Investigating Deaf Children's Vocabulary Knowledge in British Sign Language

Wolfgang Mann[1] & Chloë Marshall[1,2]

[1] City University London

[2] Institute of Education, University of London

Abstract

This study explores different aspects of the mapping between phonological form and meaning of signs in British Sign Language (BSL) by means of four tasks designed to measure meaning recognition, form recognition, form recall, and meaning recall. This is the first time that four such tasks have been explicitly put together to measure the degrees of strength of children's form-meaning mappings, at least for a signed language. The aim was to investigate whether there is a hierarchy of difficulty for these tasks, and therefore whether BSL vocabulary acquisition proceeds incrementally as is the case for spoken languages. Twenty-four deaf participants (aged 5-15 years), all of whom were BSL users, performed with greatest accuracy on meaning recognition, and least accurately on meaning recall. The results indicate that signers' knowledge of mapping between form and meaning in BSL signs is not an all-or-nothing phenomenon but depends on what the learner is required to do with that knowledge, as is the case for spoken languages.

Keywords British Sign Language; vocabulary; language testing; deafness

INTRODUCTION

Lexical items, be they spoken or signed, are mappings between a phonological form and a meaning or set of meanings. Children are able to create an initial mapping on the basis of just a single incidental encounter, or several encounters, with a new lexical item, and are able to retain these initial mappings (Carey, 1978; Goodman, McDonough, & Brown, 1998; Heibeck & Markman, 1987). Such fast mapping is an essential first step in a much longer process, namely the slow mapping of a phonological form to a more complete and conventionalised set of meanings, which requires exposure to the phonological form on different occasions and in different contexts (Clark, 2009).

The mapping between phonological form and meaning is fundamental to what it means to know a word: Given the phonological form, the user can access the meaning, and given the meaning, the language user can access the phonological form. Although there are other important aspects to vocabulary knowledge, for example, a word's morphological and syntactic properties, and its associations to other words, most standardised assessments of children's vocabulary draw on this mapping. Some, for example the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997), provide the child with a phonological form. The child then has to select the picture that matches its meaning, from a set of four. Other standardised assessments, such as the Expressive One Word Picture Vocabulary Test (EOWPVT; Brownell, 2000), provide a picture, and the child has to produce the phonological form that matches the meaning.

Researchers distinguish between mapping form to meaning, which is the first and central phase in vocabulary learning, and network building, which is a later and slower process (Aitchison, 1994; Henriksen, 1999). As children's vocabularies grow, words become organised into a semantic network, with strong links between words that are closely related and weaker links between those that are less closely related. When a word is processed, it activates other words linked to it in the semantic network. Dog for example, may activate words such as cat and hamster, which share a taxonomic relationship in that they fit within the same semantic category of "pet animal." Dog may also activate words such as kennel, bone, and leash, which are schematically or thematically linked to it. Such organisation of vocabulary knowledge can be probed by association tasks, whereby participants are asked to produce one or more words associated with a cue word (e.g., de Deyne & Storms, 2008; Sheng & McGregor, 2010), or to select, from a large set of words, those that are associated with a cue word (Read, 1993).

This paper presents an empirical study that investigated deaf children's vocabulary development in British Sign Language (BSL), and specifically tested the different degrees of strength of the mappings between form and meaning, and associations between lexical items. Before presenting the study, we discuss vocabulary acquisition in deaf children who are learning a sign language, focusing on similarities and differences compared to spoken language vocabulary, and on issues that are relevant to measuring sign language vocabulary.

VOCABULARY ACQUISITION IN SIGN LANGUAGES VERSUS SPOKEN LANGUAGES

While a great deal of research has investigated the complex nature of vocabulary acquisition and the assessment of vocabulary in spoken languages, much less is known about vocabulary development in signed languages. Vocabulary learning in deaf children is interesting for several reasons linked to the unique nature of language exposure in this group and the differences between lexical items in spoken and signed languages. Research on language development in deaf native signers[1] has shown that early exposure to sign enables children to reach developmental milestones at the same pace as their hearing peers (Anderson & Reilly, 2002; Lillo-Martin, 1994). However, only a small percentage of deaf children, 5-10%, receive sign language input from native or near-native signers (Mitchell & Karchmer, 2004), which leaves the remaining 90 - 95% of children with widely differing language backgrounds. Research on deaf children growing up with hearing parents suggests a slower pace of lexical acquisition and an overall smaller lexicon size (Anderson, 2006; Blamey, 2003; Lederberg & Spencer, 2009; Prezbindowski & Lederberg, 2003). This may be largely due to reduced incidental exposure to sign language: Hearing parents tend to use sign language only when directly addressing their deaf child and tend not to use it with hearing family members, so the child has few opportunities for picking up vocabulary through observing the interactions of others (Marschark, 1997).

There are several ways in which sign vocabulary acquisition is likely to differ from spoken vocabulary acquisition for all deaf children, and not just for those of hearing parents. For example, once hearing children learn to read, they are able to learn new words through print. In contrast, sign languages have no written form, so new signs cannot be learnt through literacy[2]. Furthermore, the structure of the lexicon in sign languages is different to that of spoken languages, due to contact between sign languages and the spoken languages around them. It has been proposed that sign languages have, in addition to a native lexicon, a non-native lexicon that comprises

[1] Native signers are defined as individuals who acquire a natural sign language, e.g., American Sign Language (ASL), BSL, from birth, and from their parents.

[2] Although literacy may be the way through which deaf children build the bulk of their spoken language lexicon.

fingerspelt[3] representations of spoken words. The native lexicon is divided in turn into the core lexicon, which comprises lexicalised signs (also called "frozen" or "established" signs, e.g., CAT, TREE, CAR), and a non-core lexicon, which comprises depicting and pointing signs that have a close relationship with gesture (Brentari & Padden, 2001; Johnston & Schembri, 2007). The distinction between these 3 components of the lexicon is not always clearcut, as over time depicting signs and fingerspelt/initialised items can become part of the core lexicon (e.g., fingerspelt e-g-g, and initialised m-m for mother and f-f for father in BSL). Even for core signs, contact with spoken language is evident, through the existence for some signs of mouthings that resemble the mouth patterns of the corresponding spoken word. For a number of signs these mouthings distinguish between lexical items that share the same manual component, and that would therefore otherwise be ambiguous.

One further way in which the acquisition of signed and spoken languages is likely to differ is in the greater iconicity of signed vocabulary. While many signs are arbitrary, with no obvious relationship between form (i.e., hand configuration, movement and location) and referent, a large number of signs have forms that do bear some visual resemblance to their referent (e.g., in BSL, SLEEP, EAT, PLANE). Recent studies have indicated that iconicity helps in the processing of signs, in both deaf adults (Thompson, Vinson, & Vigliocco, 2009) and in deaf children (Ormel, Hermans, Knoors, & Verhoeven, 2009), although, in contrast, a semantic priming study in deaf adults revealed no advantage for iconic signs (Bosworth & Emmorey, 2010), and a study of one deaf adult with sign language aphasia found that iconic and non-iconic signs were impaired to the same extent (Marshall, Atkinson, Smulovitch, Thacker, & Woll, 2004). Nevertheless, it is possible that iconicity plays a

[3] Fingerspelling is a manual alphabet used to spell out items such as people's names, place names, and technical items that do not yet have their own sign in the language, or whose sign the addressee may not be familiar with.

metalinguistic role in the learning of signs by allowing the child to map phonological features to properties of real world referents (see discussion in Perniss, Thompson, & Vigliocco, 2010).

What are the consequences of these differences for vocabulary development and the ensuing organisation of the lexicon? There are very few studies on this topic. In some ways, vocabulary acquisition in signed languages appears to be very similar to that of spoken languages. For example, Anderson and Reilly (2002) report that deaf children's acquisition of American Sign Language (ASL) vocabulary within particular semantic domains, such as question words, emotion signs, and cognitive verbs, is comparable to that found in hearing peers. Furthermore, Tweney and Hoemann (1973) found that deaf 9 to 13-year-olds, in a semantic association task with stimuli in ASL showed a clear shift from providing syntagmatic to providing paradigmatic associations[4]. This is the same type of shift, albeit delayed, that is found in hearing children.

Yet the influence of modality on semantics can be subtle. One of the features of signed languages is that their core lexicon is smaller than the lexicon of spoken languages[5] (Sutton-Spence & Woll, 1998), and there are few signs for items below the basic level. For example, in French Sign Language (LSF), as in BSL and ASL, there are signs for basic level items such as flower, dog, car etc, but not signs for particular types of flower, dog, or car (e.g., lavender, spaniel, Volvo) (Courtin, 1997). In practice, if a signer wants to discuss lavender, spaniels, or Volvos, he or she will have to use the generic sign for the basic level item first, then fingerspell the particular item. In a forced-choice association task with deaf 5 and 6 year-olds who used LSF,

---

[4] A syntagmatic association is one from a different grammatical class to the stimulus (e.g., dog-bark), whereas a paradigmatic association is from the same grammatical class (e.g., dog-cat).

[5] In contrast, the non-native (i.e., fingerspelling) and non-core (i.e., depicting and pointing signs) lexica, are extraordinarily productive.

Courtin (1997) found that deaf and hearing children associated vocabulary items below the basic level similarly when items did not share a common generic sign: Both groups produced more schematic than categorical associations (e.g., in LSF, CHARLOTTE CAKE is categorically related to APPLE and ICE CREAM – all are types of dessert – but schematically related to DESSERT SPOON and PLATE)[6]. In contrast, when items did share a common sign, deaf children were more likely to make categorical compared to schematic associations and made significantly more categorical associations than the hearing children (e.g., in LSF, the target CHARLOTTE CAKE shares a generic sign with STRAWBERRY TART and FRUIT CAKE, and was more likely to be chosen than the schematic DESSERT SPOON and PLATE). This is important because it suggests that children's semantic networks are influenced by the particular nature of form-meaning mappings.

MEASURING VOCABULARY KNOWLEDGE IN SIGN LANGUAGES

While the past decade has seen increased international activity in the field of sign language assessment, few studies have focused specifically on vocabulary. Some of these studies included tasks which required participants to demonstrate their understanding of target signs or produce signs for given meanings as part of a larger test battery to assess deaf children's expressive and/or receptive skills at different linguistic levels (Fehrmann et al., 1995a, 1995b, for German Sign Language; Hermans, Knoors, & Verhoeven, 2010, for Sign Language of the Netherlands; Hoffmeister, 1999, for ASL). Examples of such tasks include participants being asked to supply the sign representing the meaning of a target picture (equivalent to our form

[6] Categorical associations are those that are based on taxonomic relationship and similarity relations between class members. Schematic associations are those that are based on past experience with objects, scenes and events, and consist of a set of expectations about what things look like, or the order in which they occur (Mandler, 1977).

recall task, which we describe in due course), rate how well presented signs match the meaning of a target picture (i.e., form recognition), or match the correct synonym/antonym from a number of given response choices to a signed stimulus. Other studies have used vocabulary subtests as screening tools to determine the extent to which the language levels of the main assessment fall within the competence of the participants (Herman, Holmes, & Woll, 1999, for BSL; Kuntze, 2004, for ASL).

Currently, only one test developed specifically to assess deaf children's vocabulary has been published, namely the Prüfverfahren zur Erfassung Lexikalisch-Semantischer Kompetenz (PERLESKO; Bizer & Karl, 2002), a receptive vocabulary test for German Sign Language (DGS). The PERLESKO is normed for deaf children between 7-13 years, and can be used to assess individuals' comprehension skills in three language modalities, namely DGS, spoken German, and written German. With the exception of one empirical study (Wildemann, 2008), little information about this test has been published.

In the context of the present study, we hypothesise that those aspects of vocabulary that transcend modality are likely to be acquired in the same way in spoken and signed languages, whereas those aspects that are more closely related to modality are likely to show differences in their acquisition. Moreover, we focus on an aspect of vocabulary acquisition that is likely to be similar between signed and spoken languages: That vocabulary acquisition is incremental in nature and that some components of word knowledge are acquired before others. Researchers have endeavoured to capture the construction of lexical knowledge from a number of different angles, whether describing it as combination of a number of inter-related subknowledges or a continuum which consists of various levels of knowledge, ranging from superficial familiarity to the ability to successfully use the word in free production (for studies of spoken languages, see Laufer, Elder, Hill, & Congdon, 2004; Laufer & Goldstein, 2004; Read, 2000; inter alia).

One important distinction that has been made in spoken language research is between receptive and productive knowledge, that is, the direction of the mapping between form and meaning (Laufer & Goldstein, 2004). Receptive word knowledge enables the language learner to retrieve the meaning of the spoken or written word, whereas productive word knowledge empowers the learner to retrieve the form of the word. Generally, learners tend to comprehend new words before being able to produce them. Consequently, a learner who successfully retrieves the word form for a given meaning is likely to also retrieve its meaning when s/he is presented with the form. It is generally assumed that only a limited number of the words we know receptively will become available for production (see discussion in Henriksen, 1999). Receptive vocabularies are larger than productive vocabularies across a range of learners, and this includes young deaf children learning BSL (as measured by the BSL version of the MacArthur-Bates Communicative Development Inventories; Woolfe, Herman, Roy, & Woll, 2010). There are therefore signs that young learners of BSL understand when others use them, but that they do not yet use themselves.

A second important distinction has been made between the recall of words and their recognition, in other words, the depth to which the mapping between form and meaning is probed. Learners who can recall the meaning or form of a word are also likely to successfully recognize that meaning or form when presented with a number of options. If a word can be recognised but not recalled, that suggests a weaker mapping than if the word can be both recognised and recalled. However, the extent to which learners are able to use words receptively versus productively, and to recall versus recognise, is likely to also be related to how easily and automatically the mapping can be accessed. It should also be remembered that inferential strategies can work well to fill in the gaps in receptive knowledge, but are little help in production.

These two distinctions – receptive versus productive knowledge, and recall versus recognition – are relevant to vocabulary assessment. For instance, the PPVT and the EOWPVT, mentioned in the Introduction, differ from one another along exactly these dimensions. The PPVT tests the mapping in the direction of phonological form to meaning and assesses recognition of that meaning. The EOWPVT tests the mapping in the opposite direction, from meaning to phonological form, and assesses recall of that form. The advantage of both types of assessment is that they allow many different words to be tested, and hence are able to provide a measure of children's total vocabulary size. Their disadvantage is that they do not provide an indication of how robust the child's form-meaning mappings are.

Laufer and colleagues (Laufer, et al., 2004; Laufer & Goldstein, 2004) propose a new approach to vocabulary assessment in which the relationship between word form and word meaning can be expressed in four degrees of strength: active recognition, passive recognition, active recall, and passive recall. Aligning these terms with the terminology that we have used so far, "passive" indicates the mapping in the direction of form to meaning (i.e., requiring the participant to recognise/recall meaning), and "active" indicates the mapping in the direction of meaning to form (i.e., requiring recognition/recall of form). The four degrees of strength are hypothesised to form a hierarchy, with active (i.e., form) recall requiring the strongest degree of word knowledge and passive (i.e., meaning) recognition the weakest, with passive (meaning) recall and active (form) recognition in-between. While it is likely that the incremental nature of vocabulary acquisition discussed for spoken language also applies to signed languages, there is no empirical documentation to date to support such a claim.

THE PRESENT STUDY

The present study is the first of its kind to investigate deaf children's vocabulary knowledge in BSL by specifically measuring the degree of strength of the mappings between form and meaning for items in the core lexicon. We tested a group of children who were born deaf and who, on the report of teachers and parents, use BSL as their primary language. Four new tasks were created, each of which tapped a different degree of strength of vocabulary knowledge: meaning recognition, form recognition, meaning recall, and form recall. Following Laufer and colleagues (Laufer, et al., 2004; Laufer & Goldstein, 2004), each item was tested in all 4 tasks. To the best of our knowledge, this is the first time that four such tasks have been explicitly put together to measure the degrees of strength of children's form-meaning mappings, at least for a signed language.

Our approach was motivated by the following principal research question: Do signers show variation in their understanding of different mappings between form and meaning in BSL signs? Related to this is a secondary question: Do these mappings form the same hierarchy of the degrees of strength as is found in spoken languages? Put differently, are some degrees of knowledge more advanced than others and do they presuppose the less advanced degrees of knowledge? Three of the tasks were straightforward to design. As we describe in greater detail in the Methods section, for meaning recognition children were shown a BSL sign and had to select the meaning from a set of 4 pictures; for form recognition, children were shown a picture and had to select the appropriate form from a set of 4 signs (presented as short videos); for form recall, children saw a picture and were asked to produce the sign. The fourth task, meaning recall, was harder to design. Meaning recall has been tested in adult second language (L2) learners by asking them to complete a phrase, or a short sentence, in which the target word is embedded, for example, "When something melts, it turns into ……" (Laufer, et al., 2004, p. 207), or to provide the first language (L1) translation for a word supplied in the L2 (Laufer & Goldstein, 2004). Neither approach seemed appropriate for our purpose/target group, firstly because by using whole BSL phrases or sentences we might run the risk of participants not knowing all of the signs in those phrases/sentences, and secondly because we were testing L1 learners of BSL and did not want to bring English into the testing situation. An alternative would have been to show participants signs and ask them to define them, but an unrelated study of BSL vocabulary has shown that eliciting definitions proved difficult for children under 10 (Mason & Rowley, personal communication, December 1, 2010). This could be because providing definitions requires a high level of metalinguistic awareness in addition to good language skills at the sentence level, skills which would be expected to be challenging for young signers.

One solution was to test meaning recall by asking participants to produce a sign (e.g., BONE), related in meaning to the target sign (e.g., DOG) they saw. We hence tested meaning recall through association, which makes this a very different task to that used by Laufer and Goldstein (2004) and Laufer et al. (2004) for spoken language. As will be shown later (cf. Table 2), the meaning recall task additionally requires the recall of a form, and so might be predicted to be even more difficult than the straight form recall task, which would be the reverse to what Laufer et al. (2004) and Laufer and Goldstein (2004) found in their studies. In addition to enabling us to measure the degrees of strength of children's form-meaning mappings for the first time in a signed language, our four task battery allowed us to explore BSL vocabulary acquisition in one further way, namely, by asking a third research question: On the meaning recall task, do children show a shift from schematic (i.e., thematic) to categorical (i.e., taxonomic) responses as they get older?

METHODS

Participants

Twenty-four deaf participants (aged 5-15 years) were recruited from five programmes, namely three Deaf schools and two units/resource bases, in different regions of the UK. All programmes used BSL as a means of instruction. Their average age (12 males/ 12 females) was 11 years, 2 months. Table 1 shows individual participant characteristics. All participants had a hearing loss in the severe to profound range (>70dB) in their better ear. They were either native signers from deaf families or strong signers and all used BSL as their preferred means of communication. Testing was performed in a quiet room at each participant's school.

Table 1 – Insert here

Materials

Participants completed a biographical questionnaire, a nonverbal IQ (NVIQ) task, and the BSL vocabulary tasks. Each task is presented below.

Biographical questionnaire

Biographical information on each participant was collected by means of an online questionnaire. This questionnaire included questions on children's age, type of hearing loss, amplification used, parental hearing status, and communication used at home. It was filled out by the child's teacher or the speech language therapist.

Nonverbal IQ task

NVIQ was measured using the Raven's Progressive Matrices (Raven, 1994). A measure of non-verbal cognitive ability was needed in order to identify any participants with cognitive delays. The assessment took place during one of the two administration sessions of the vocabulary tasks.

BSL vocabulary tasks

Four tasks were developed to measure participants' mappings between form and meaning in BSL each of which consisted of the same 120 vocabulary items. This approach allowed the comparison of children's mappings across the four tasks and, thus, provide us with more detailed information about their level of knowledge for each item.

Our stimuli included pictures and BSL videos, depending on the task. To maintain a similar picture style, we used Clipart (www.clipart.com), an online collection of royalty-free images. The tasks were developed for children between the ages of 5-15 years in order to cover an as wide as possible range in response to the lack of BSL vocabulary assessments for deaf children at school age. The need to investigate a wide(r) age range is supported by a recent study by Wildemann (2008), who used the PERLESKO with younger children (5-6 years). Her findings suggest that the sign language subtest of the PERLESKO does not differentiate sufficiently between younger and older (7-13 years) children.

Two of the four tasks were receptive tasks and followed a multiple-choice (and hence, limited set) format In the meaning recognition task, participants were shown the target BSL sign, followed by 4 pictures, and had to select the picture that corresponded to the target sign. In the form recognition task, participants were shown a picture, followed by 4 BSL signs, and had to select the target BSL sign that matched the picture. The two remaining tasks were production (and hence, open set) tasks. In the form recall task, participants were shown a picture and had to produce the target BSL sign. For the meaning recall task, participants saw the target BSL sign and had to supply a different BSL sign with an associated meaning. Table 2 illustrates the four tasks, using the example of the BSL target sign BUS.

Table 2 – Insert here

The selection process of the vocabulary for the tasks was informed by a number of sources, including a BSL norming study (Vinson, Cormier, Denmark, Schembri, & Vigliocco, 2008), a receptive vocabulary test for German Sign Language (PERLESKO, Bizer & Karl, 2002), a number of standardized English vocabulary tests, and feedback from a group of deaf and hearing researchers and teachers of the Deaf [7]all of whom were fluent signers, who collaborated with the first author on the item development. In each of the four tasks, the 120 items were arranged in their order of difficulty, resulting in 12 sets of 10 items per age group (4-15 years). The order of difficulty was based on their position in any of the other assessments (where available) and on judgment by the deaf and hearing teachers.

Our aim was to tap vocabulary within the core lexicon, to the exclusion of fingerspelt signs and depicting/pointing signs (see our earlier explanation of the different components of the BSL lexicon). Within each set, the item order was randomized, resulting in a different order for each participant. The grammatical

[7] Following the conventions of the sign language literature, we use Deaf with an uppercase (D) to refer to members of the community that use BSL and deaf with a lowercase (d) when discussing the effects of hearing loss.

categories assessed were nouns, verbs, and adjectives[8]. The ratio of items across categories was 8:1:1 for participants (<10 years) and 6:2:2 for older participants (>10 years), based on well-documented findings from research on spoken language acquisition which show that children's first words are primarily nouns (Goldin-Meadow, Seligman, & Gelman, 1976; Macnamara, 1972; Nelson, 1973). While recent studies on early lexical development in sign language suggests a higher frequency of certain verbs, or action words/signs, such as CATCH, PLAY, SWIM, in children's signed BSL utterances over English (Woll, personal communication; for similar results in Sign Language of the Netherlands, see Hoiting, 2006), more research is needed with specific focus on deaf children's vocabulary development across grammatical categories at the age of 5 and beyond.

Out of the five participating schools, three were located in the South East region, one in the West Midlands, and one in the South West of the UK. To account for potential effects of regional variation in signing, we took the following steps. During the selection process of the items, any groups of signs that were known to be subject to regional variation were excluded. This included numbers and colours. In addition, all schools received a list with screenshots of each target sign prior to the testing visit. This allowed teachers to familiarize themselves with the signs and to anticipate any items that the student may not understand. Furthermore, in those schools that were located outside of the South Eastern region, a staff member (e.g., teaching assistant) stayed in the room during testing and could facilitate in the rare case when the child did not recognize a target sign or used a regional sign that the first author was not familiar with.

[8] We acknowledge that, in some instances, it is difficult to distinguish between nouns and verbs in sign languages, and there has been some debate regarding how obvious this distinction actually is.

Tables 3 and 4 below provide a schematic layout of the item design for the tasks.

Table 3 – Insert here

Table 4 – Insert here

Distractors for Receptive Tasks

The multiple-choice format used for the two receptive tasks (i.e., meaning recognition; form recognition) consisted of four types of responses: the target, a phonological distractor, a semantic distractor, and an unrelated distractor or, in the case of the meaning recognition task, a visual instead of an unrelated distractor wherever possible. These were presented in random order within a 2 x 2 arrangement. Selecting distractors in sign languages is not an easy task. In the PPVT, distractors are picked according to frequency in that all three distractors are chosen to match the target with respect to frequency. Frequency counts for BSL do not yet exist for the range of vocabulary items that we wanted to test. Given that our research focus is on the strength of the form-meaning mapping, we selected instead a form (i.e., phonological) distractor and a meaning (i.e., semantic) distractor. We also hypothesised that, for the meaning recognition task, if the child has no phonological or semantic knowledge of the target item, he or she might be tempted to pick an item where the sign is visually similar to the referent, in other words, to make a choice based on iconicity. The possibility of deaf children exploiting the iconic features of an unknown sign to correctly guess its meaning is not new and has been discussed in the literature, particularly in the context of test development (Herman et al., 2010; Jansma et al., 1997; White & Tischler, 1999). To account for this problem, we minimized the number of iconic items in the test by excluding items from certain categories used in vocabulary tests for spoken language, which are known for their high level of iconicity in sign language. These categories included animals, numbers, body parts, and vehicles. For any remaining items with iconic features, we selected visual distractors.

Phonological distractors were similar to the target item in that both shared one or more phonological parameter (e.g., hand configuration, location, movement). In the example for DOG, shown in Tables 3 and 4, the BSL sign SHOP is the phonological distractor because it shares movement and location with the target item, and only hand configuration differs. All signs selected as semantic distractors were semantically related to the target item. For example, CAT was selected as the semantic distractor for DOG. The two signs have different hand configurations, movements and locations. Finally, unrelated distractors, with no phonological or semantic relation with the target sign, were used for both tasks, except for the meaning recognition task if a visual distractor could be created. For example, “cutlery” was selected as a visual distractor for the target sign DOG in the meaning recognition task, because the two hands have a configuration and movement that resemble a knife and fork being held for eating (see Table 4, choice 3 for the form of the sign DOG).

Procedure

Each child was tested individually by the first author, who is hearing and a fluent signer with more than ten years experience in carrying out research with deaf children in schools. The test was presented to each child individually in a quiet room at the participant’s school, using a computer (either Mac or PC) with a 17-19 inch monitor and internet access. Both of the recognition tasks allowed participants to independently operate the mouse and either select an image corresponding to the stimulus sign (meaning recognition task) or a sign corresponding to the stimulus image (form recognition task). Upon selecting a response and clicking on the 'enter' button, responses were automatically saved onto an Excel datasheet on the web server. For the production tasks, the children had to produce a sign corresponding to the stimulus image (form recall task) or supply another sign with similar meaning to the stimulus sign (meaning recall task). BSL responses for these tasks were entered manually with English glosses by the first author during test administration and automatically saved to the database on the web server.

The four tasks were administered over the course of two sessions, with two tasks per session. To reduce fatigue, within each session one of the tasks was administered in the morning and the second task during the afternoon. In addition, to minimize learning, there was a minimum of one week between the two sessions. During the first session, participants completed the form recall and form recognition task. During the second session, participants completed the meaning recall task, followed by the meaning recognition task. Participants completed all 120 items in each of the four tasks and the administration time for each task was about 30 minutes. Prior to the beginning of each task, participants saw a video with pre-recorded instructions in BSL on their computer screen[9]. Next, they were given time to practice on two items during which they could ask questions. All participants completed the four tasks in the same order (i.e., form recall, form recognition, meaning recall, meaning recognition) spread over the morning and afternoon of the two sessions. This order/format was chosen to minimize, where possible, learning effects. As an additional measure, the item order within each item set was randomized. For the form

[9] Located below the video with the pre-recorded instructions in BSL, participants had the option to read the instructions in English, where preferred. Notwithstanding their language preference, all participants were shown the BSL instructions first.

recall task, prompts (e.g., what is this? what is s/he doing? how does s/he feel?) were used, where appropriate, to focus the participants' attention. To minimize fingerspelt responses or responses in phrase format, participants were encouraged during the practice to respond with a single sign.

The non-verbal IQ test was administered by the first author in a separate session. Children's biographical information was submitted by the teacher or speech language therapist around the time of the test administration, using the online questionnaire.

Scoring

The answers provided by participants for the recognition tasks were matched against a pre-specified marking key and automatically scored as correct ('1') or incorrect ('0'). These scores were saved in a database on the web-server.

For the two recall tasks, participants' responses were manually entered by the test administrator – the first author – in a scoring form, which appeared together with the stimulus on screen. The decision to score "live" was made to facilitate future administration of the test by practitioners (also see Hermans, Knoors, & Verhoeven, 2010). Four scoring choices were provided. The coding scheme consisted of "correct sign," "partially correct sign," "wrong sign/different sign" (i.e., incorrect), and "do not know" responses, which appeared as "CS," "PCS," "WS/DS," and "DNK" respectively, so that the participant did not know whether their response was correct or not, and would therefore not become less motivated. A response was coded as "correct" and scored as '1' if the participant demonstrated the ability to provide the expected BSL sign to name the target item (form recall task) or supply a semantically related sign (meaning recall task), and as "partially correct", scored as '0.5', if the participant supplied a sign that was outside the immediate range of expected answers, yet suggested that they knew the meaning of the target. An example of a partially correct response on the form recall task for the target BSL sign HAPPY was the sign EXCITED. On the meaning recall task, examples of partially correct responses for the target item GIRL were BEAUTIFUL and FASHIONABLE, while the responses BOY (i.e., the opposite of girl) and DRESS (something a girl would wear) were scored as correct.

In addition to scoring each response, the gloss for each sign was entered into a textbox below the score choices and all signs were videotaped. Inter-rater reliability on the judgement of responses was carried out, with a deaf native signer coding the responses of three of the children, in comparison with the first author's coding. Mean reliability was 94.5% (range 91-97%) for the form recall task, and 95.2% (range 93-97.5%) for the meaning recall task.

RESULTS

In order to get a developmental picture of vocabulary growth, we first tested whether performance on the different measures/tasks correlated with age, and correlated with raw score on the Raven's Progressive Matrices when age was partialled out (since non-verbal IQ has been shown to be weakly related to vocabulary size in spoken languages). We then investigated whether performance on each pair of vocabulary tasks was correlated once age and Raven's Progessive Matrices Raw Score had been partialled out. Scores for the vocabulary tasks and the Raven's Progressive Matrices are set out in Table 5.

Table 5 – Insert here

Bivariate correlations between each of the tasks and age, with the alpha level reduced to 0.013 to compensate for multiple comparisons (k=4) revealed significant correlations between age and meaning recognition, $R(24) = 0.766$, $p < 0.001$, age and form recognition, $R(24) = 0.758$, $p < 0.001$, age and form recall, $R(24) = 0.845$, $p < 0.001$, and age and meaning recall, $R(24) = 0.511$, $p = 0.011$. However, raw scores on the Raven's Progressive Matrices, with age partialled out and the alpha level again set at 0.013, did not correlate significantly with any tasks, although there was a trend for such a relationship for the two recognition tasks: meaning recognition, $R(21) = 0.451$, $p = 0.031$, form recognition, $R(21) = 0.465$, $p = 0.025$, form recall, $R(21) = 0.392$, $p = 0.064$, and meaning recall , $R(21) = 0.305$, $p = 0.157$.

Partial correlations between the different tasks, controlling for age and Raven's raw scores, with the alpha level reduced to 0.008 (for k=6 comparisons), were significant for: the two recognition tasks, $R(20) = 0.818$, $p < 0.001$; form recognition and form recall, $R(20) = 0.645$, $p = 0.001$; form recognition and meaning recall, $R(20) = 0.696$, $p < 0.001$; and the two recall tasks, $R(20) = 0.626$, $p = 0.002$. The partial correlations were nearing significance for meaning recognition and form recall, $R(20) = 0.489$, $p = 0.021$; and meaning recognition and meaning recall, $R(20) = 0.520$, $p = 0.013$.

The results of the correlational analyses indicate that knowledge of form-meaning mappings improves with age but is only weakly related to non-verbal abilities. This was important to ascertain because it reveals that BSL vocabulary level is much more than just part and parcel of general intelligence.

We next investigated our first research question, namely whether there is a hierarchy of difficulty for the different tasks, indicating that the tasks tap the strength of the form-meaning mapping to differing extents. We used a series of paired samples t-tests to compare performance across the 4 tasks, and with the alpha level reduced to 0.008 to compensate for multiple (k=6) comparisons. Performance on the form recall task was not significantly different to performance on the meaning recall task under this strict alpha level, $t(23) = 2.640$, $p = 0.015$, but all other comparisons were significant: Children scored higher on meaning recognition than form recognition, $t(23) = 3.686$, $p = 0.001$; higher on meaning recognition than form recall, $t(23) = 17.024$, $p < 0.001$; higher on meaning recognition than meaning recall, $t(23) = 8.767$, $p < 0.001$; higher on form recognition than form recall, $t(23) = 16.887$, $p < 0.001$, and higher on form recognition than meaning recall, $t(23) = 8.288$, $p < 0.001$. The results indicate that meaning recognition is the easiest task, followed by form recognition, with form recall and meaning recall being the hardest[10]. These results speak to our second research question, namely whether these mappings form the same hierarchy of the degrees of strength as is found in spoken languages. In spoken languages the pattern is identical, except for form recall being harder than meaning recall, whereas for our BSL tasks they are not significantly different. However, as we discussed in the introduction, our meaning recall task was by necessity rather different to that used in previous studies of meaning recall.

In order to address our third question, that is, whether children show a shift from schematic (i.e., thematic) to categorical (i.e., taxonomic) responses as they get older, we further analyzed the types of associations that children make for the meaning recall task. Four children[11] gave a high proportion of "do not know"

---

[10] As an anonymous reviewer points out, the fact that in the recognition tasks the child could get the answer right just by chance, whereas this is not likely for the two recall tasks, means that the magnitude of difference between the two recognition tasks and the to recall tasks is probably smaller than our data suggest. Nevertheless, given the overall very high performance on the two recognition tests and the large t values in the paired comparisons, we are confident that this difference between recognition and recall scores is due to more than just the differences in task format.

[11] These children were the three youngest children, aged 5;1, 6;5 and 6;6, and also an older child, aged 13;3.

responses, and testing was discontinued, so their data for this task are not complete. Therefore only the data from 20 children are included in this analysis.

Because we coded for categorical versus schematic responses, we excluded verbs and adjectives from the item list (k =20 for each). This left us with 80 nouns. Of these 80, we selected only those nouns for which there were 3 or fewer “do not know” responses. Correct responses to the remaining 46 nouns were coded as follows:

- Categorical, e.g., dog – cat; plane – helicopter; dress – jumper. This class also included superordinate associations, e.g., cherry – fruit; chocolate – food.
- Schematic, e.g., dog – bone; dress – lady; plane - flying; fish – sea. Accepted responses were associated in some way with the schema for that object. (Schematic associations are based on past experience with objects, scenes, and events, and consist of a set of expectations about what things look like, or the order in which they occur. These could be things the object does, how it is used, what it looks like etc.)

Errors were of three types:

- “Don’t know” responses
- Responses that have no semantic association with the target but that are phonologically similar, e.g., subtitles – ruler, cherries – earrings.
- Responses that have no semantic or phonological association with the target

Items were coded by the second author, and 11 items were also coded across all 20 children by the first author. Inter-rater reliability was 97.0%. The different error types, along with categorical and schematic responses, are set out in Table 6 for the sake of completeness, but only the two types of semantic association are analysed further.

Table 6 - Insert here

A paired samples t-test revealed that, as expected from the difference in their means (30% for categorical responses and 54% for schematic responses), children were significantly more likely to give a schematic than a categorical response, $t(19) = 3.241$, $p = 0.004$. We investigated whether the type of response changed with age, by splitting the children into 2 groups: the youngest 10 children (range 8;0 – 11;1) and the oldest 10 (range: 11;3 – 16;5). There was no difference between the groups. The younger group produced 28.91% (SD=13.40) categorical and 55.43% (SD=16.60) schematic responses, and the older group produced 31.30% (SD=23.66) categorical and 53.26% (SD=16.08) schematic responses.

It is of course possible that the number of categorical and semantic associations that a child makes varies more closely as a function of vocabulary ability or non-verbal IQ than as a function of age. However, splitting the participants into 2 groups based on overall vocabulary score (i.e., a composite of the 4 vocabulary tasks) yielded no difference between the low scoring and high scoring groups: The lowest scoring 10 children produced a mean of 27.17% (SD=18.57) categorical responses and 51.09% (SD=16.21) schematic responses, while the highest scoring 10 children produced a mean of 33.04% (SD=19.47) categorical and 57.61% (SD=15.82) schematic responses. The same was the case with respect to non-verbal IQ: The lowest scoring group produced 28.91% (SD=18.36) categorical responses and 54.78% (SD=18.55) schematic responses, and the highest scoring 10 children produced 31.30% (SD=20.06) categorical and 53.90% (SD=13.86) schematic responses. Therefore, there is no evidence in these data for shifts from schematic to categorical responses as children get older, or as their vocabulary or non-verbal IQ improve.

DISCUSSION

A sign is a mapping between a phonological form and a meaning. In this study, four tasks were devised to measure different aspects of this mapping BSL: meaning recognition, form recognition, form recall, and meaning recall. In addition, the meaning recall task tapped into the associative links between signs. The main aim was to investigate whether there is a hierarchy of difficulty for these tasks, and therefore whether BSL vocabulary acquisition proceeds incrementally, as is the case for spoken languages.

We tested 24 deaf signers between the ages of 5 and 15. Children's performance on all four tasks was correlated with chronological age but correlated only weakly with non-verbal IQ scores on the *Raven's Progressive Matrices* once age had been factored out. The findings indicate that knowledge of form-meaning mappings improves with age but only weakly with increasing non-verbal ability. Performance across most of the four vocabulary tasks was correlated, even with age and *Raven's* scores partialled out, but after correcting for multiple comparisons there was only a trend for a significant correlation between meaning recognition and form recall, and between meaning recognition and meaning recall. Nevertheless, these results indicate that all tasks tap into the child's knowledge of vocabulary.

Furthermore, the tasks form a hierarchy of difficulty, with the meaning recognition task being the easiest, followed by the form recognition task. The two recall tasks – form recall and meaning recall – were the hardest. Meaning recall produced the lowest scores, although not significantly lower than for the form recall task at the strict alpha level adopted in our analyses. The prediction had been that this particular meaning recall task would be harder than form recall because of its additional requirement for recall of a phonological form. However, the result might also suggest that deaf children do not have as large a network of links between items in their BSL core lexicon, which might be related to the smaller vocabulary size of signed languages compared to spoken languages (Sutton-Spence & Woll, 1998). This would make an interesting topic for future research.

Overall, our results indicate that the signer's knowledge of the mapping between form and meaning is not an all-or-nothing phenomenon, but depends on what the learner is required to do with that knowledge. For a signed language, just as for spoken languages, it appears that recall requires a stronger mapping than recognition, and signers who can recall the form of a sign or supply another sign with an associated meaning are also likely to recognise that sign's form or meaning. Vocabulary acquisition in BSL therefore proceeds incrementally, as is the case with spoken languages, with the strength of the mapping between form and meaning increasing over time.

Further analysis of a subset of the data allowed us to investigate those form-meaning mappings in more detail. For the meaning recall task, we had expected to find a shift from schematic to categorical responses with age, but this was not the case. It may be that this shift would have been evident if we'd been able to test participants over an even wider range, at a younger age (the youngest children we were able to include in this analysis were 8 years old) and into early adulthood.

The methodology used in this study could potentially form part of a larger assessment battery for measuring deaf children's BSL skills, and be used by teachers for program placement. This is of particular relevance in a sign bilingual education environment, where special consideration needs to be given to the question of how deaf pupils' sign language can be properly assessed. By measuring different degrees of strength of BSL vocabulary, children's understanding of different form-meaning mappings can inform practitioners in more detail about a student's level of knowledge for any sign, thereby facilitating the decision about appropriate ways to effectively support student learning efforts. Frequently, vocabulary instruction is limited to direct instruction of signs, which is often confined to memorizing definitions. A more meaningful approach could see the combination of repeated exposure to signs and their definitions with contextual information, thereby enabling the child to explore the meaning of the new signs rather than simply memorizing them.

While one of the main strengths of using of the same items across tasks is that it provides a more in-depth look at possible differences in deaf children's knowledge of 120 individual signs, this approach is not without potential weakness(es). One such weakness may be the period of time between administering different sets of tasks, which is significantly shorter in comparison to most conventional language tests (e.g., six months or more between test and re-test). Although our situation cannot be directly compared because the format of each of our tasks is different, it does not completely exclude/remove the possibility of the test taker having acquired his or her knowledge while completing the fours tasks. In our study, we tried to minimize this learning effect as much as possible by randomizing the item order within each task to make it more difficult for the test taker to memorize items. This problem is likely to be further reduced in future use by practitioners where we envision that a teacher will not necessarily administer all four tasks in order to assess a student's level of vocabulary knowledge, but may do so only for certain signs of interest, to really probe particular item knowledge.

For the present study, we intentionally selected deaf children who were (known to be) good signers (with frequent access to sign language through interaction with family members or friends). However, given that the majority of deaf children come from widely varying language backgrounds, their acquisitional pathways may be quite different. In order to further explore these differences, additional research is needed to investigate knowledge of form-meaning mappings in BSL on a larger and more diverse sample, by including deaf children with different levels of signing skills and less frequent exposure to sign language, and also by tracking their lexical acquisition longitudinally. We are currently collecting data from a larger sample of deaf children with different levels of signing from schools for the Deaf and support units throughout the UK. Finally, a valuable aim for future investigations using the tasks we have developed would be to investigate which strength-associated mapping is best associated with wider sign language skills. This could inform both teachers and learners in their decision of which aspect of the form-meaning mapping to concentrate on.

CONCLUSION

A theory or model of language development has to account for all languages, whether spoken or signed. In this context, particular focus needs to be given to deaf users of signed languages, given the unique situation of this group and the varying circumstances under which they acquire language. Taking these variables into account, we introduced a new approach to investigating deaf children's vocabulary knowledge in BSL by specifically measuring the degree of strength of the mappings between form and meaning for items in the core lexicon, and associations between signs. The data we presented in this paper show that the idea of degrees of strength of the mapping between form and meaning to measure language users' vocabulary knowledge appears to hold for signed languages as well as spoken. This is important for at least two main reasons. Firstly, it facilitates the comparison of vocabulary development between deaf and hearing language learners. Secondly, given the overall lack of suitable assessment tests for sign language and the over-reliance of teachers of the Deaf on tests that have been developed and normed on hearing language users, our methodology shows that existing models of language development can be successfully adapted across modality in a way that is appropriate for deaf test-takers.

Revised version accepted 30 July 2011

REFERENCES

Aitchison, J. (1994). Words in the mind: An introduction to the mental lexicon (2nd edition). Oxford: Blackwell.

Anderson, D. (2006). Lexical development of deaf children acquiring signed languages. In B. Schick, M. Marschark, & P. E. Spencer (Eds.), Advances in the sign language development of deaf children (pp. 135-160). Oxford University Press.

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

Bizer, S. & Karl, A.-K. (2002). Entwicklung eines Wortschatztests für gehörlose Kinder im Grundschulalter in Gebärden-, Schrift- und Lautsprache. Dissertation, Fachbereich Erziehungswissenschaften, Universität Hamburg.

Bizer, S. & Karl, A.-K. (2002). Perlesko – Prüfverfahren zur Erfassung lexikalisch-semantischer Kompetenz gehörloser Kinder im Grundschulalter. Manual. Dissertation, Fachbereich Erziehungswissenschaften, Universität Hamburg.

Blamey, P. J., (2003). Development of spoken language by deaf children. In M. Marschark & P. E. Spencer (Eds.), Oxford Handbook of Deaf Studies, Language, and Education (pp. 232-246). Oxford University Press.

Bosworth, R. & Emmorey, K. (2010). Effects of iconicity and semantic relatedness on lexical access in American Sign Language. Journal of Experimental Psychology: Learning, Memory and Cognition, 36, 1573-1581.

Brentari, D. & Padden, C. A. (2001). Native and foreign vocabulary in American Sign Language: A lexicon with multiple origins. In D. Brentari (Ed.), Foreign vocabulary: A cross-linguistic investigation of word formation (pp. 87-119). Mahwah, NJ: Lawrence Erlbaum Associates.

Brownell, R. (2000). Expressive One-Word Picture Vocabulary Test. Novato, CA: Academic Therapy Publications.

Carey, S. (1978). The child as word learner. In J. Bresnan, G. Miller, & M. Halle (Eds.), Linguistic Theory and Psychological Reality (pp. 264-293). Cambridge, MA: MIT Press.

Clark, E. (2009). First Language Acquisition. 2nd edition. Cambridge, UK: Cambridge University Press.

Courtin, C. (1997). Does sign language provide deaf children with an abstraction advantage? Evidence from a categorisation task. Journal of Deaf Studies and Deaf Education, 2, 161-171.

De Deyne, S. & Storms, G. (2008). Word associations: Norms for 1,424 Dutch words in a continuous task. Behavior Research Methods, 40, 198-205.

Dunn, L. M. & Dunn, L. M. (1997). Peabody Picture Vocabulary Test. 3rd Edition. Circle Pines, MN: American Guidance Services.

Fehrmann, G., Huber, W., Jäger, L., Sieprath, H., & Werth, I. (1995a). Linguistische Konzeption des Aachener Tests zur Basiskompetenz in Deutscher Gebärdensprache (ATG). Unpublished Report, RWTH-Aachen, Germanistisches Institut & Neurologische Klinik, Projekt DESIRE.

Fehrmann, G., Huber, W., Jäger, L., Sieprath, H., & Werth, I. (1995b). Aufbau des Aachener Tests zur Basiskompetenz in Deutscher Gebärdensprache (ATG). Unpublished Report, RWTH-Aachen, Germanistisches Institut & Neurologische Klinik, Projekt DESIRE.

Fenson, L., Dale, P.S., Reznick, J.S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J.S. (1993). Guide and technical manual for the MacArthur Communicative Development Inventories. San Diego, CA: Singular Press.

Goldin-Meadow, S., Seligman, M., & Gelman, R. (1976). Language in the two-year-old. Cognition, 4, 189-202.

Goodman, J.C., McDonough, L., & Brown, N.B. (1998). The role of semantic context and memory in the acquisition of novel nouns. Child Development, 69, 1330-1344.

Haug, T. (2009). Review of sign language assessment instruments. In A. Baker & B. Woll (Eds.), Sign Language Acquisition (pp. 51–85). Amsterdam: John Benjamins.

Heibeck, T.H. & Markman, E.M. (1987). Word-learning in children: An examination of fast mapping. Child Development, 58, 1021-1034.

Henriksen, B. (1999) Three dimensions of vocabulary development. Studies in Second Language Acquisition, 21, 303-317.

Herman, R., Holmes, S., & Woll, B. (1999). Assessing BSL Development - Receptive Skills Test. Coleford, UK: The Forest Bookshop.

Hermans, D., Knoors, H., & Verhoeven, L. (2010) Assessment of Sign Language Development: The case of Deaf children in the Netherlands. Journal of Deaf Studies and Deaf Education, 15, 107-119.

Hoffmeister, R. J. (1999). American Sign Language Assessment Instrument (ASLAI). Manuscript, Center for the Study of Communication & the Deaf, Boston University.

Hoiting, N. (2006). Deaf children are verb attenders: Early sign vocabulary development in Dutch toddlers. In B. Schick, M. Marschark, & P. E. Spencer (Eds.), Advances in the Sign Language Development of Deaf Children (pp 161-188). Oxford: Oxford University Press.

Jansma, S., Knoors, H., & Baker, A.E. (1997). Sign language assessment: A Dutch project. Deafness & Education, 21, 39-46.

Johnston, T. & Schembri, A. (2007). Australian Sign Language: An introduction to sign language linguistics. Cambridge: Cambridge University Press.

Kuntze, M. (2004). Literacy acquisition and deaf children. A study of the interaction of ASL and written English. Unpublished dissertation. Stanford University.

Kyle, J. & Woll, B. (1983). Language in sign. London: Croom Helm.

Laufer, B. & Goldstein, Z. (2004). Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness. Language Learning, 54, 469-523

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? Language Testing, 21, 202-226.

Lederberg, A. R. & Spencer, P.E. (2009). Word-learning abilities in deaf and hard-of-hearing preschoolers: effect of lexicon size and language modality, Journal of Deaf Studies and Deaf Education, 14(1), 44-62.

Lillo-Martin, D. (1994). Setting the null argument parameters: Evidence from American Sign Language and other languages. In B. Lust, M. Suñer, & G. Hermon (Eds.), Syntactic Theory and First Language Acquisition: Cross Linguistic Perspectives, Volume 2: Anaphora (pp. 301-318). Hillsdale, NJ: Lawrence Erlbaum Associates.

Macnamara, J. (1972). Cognitive basis of language learning in infants. Psychological Review, 79, 1-13.

Mandler, J. M. (1977). Categorical and schematic organization in memory. In C. R. Puff (Ed.), Memory organization and structure (pp. 259-299). New York: Academic Press.

Marschark, M. (1997). Raising and educating a deaf child. New York: Oxford University Press.

Marshall, J., Atkinson, J., Smulovitch, E., Thacker, A., & Woll, B. (2004). Aphasia in a user of British Sign Language: Dissociation between sign and gesture. Cognitive Neuropsychology, 21, 537-554.

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.

Nelson, K. (1973). Structure and strategy in learning to talk. Monographs of the Society for Research in Child Development, 38, 149.

Ormel, E., Hermans, D., Knoors, H., & Verhoeven, L. (2009). The role for sign phonology and sign iconicity during sign processing: The case of deaf children. Journal of Deaf Studies and Deaf Education, 14, 436-448.

Perniss, P. M., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages [Review article]. Frontiers in Psychology, 1, E227. doi:10.3389/fpsyg.2010.00227.

Prezbindowski, A.K. & Lederberg, A.R. (2003). Vocabulary assessment of deaf and hard-of-hearing children from infancy through the preschool years. Journal of Deaf Studies and Deaf Education, 8 (4), 383- 400.

Raven, J.C. (1994). Standard progressive matrices sets A, B, C, D, & E. Oxford Psychologists Press.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. Language Testing, 10, 355-371.

Read, J. (2000). Assessing vocabulary. Cambridge University Press.

Sheng, L. & McGregor, K. (2010). Lexical-semantic organisation in children with Specific Language Impairment. Journal of Speech, Language and Hearing Research, 53, 146-159.

Sutton-Spence, R. & Woll, B. (1998). The linguistics of British Sign Language: An introduction. Cambridge, UK: Cambridge University Press.

Swanwick, R. & Gregory, S. (2007). Sign bilingual education: policy and practice. Coleford: Douglas McLean.

Thompson, R.L., Vinson, D.P., & Vigliocco, G. (2009). The link between form and meaning in American Sign Language: Lexical processing effects. Journal of Experimental Psychology: Learning, Memory and Cognition, 35, 550-557.

Tweney, R. D. & Hoemann, H. W. (1973). The development of semantic associations in profoundly deaf children. Journal of Speech and Hearing Research, 16, 309-318.

Vinson, D., Cormier, K., Denmark, T., Schembri, A., & Vigliocco, G. (2008). The British Sign Language (BSL) Norms for Age of Acquisition, Familiarity and Iconicity. Behavior Research Methods, 40, 1079-1087.

White, A. & Tischler, S. (1999). Receptive sign vocabulary tests: Tests of sign-word vocabulary or iconicity. American Annals of the Deaf, 144, 334-338.

Wildemann, A. (2008). Perlesko – ein Testinstrument zur Überprüfung sprachlicher Kompetenz. Ein Erfahrungsbericht. Das Zeichen, 78, 96-101.

Woolfe, T., Herman, R., Roy, P., & Woll, B. (2010) Early vocabulary development in deaf native signers: a British Sign Language adaptation of the communicative development inventories. Journal of Child Psychology and Psychiatry, 51, 322 – 331.

Table 1. Participant characteristics

Participant	Age	Gender	Cochlear Implant	Parental hearing status	Age of first exposure to BSL	School
1	8;11	F	N	Hearing	36	Unit/Resource base
2	5;1	M	N	Deaf	0	Unit/Resource base
3	6;6	F	N	Hearing	0	Unit/Resource base
4	6;5	F	Y	Hearing	18	Unit/Resource base
5	10;3	F	N	Hearing	Not available	Unit/Resource base
6	15;5	F	N	Hearing	Not available	Deaf school
7	13;3	F	N	Hearing	60	Deaf school
8	11;3	M	N	Hearing	Not available	Deaf school
9	11;8	F	Y	Hearing	Not available	Deaf school
10	10;2	F	N	Deaf	0	Deaf school
11	11;1	M	N	Hearing	Not available	Deaf school
12	10;11	M	Y	Hearing	Not available	Deaf school
13	9;3	F	Y	Hearing	Not available	Deaf school
14	14;6	M	N	Deaf	0	Deaf school
15	16;5	M	N	Hearing	120	Deaf school
16	15;1	F	N	Deaf	0	Deaf school
17	11;0	M	N	Deaf	0	Deaf school
18	11;9	M	N	Hearing	120	Deaf school
19	15;2	M	N	Deaf	0	Deaf school
20	14;7	M	N	Deaf	0	Deaf school
21	11;5	M	N	Hearing	60	Deaf school
22	10;11	F	Y	Hearing	36	Unit/Resource base
23	8;0	F	N	Hearing	36	Unit/Resource base

24	9;4	M	N	Hearing	36	Unit/Resource base

Table 2: Item design for the four mapping tasks, using the BSL sign BUS

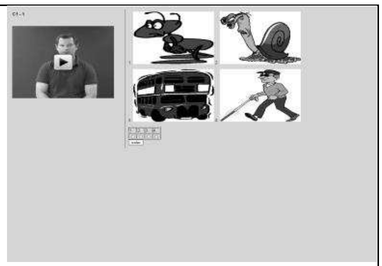
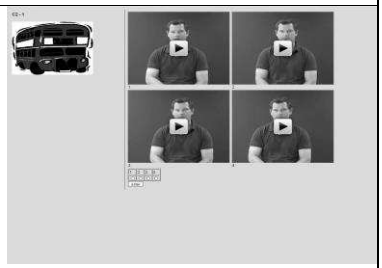
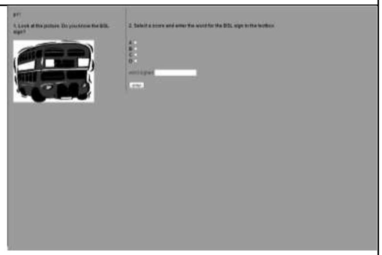
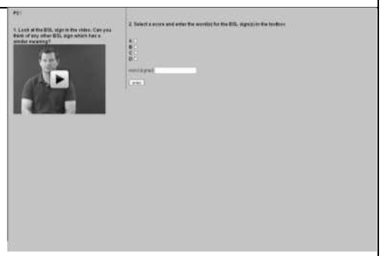
Meaning recognition	Stimulus: BSL sign (form): Set: limited Task: choose target from 4 pictures Mapping: form → meaning	
Form recognition	Stimulus: Picture (meaning) Set: limited Task: choose target from 4 BSL signs Mapping: meaning → form	
Form recall	Stimulus: Picture (meaning) Set: open Task: produce BSL sign Mapping: meaning → form	
Meaning recall	Stimulus: BSL sign (form) Set: open Task: produce BSL sign (form for related meaning) Mapping: form → meaning → meaning → form	

Table 3: Target and distractor pictures for the BSL sign DOG in the meaning recognition task

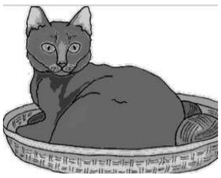
Choice #1: Target picture 	Choice #2: Phonological distractor 
Choice #3: Semantic distractor 	Choice #4: Visual distractor 

Table 4: Target and distractor signs for the picture dog in the form recognition task

Choice #1: Unrelated distractor  CUTLERY	Choice #2: Phonological distractor  SHOP
Choice #3: Target sign  DOG	Choice #4: Semantic distractor  CAT

Table 5. Scores for the vocabulary tasks and non-verbal IQ

Test raw scores	Mean	SD	Minimum-Maximum
Raven's Progressive Matrices	31.42	10.14	10 - 40
Meaning recognition	100.54 (87.78%)	12.79	60 - 115
Form recognition	97.21 (81.02%)	11.36	61 - 111
Form recall	73.71 (61.43%)	14.45	34 - 97.50
Meaning recall	62.96 (52.47%)	28.08	1-103.50

Note. Each vocabulary test has a maximum score of 120, and the Raven's Progressive Matrices has a maximum score of 60.

Table 6. Types of responses to the meaning recognition task

Response	Mean	SD	Minimum-Maximum
Categorical	13.80 (30.00%)	8.67	1-28
Schematic	25.00 (54.35%)	7.33	14-38
Phonological association	0.45 (0.98%)	0.83	0-3
"Do not know" error	3.80 (8.26%)	4.54	0-18
Other error	2.90 (6.30%)	2.57	0-8
Total error	7.65 (16.63%)	4.38	1-19

Note. A total of 46 items were analysed.