assessment of deaf children's signing skills. *Deafness and Education.* 22:3–8.

Herman, R., Holmes, S., & Woll, B. (1999). *Assessing British Sign Language Development: Receptive Skills Test.* Gloucestershire, UK: Forest Bookshop.

Hoffmeister, R. J. (1994). Metalinguistic skills in deaf children: Knowledge of synonyms and antonyms in ASL. In J. Mann (ed.), *Proceedings of the Post Milan: ASL and English Literacy Conference* (pp. 151-175). Washington, DC: Gallaudet University Press.

Hoffmeister, R. J. (2000). A Piece of the puzzle: ASL and reading comprehension in deaf children. In C. Chamberlain, J. P. Morford & R. Mayberry (eds.), *Language Acquisition by Eye* (pp. 143-163). Mahwah, NJ: Lawrence Erlbaum Publishers.

Johnson, R. E., Liddell, S., & Erting, C. J. (1989). *Unlocking the curriculum: Principles for achieving access in deaf education* (working paper 89-3). Washington, DC: Gallaudet University, Gallaudet Research Institute.

Johnston, T. (2004). The assessment and schievement of proficiency in a native sign language within a sign bilingual program: The pilot Auslan receptive skills test. *Deafness and Education International, 6* (2), 57-81.

Klima, E. & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Knight, P. & Swanwick, R. (2002). *Working with deaf pupils: Sign bilingual policy into practice*. London, UK: David Fulton Publishers.

Lane, H. (1984). *When the mind hears: A history of the deaf*. New York: Random House.

Lillo-Martin, D. (1999). Modality effects and modularity in language acquisition: The acquisition of America Sign Language. In W. Ritchie & T. Bhatia (Eds.), *Handbook of child language acquisition.* San Diego, CA: Academic Press.

Mahshie, S.N. (1995). *Educating deaf children bilingually*. Washington, D.C.: Gallaudet University Press.

Maller, S. J., Singleton, J. L., Supalla, S. J., & Wix, T. (1999). The development and psychometric properties of the American Sign Language Proficiency Assessment (ASL-PA). *Journal of Deaf Studies and Deaf Education, 4*, 4, 249-269.

Marschark, M., Lang, H., & Albertini, J. (2002). *Educating deaf students: From research to practice.* New York: Oxford University Press.

Mason, T. C. (2005). Cross-cultural instrument translation: Assessment, translation, and statistical application. *American Annals of the Deaf, 151*, 356-370.

Mitchell, R. & Karchmer, M. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, 4, 138-163.

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (2001). *The Syntax of American Sign Language: Functional categories and hierarchical structure*. Cambridge, MA: MIT Press.

Newport, E. & Meier, R. (1985). The acquisition of American Sing Language. In D. I. Slobin (Ed.), The crosslinguistic study of language acquisition, Vol. 1. The Data (pp. 881-938). Hillsdale, NJ: Lawrence Erlbaum Associates.

Oakland, T. & Lane, H. (2004). Language, reading and readability formulas: Implication for developing and adapting tests. *International Journal of Testing, 4,* 239-252.

Schembri, A., Wigglesworth, G., Johnston, T., Leigh, G., Adam, R., & Barker, R. (2002). Issues in development of the Test Battery for Australian Sign Language Morphology and Syntax. *Journal of Deaf Studies and Deaf Education, 7,* 1, 18-40.

Schick, B. (2003). In Marshcark & Spencer (eds.) *Deaf Studies, Language, and Education*. New York: Oxford University Press.

Singleton, J. L., & Supalla, S. J. (2005). Assessing children's proficiency in natural signed languages. In M. Marschark & P. Spencer (Eds.), *The Oxford handbook of deaf studies, language, and education* (pp. 289-302). Oxford, UK: Oxford University Press.

Strong, M. (1988). *Language, learning, and deafness*. Cambridge, MA: Cambridge University Press.

Strong, M. & Prinz, P. (1997). A Study of the relationship between American Sign Language and English literacy. *Journal of Deaf Studies and Deaf Education, 2,* 1, 37-46.

Strong, M. & Prinz, P. (2000). Is American Sign Language skill related to English literacy? In C. Chamberlain, J. P. Morford & R. Mayberry (eds.), *Language Acquisition by Eye* (pp. 131-142). Mahwah, NJ: Lawrence Erlbaum Publishers.

Valli, C. & Lucas, C. (1992). *A resource text for ASL users*. Washington, DC: Gallaudet University Press.