The British Sign Language (BSL) norms for age of acquisition, familiarity, and iconicity

**DAVID P. VINSON, KEARSY CORMIER, TANYA DENMARK,
ADAM SCHEMBRI, AND GABRIELLA VIGLIOCCO**
University College London, London, England

Research on signed languages offers the opportunity to address many important questions about language that it may not be possible to address via studies of spoken languages alone. Many such studies, however, are inherently limited, because there exist hardly any norms for lexical variables that have appeared to play important roles in spoken language processing. Here, we present a set of norms for age of acquisition, familiarity, and iconicity for 300 British Sign Language (BSL) signs, as rated by deaf signers, in the hope that they may prove useful to other researchers studying BSL and other signed languages. These norms may be downloaded from www.psychonomic.org/archive.

For many years, psycholinguistic research has been dominated by studies of spoken languages,[1] leading to the present situation in which evidence from spoken language is generally taken to represent language in general. Within this kind of framework, important advances have been made by contrasting languages that have different properties, leading to theoretical refinements at various levels of detail (see, e.g., Slobin, 1996). But in many cases, because these contrasts are limited to comparisons between spoken languages, researchers have neglected signed languages, which also can provide important evidence about language in general. For example, spoken languages rely primarily on a single articulator, whereas signed languages involve multiple articulators (both hands, the mouth, the face, the body, etc.) operating in tandem, a difference that has numerous implications for models of comprehension and production (Emmorey, 2002). Although such differences do not of course rule out the possibility that the processing of spoken and that of signed languages are very similar in nature (especially if one redefines "spoken language" as speech and coverbal gesture), research on signed languages is crucial to identify convergence and divergence between the two (e.g., Thompson, Emmorey, & Gollan, 2005, revealed both important similarities and differences in tip-of-the-tongue states in spoken languages and tip-of-the-finger states in signed languages).

One important factor limiting the extent of signed language research is the absence of normative information for lexical signs. In spoken language research, hundreds of studies have highlighted the importance of various lexical properties, especially when previous theoretical conclusions have been called into doubt because of experimental confounds with one or another uncontrolled lexical variable (see, e.g., Chiarello, Liu, Shears, & Kacinik, 2002). In

turn, this has led to more and more sophisticated studies designed to unravel the contributions of various lexical variables, using normative samples of thousands of words (e.g., subjective ratings of age of acquisition [AoA], imageability, and familiarity for thousands of English words; Bird, Franklin, & Howard, 2001; Cortese & Fugett, 2004; Cortese & Khanna, 2008; Gilhooly & Logie, 1980; Stadthagen-Gonzalez & Davis, 2006), in addition to massive amounts of information derived from text corpora (e.g., frequency of occurrence in the British National Corpus sample of 100 million words, www.natcorp.ox.ac.uk) and extremely large data sets, such as the English Lexicon Project (Balota et al., 2007), containing word naming and lexical decision latencies for more than 40,000 English words (see Baayen, 2005, for further discussion). Such studies provide crucial information about the role of such variables and highlight the importance of controlling them in various kinds of experimental studies not specifically related to them. To date, however, there has been no known large-scale study to collect such normative data on a signed language.[2]

The present study is the first step in this direction, reporting deaf participants' subjective ratings of familiarity, AoA, and iconicity for a set of lexical signs from British Sign Language (BSL). We included familiarity and AoA on the basis of spoken language research indicating the important effects of these variables on various kinds of language tasks, and we included iconicity because of its potential role in signed languages (e.g., Vigliocco, Vinson, Woolfe, Dye, & Woll, 2005).

As Stadthagen-Gonzalez and Davis (2006) have discussed, the exact components of familiarity are not clear, and there is some debate in the literature about the precise role of familiarity in lexical processing. Subjective ratings of familiarity are closely linked to lexical frequency, a

D. P. Vinson, d.vinson@ucl.ac.uk

variable that has been shown to be a very strong predictor of performance in any number of lexical tasks (see, e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). Although some researchers have claimed that familiarity is a superior predictor (e.g., Gernsbacher, 1984), current evidence seems to suggest that of the two, objective measures of lexical frequency (e.g., corpus counts) are stronger predictors of lexical processing effects (Balota et al., 2004), and the unique contribution of familiarity is quite small once effects of frequency have been factored out (see Stadthagen-Gonzalez & Davis, 2006). In signed languages, however, objective frequency measures are far more difficult to come by (although see note 2) and are nonexistent for BSL. Even given the advent of digital video and the greater availability of sign language production data in recent years, obtaining frequency counts involves a large-scale coding effort, further hampered by the fact that sign languages do not have writing systems. However, psycholinguistic experimentation in BSL is ongoing, and the need to exert appropriate control on materials is becoming more and more evident. Here, we take the approach of collecting subjective familiarity ratings not only to allow us to control for the variable of familiarity itself, but also as a proxy for frequency until such time as objective measures of BSL frequency become available (at which time the dissociation between them can be investigated). This approach is indirectly supported by Stadthagen-Gonzalez and Davis's findings that English familiarity ratings are more highly correlated with measures of spoken frequency (e.g., British National Corpus spoken, $r = +.72$) than with measures of written frequency (e.g., British National Corpus written, $r = +.57$).

AoA (i.e., the reported age at which a speaker/signer first learns a word/sign) has been studied extensively over the years, with studies revealing that adult ratings of AoA are valid estimates of the actual age at which words are learned (Gilhooly & Gilhooly, 1980; Morrison, Chappell, & Ellis, 1997). Early research showed that participants were faster in picture naming for pictures whose names are acquired earlier in life (Carroll & White, 1973). As is the case for familiarity, however, the exact role of AoA in processing has since been a matter of extensive debate in the literature, mainly because of its correlation with other factors. Current results suggest that estimates of AoA do significantly predict processing effects once effects of other variables are factored out (e.g., Brysbaert & Ghyselinck, 2006; Cortese & Khanna, 2007; Ghyselinck, Custers, & Brysbaert, 2004), a finding that has numerous implications for models of language acquisition (see Juhasz, 2005, for a comprehensive review). For signed language research, measures of AoA are especially important, since most learners of signed languages do not acquire a signed language from infancy, instead exhibiting a wide variety of acquisition profiles (e.g., Emmorey, 2002). As a result, this variable may play very different roles in signed and spoken languages.

Finally, ratings of iconicity are especially important when it comes to research involving signed languages. *Iconicity*—the presence of a relationship between a lexical item's meaning and its phonological form—is extremely limited in most spoken languages, being restricted mostly to onomatopoeia (e.g., in English, words such as *crash*, *pow*, *buzz*, etc.). In signed languages, however, iconicity based on visual links between form and meaning is pervasive, as is illustrated in Figure 1. For example, the BSL lexical signs for *camera* and *eat* (upper panel of Figure 1) use the hands to imitate the way humans hold objects or perform actions. Not all signs are iconic, however; some signs' phonological forms do not resemble their referents, and others refer to abstract concepts with no physical referents, such as the BSL signs for *boy*, *easy*, *important*, and

Figure 1. Still images from BSL signs. Top row, iconic signs (*camera*, *eat*). Second row, noniconic signs (*boy*, *easy*). Third row, noniconic signs (*important*, *right*). Bottom row, iconic signs referring to abstract concepts (*time*, *think*).

right (as in "correct"), illustrated in the middle panels of Figure 1. However, iconicity is not restricted to concrete objects; it can also be used metaphorically, as in the BSL signs for *time* or *think* (lower panel of Figure 1). There is only very limited research on iconicity in spoken languages, although much attention has been paid to imageability—how easily a word evokes a mental image—and to the closely related concept of concreteness, both of which may be related to iconicity to an important extent and both of which predict lexical processing effects beyond other variables (see Balota, 1990; Balota et al., 2004). Many studies investigating imageability or concreteness have treated the two interchangeably (e.g., Binder, Westbury, McKiernan, Possing, & Medler, 2005), although they can be dissociated (Bird et al., 2001) and their precise relationship remains a matter of debate (Kousta, Vinson, & Vigliocco, 2007). Most important here, although iconicity is related to both imageability and concreteness (most highly imageable concrete signs are also iconic), there are also many cases where they dissociate. For example, many highly abstract, less imageable signs are also extremely iconic, as in the metaphorically iconic signs for *time* and *think* illustrated in Figure 1. And although there is a strong tendency for imageable and/or concrete signs to be iconic, many counterexamples can be found. For example, signs referring to people (e.g., *boy*, *girl*, *parents*) tend to be noniconic despite high concreteness and imageability (at least on the basis of ratings of their English translations), and colors, which tend to be extremely imageable (again according to English ratings), also tend to be strongly noniconic. Ratings of iconicity can provide crucial information for experimental studies of various kinds. Beyond the obvious example of selecting items for studies investigating the role of iconicity itself (and eventually exploring its relationship with imageability and concreteness), such ratings can also allow the selection of items that are well controlled in order to contrast spoken and signed languages with iconicity held constant between the two, which is useful, for example, in studies designed to investigate effects of language modality without this factor being confounded with iconicity.

METHOD

Participants

Ratings were obtained from a set of 20 deaf participants for each of the three variables (AoA, familiarity, and iconicity). Ratings for each variable were obtained separately, although many participants completed more than one rating task. In all, 33 deaf BSL signers (18 women, 15 men) participated. Because of the duration of the task, only 11 completed all three rating tasks; 6 completed two of the three, and the remaining 16 each completed only one.

Participants were selected on the basis of their responses to a thorough online questionnaire, which all potential participants filled in before taking part. Eligible participants were those who identified themselves as deaf users of BSL who were British-born and between the ages of 18 and 62. Participants were recruited at various events held by deafness-related organizations around the United Kingdom and through the authors' personal contacts. Most of the participants (*n* = 29) reported that they were born deaf, with the remainder having become deaf before the age of 5.

Concerning language, all participants were BSL signers, and the vast majority considered their preferred language to be BSL.[3] Par-

ticipants self-rated their BSL skills on average as 6.2 on a 7-point scale (range = 5–7, *SD* = 0.8). Most began to sign early in life; 14 participants started signing before age 3, 7 from between the ages of 3 and 5, and 6 participants began to sign at the age of 15 or later. They learned to sign from a variety of sources (16 from family at home, 11 from friends, 9 from other students at school, and 3 from other sources).[4] Most of the participants (*n* = 24) were born in the southeast of England, with a few participants from other areas (5 from southwest England, 1 from northwest England, 1 from northeast England, and 2 from Scotland). Most participants (*n* = 27) also currently lived in southeast England.

Materials

We initially selected a very large set of BSL signs from a variety of sources, including dictionaries of BSL based on linguistic principles (e.g., Brien, 1992) and other collections of signs specifically created as teaching materials (Microbooks, 2005). We also selected as many signs as possible for clearly pictureable objects and events, on the basis of lists of English words for which norms exist (e.g., Snodgrass & Vanderwart, 1980) together with materials used in previous research in various domains (e.g., Cree & McRae, 2003; McRae, Cree, Seidenberg, McNorgan, 2005; Vinson & Vigliocco, 2008; among many others). We also selected a wide range of items varying in familiarity and AoA on the basis of available English norms and varying in iconicity on the basis of our own judgments. Furthermore, we selected specific subsets of signs that might be less familiar to many signers, including those identified as regionalisms (i.e., signs listed in the BSL dictionary as used only in particular regions of the United Kingdom), colloquialisms, and signs believed to be recent borrowings from foreign sign languages. Drawing on these sources, the third author (herself a native signer) worked closely with a deaf native signer to compile an initial list of 2,490 BSL signs. The list was then more closely examined by the sign language linguists on the research team, who used two main criteria to select a smaller subset of 309 signs (300 test items plus 9 practice items) for the norming study.

The first criterion was based on known organizational principles of sign language lexicons (Brentari & Padden, 2001; Johnston & Schembri, 2007). Briefly, signs in BSL may be analyzed as fitting into three major categories, based on differences in phonological, grammatical, and semantic properties: *core native signs* (signs that behave similarly to words in a spoken language, known as *lexical signs*), *nonnative signs* (signs based on the manual alphabet, known as fingerspelled loan signs), and *noncore native signs* (highly iconic nonlexicalized signs known as *classifier constructions*). In particular, we wished to exclude classifier constructions, which are highly iconic, productive, complex forms that appear to share some properties with gestural communication (Emmorey & Herzig, 2003; Schembri, Jones, & Burnham, 2005) and have been the focus of very little psycholinguistic research. Although we included a small number of fingerspelled loan signs (e.g., the fingerspelled sequence T–H, meaning "Thursday"), the list was dominated by lexical signs.

The second criterion we used in the selection of the task was to select signs that exhibited as little polysemy as possible and that did not have known examples of homonyms. This was motivated by a desire to ensure that the ratings for the signs used in the task actually reflected ratings for the target lexical item and not some sign with a related meaning or a formationally similar sign with a different meaning. Therefore, 3 deaf native signers who took part in our later stages of piloting were also asked to provide English translations in addition to their ratings. Any signs that were found to be polysemous in piloting were then excluded from the final list.

Although norming studies on spoken languages use a significantly greater number of lexical items than our set of 300 signs (e.g., Stadthagen-Gonzalez & Davis, 2006, normed a set of 1,526 English words, which, when combined with the norms of Gilhooly & Logie, 1980, yielded ratings for 3,394 words), this difference reflects the relatively longer time required for a lexical norming task in a signed language, as well as our expectation that deaf BSL signers would be

much more difficult to find than native English speakers. Tasks involving written stimuli can be completed very quickly; Stadthagen-Gonzalez and Davis reported that each of their participants took no more than 30 min to rate 366 English words (a maximum of 12 sec per word), typing their responses directly into a spreadsheet. For signed languages, participants must view and rate video clips of signs, which, themselves, are not minimal in duration. With 300 signs, each task took between 60 and 90 min to complete. In addition, we collected ratings on iconicity, familiarity, and AoA separately, because the duration of the combined task would have been unmanageable. Thus, participants could choose whether to do one, two, or all three tasks, at different times and on different days if they preferred.

The final set of 309 signs was filmed by a professional film production company. Four deaf BSL signers presented the stimuli for filming; each of them presented approximately 75 signs. Before filming, two of the researchers went through the sign list with the presenters. The signs were then delivered to the presenters by the third author, a native BSL signer. The production of 90% of all 309 signs used in the task was spontaneously accompanied by the mouthing of the equivalent English word. Because mouthing is a common feature of BSL discourse (Sutton-Spence & Day, 2001), we did not attempt to prevent the presenters from mouthing while being filmed.

Procedure

The resulting 300 video clips were inserted into an online survey tool, RiddleMeThis (www.riddlemethis.com), which is capable of presenting video stimuli and collating responses in text files. Three different surveys were created for each of the three variables (AoA, familiarity, and iconicity), presenting each sign on a separate screen. The 300 signs were always presented in an individually randomized order. All consent information and instructions for the tasks were provided in written English and in BSL (video format). Breaks were provided every 75 signs, and participants were also permitted to stop the task at any time and to restart later; the system was set to resume at the stopping point as long as the participant continued within a month's time.

Raters for AoA were asked to estimate the age at which they thought they had learned each of the 300 signs. Each screen showed the target sign and a list of age ranges—0–2, 3–4, 5–6, 7–8, 9–10, 11–12, 13–14, 15–16, and 17 or older—plus an additional option to select if they did not know that sign. Raters for familiarity were asked to estimate, on a scale of 1–7, how often they see each of the 300 signs (1, *participant had never seen the sign before*; 7, *participant sees the sign very often*). Each screen showed the target sign and the 1–7 response scale. Raters for iconicity were asked to estimate, on a scale of 1–7 (1, *not at all iconic*; 7, *highly iconic*), the extent to which they thought each of the 300 signs was iconic. The English version of the instructions used for each variable is shown in the Appendix. The instructions for AoA and familiarity were based largely on those used by Stadthagen-Gonzalez and Davis (2006).

During our piloting, some participants gave high iconicity ratings to signs based on the fingerspelling system of BSL (e.g., the BSL lexical sign for *father*, which is based on the fingerspelled letter F), in addition to the meaning–form correspondences that we expected. Therefore, we added to our instructions for iconicity a warning to participants that they should avoid conflating iconicity with motivation. That is, in order to be considered iconic, the sign had to visually resemble or represent the referent, rather than a letter from the manual alphabet associated with the semantically equivalent English word.

RESULTS AND DISCUSSION

Each participant's responses were saved as a text file. We first examined their responses to the initial questions to ensure that each participant had indeed given their con-

sent to participate. Next, we extracted the participants' ratings for each of the BSL signs. In some cases (1.3% of all signs), a sign was presented to a participant more than once in the course of the task; this usually occurred because of software error and/or when a participant stopped the task and resumed later. In any such instances, the average of the multiple responses was taken.

For the familiarity task, participants' responses were simply values on a 1–7 scale and required no further processing. The iconicity task also permitted a response of 8 (*I don't know the meaning of the sign*); we considered only values of 1–7 as iconicity ratings. Finally, we converted AoA ratings from ranges into numeric values; for all intermediate values, this was done by taking the midpoint of a range (e.g., "age 5–6" was assigned a value of 5.5). The endpoints of the scale were treated differently; the range of 0–2 was given a value of 1.5, and the range of 17+ a value of 17.5. The latter conversion will necessarily result in an underestimate of the age of signs learned in adulthood, so AoA values in the high teens should not be thought of as an accurate reflection of the actual AoA. More generally, the varied language experience of the "typical" BSL learner implies that AoA values are best considered in relation to each other, rather than as a valid estimate of objective AoA—the actual age at which any sign might be expected to be learned. Average iconicity, familiarity, and AoA ratings for all items across participants are provided as supplementary materials (available for download; see the Archived Materials section).

Our next step was to remove the signs that were generally unfamiliar to a majority of our participants, because the reliability of average ratings is in doubt in such cases. We decided to exclude any signs that were given ratings of *I don't know this sign* by more than half of the participants in either the iconicity or the AoA rating tasks. There were 15 such signs, of which all but 1 were regional signs unknown to the majority of our participants (see Brien, 1992). Familiarity ratings confirmed that these signs were highly unfamiliar overall (the average familiarity rating was 1.89 on a 7-point scale). These items were removed before further analyses were conducted. For AoA, we also conducted one additional set of analyses to see whether responses by the 5 participants in this task who reported learning BSL late in life were similar to those for participants who had learned BSL early. To do this, we calculated the correlations between the AoA ratings by each possible pair of participants. If late learners were approaching the task differently, or if their responses lacked enough variability to be useful, ratings by late learners should exhibit far less correlation with the group of participants as a whole, compared with those of early learners. The average interparticipant correlation across all participants was $r = +.566$ ($p < .01$). Although late learners' ratings were slightly less correlated to those of the other participants (average $r = +.547, p < .01$), this difference was minimal, and some individual late learners' responses were among the most highly correlated of all (rs $= +.616$ and $+.611$, third and fourth highest among all participants). We therefore kept all responses, whether participants had acquired BSL early or later in life.

Our next step was to carry out an informal assessment of the items that received ratings at the extremes of each scale, to see whether average ratings matched our intuitions. The signs that received the highest iconicity ratings are those signs that would generally be considered transparently iconic; that is, one would expect their meanings to be easily guessed, even by nonsigners (Lieberth & Gamble, 1991). Signs rated as most highly iconic included examples such as *camera* ($M = 6.85$) and *eat* ($M = 6.80$), which are highly recognizable even to nonsigners (illustrated in the upper panel of Figure 1). The motivation of these signs is similar to that of American Sign Language (ASL) signs rated as transparently iconic by nonsigners (Lieberth & Gamble, 1991). Signs with low iconicity ratings included examples such as *easy* ($M = 1.60$), *boy* ($M = 1.85$), *important* ($M = 2.45$), and *right* ($M = 2.30$), which do not exhibit any apparent resemblance to referents (illustrated in the middle panels of Figure 1). Note that a few of the signs with the very lowest iconicity ratings also received very low familiarity ratings (see below). It is not surprising that less familiar signs will also receive low iconicity ratings, because participants may be unable (or somewhat unable) to visualize the referent. For this reason, iconicity ratings for items rated low in familiarity may underestimate true iconicity among individuals who are familiar with a particular sign, an important consideration for use of such items in experimentation.

The items rated as the most familiar are all concepts regularly used in everyday conversation: *work* ($M = 6.90$), *eat* ($M = 6.85$), and *what* ($M = 6.80$), whereas the least familiar signs include *Basingstoke*[5] ($M = 1.95$), *address* ($M = 2.35$), *of course* ($M = 2.35$), *shabby* ($M = 2.45$), and *people* ($M = 2.55$).[6] For AoA, signs rated as the earliest acquired are *eat* ($M = 3.17$ years), *ice cream* ($M = 3.61$ years), *cry* ($M = 3.61$ years), *draw* ($M = 4.32$ years), *duck* ($M = 4.50$ years), and *rabbit* ($M = 4.50$ years).[7] The latest acquired signs were a mix of abstract signs, specific locations, and modern technology added only recently to the BSL lexicon (e.g., *metaphor*, $M = 17.14$ years; *Basingstoke*, $M = 16.83$ years; *DVD*, $M = 16.61$ years; *webcam*, $M = 16.57$ years; *email*, $M = 16.50$ years; *Malaysia*, $M = 16.50$ years; and *Melbourne*, $M = 16.50$ years). In addition to examining the ends of the scales, we also looked at the distribution of ratings across each scale by creating frequency histograms (see Figure 2).

AoA ratings (top panel of Figure 2) were distributed across the scale, illustrating that participants tended to use the entire range of possible ages (most participants learned BSL early in life, so this was possible). Ratings of iconicity, instead, were bimodally distributed, perhaps indicating a tendency toward a distinction between signs considered to be iconic and signs that are not. Finally, familiarity ratings were heavily skewed toward the higher end of the scale.[8] One explanation for this might be that our item selection process favored signs that are highly familiar. We sought to include many signs that are nonpolysemous (or with limited polysemy) and less regionally variant, so that they would be unambiguous to most of our participants. This may also reflect lexicalization processes in sign language lexicons, because it may be that core native signs tend to develop for frequently encountered phenomena, whereas lexical borrowing by means of fingerspelling or lexicalization of classifier constructions is used for rarer concepts (Brentari & Padden, 2001).

Because we included the same signs on all three tasks, we were also able to examine the degree of correspon-
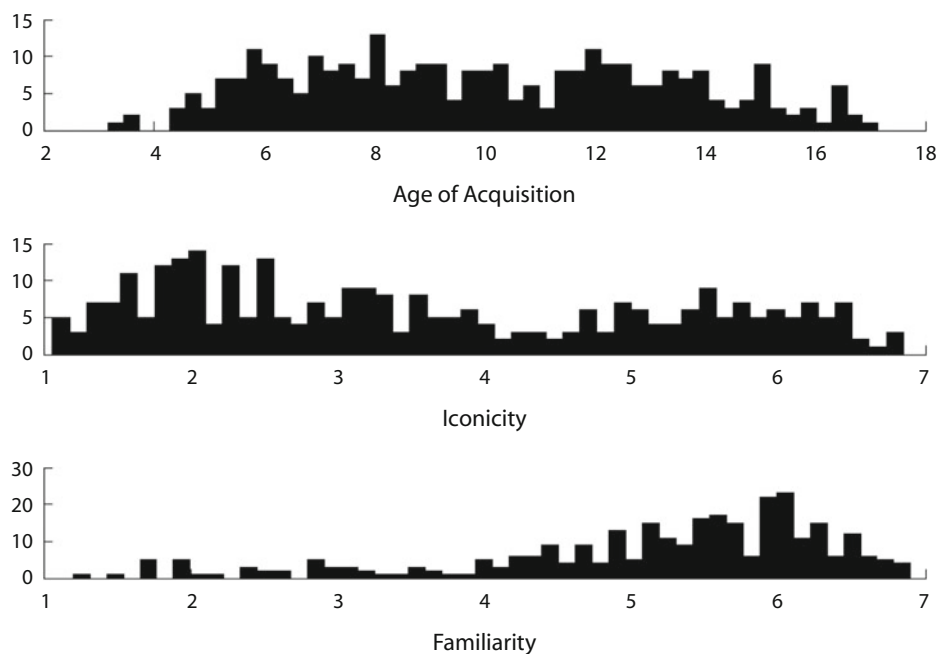

Figure 2. Frequency histograms reflecting the distribution of ratings for the 300 BSL signs normed here.

dence between AoA, iconicity, and familiarity rating scales (also illustrated as scatterplots in Figure 3). We did so by calculating correlation coefficients for the average ratings for each item. The greatest degree of correlation was between AoA and iconicity ($r = -.463$, $p < .01$); early acquired signs tended to be rated the most iconic. This trend may seem to contradict claims from research in ASL that iconic signs are not overly represented in children's earliest signs (Orlansky & Bonvillian, 1984) and that children's errors in producing iconic signs are not more iconic than the adult form (Meier, Mauk, Cheek, & Moreland, 2008). However, Orlansky and Bonvillian's claim was made only on the basis of parental reports from a limited number of observers, and even if this is true, a correlation between iconicity and AoA does not strictly imply that the earliest acquired signs will be noticeably more iconic.[9] Furthermore, Meier et al. pointed out that their data indicated only that children's errors did not seem to enhance iconic properties of ASL signs and could not be taken as evidence that children are insensitive to all types of iconicity. Slobin et al. (2003) also claimed

that iconicity may have a role in the acquisition of classifier constructions. In any event, existence of a negative correlation between iconicity and AoA does not necessarily mean that iconicity will affect children's language processing.

AoA was also highly correlated with familiarity ($r = -.433$, $p < .01$); signs acquired early also tended to be more familiar. Contrary to what might be expected, the correlation between iconicity and familiarity was much lower ($r = +.146$, $p < .05$); there was a reliable tendency for iconic signs to be more familiar, but this was not particularly strong. This was especially true at the high end of the familiarity scale, where many of the most familiar signs are extremely low in iconicity, such as the signs depicted in the middle panels of Figure 1, all with iconicity ratings less than 2.5: *boy* (familiarity = 6.35), *easy* (familiarity = 6.70), *important* (familiarity = 6.50), and *right* (familiarity = 6.65). This highlights the fact that although iconicity is pervasive in signed languages, many arbitrary (noniconic) signs are highly familiar and, thus, probably also frequently occurring.
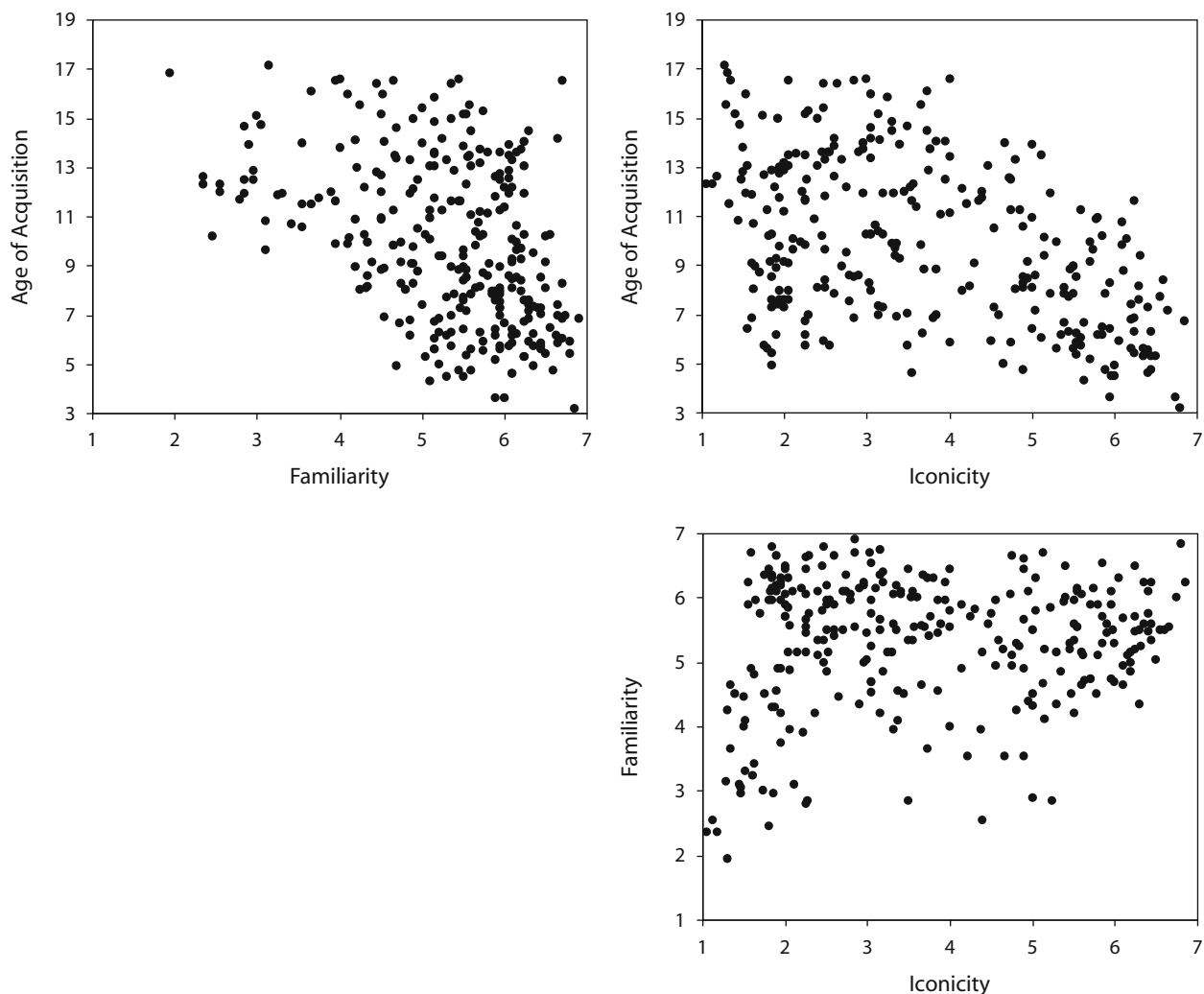

Figure 3. Scatterplots depicting the relations between age of acquisition, iconicity, and familiarity ratings.

We were also able to compare our AoA data with an independently obtained data set (Gaynor, Woolfe, & Woll, in press), in order to obtain an additional measure of reliability. Unlike in the present study, in which adults were asked to estimate the age at which they learned a sign, the data collected by Gaynor et al. were based on longitudinal testing of children whose parents were asked to complete a checklist of signs that a child comprehends at various ages. From this data set, we selected the data from the last report from each deaf child tested (*n* = 34), thus reflecting the greatest extent of language development. The average age at the time of testing was 29.3 months (range = 13–47).

There were 85 signs in common between the two data sets, with an average AoA rating of 7.3 years (range = 3.2–13.9, with *eat* at the low and *turtle* at the high end). For each of these signs, we calculated the proportion of children who comprehended and produced each sign at the time of testing. These measures should be negatively correlated with AoA; the more children who comprehend/produce a sign at a given age, the earlier acquired that sign should be. Across these signs, the average proportion of comprehension was .61 (range = 0–.94, with *not care* at the low and *mother* at the high end), and the average proportion of production was .54 (range = 0–1.00, with *not care* at the low and *dog* at the high end). We then calculated the correlations between these measures and our participants' estimates of AoA. AoA was significantly negatively correlated with both the proportion of children who comprehended each sign ($r = -.526$, $p < .01$) and the proportion of children who produced each sign ($r = -.499$, $p < .01$).[10] Despite the numerous differences in the methodology and populations tested in these two data sets, these results nicely illustrate a strong degree of correspondence between adults' estimates of the relative ages at which they acquired a sign and parents' judgments of what their children can actually comprehend and produce at a given age. This is especially important given the number of late BSL learners who contributed to our ratings of AoA.

We provide these norms in the hope that they can prove useful to other researchers of British Sign Language and of other signed languages.

REFERENCES

BAAYEN, R. H. (2005). Data mining at the intersection of psychology and linguistics. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 69-83). Hillsdale, NJ: Erlbaum.

BALOTA, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9-32). Hillsdale, NJ: Erlbaum.

BALOTA, D. A., CORTESE, M. J., SERGENT-MARSHALL, S. D., SPIELER, D. H., & YAP, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, **133**, 283-316.

BALOTA, D. A., YAP, M. J., CORTESE, M. J., HUTCHISON, K. I., KESSLER, B., LOFTIS, B., ET AL. (2007). The English Lexicon Project. *Behavior Research Methods*, **39**, 445-459.

BINDER, J. R., WESTBURY, C. F., MCKIERNAN, K. A., POSSING, E. T., & MEDLER, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, **17**, 905-917.

BIRD, H., FRANKLIN, S., & HOWARD, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verb and function words. *Behavior Research Methods, Instruments, & Computers*, **33**, 73-79.

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: Erlbaum.

BRIEN, D. (ED.) (1992). *Dictionary of British Sign Language/English*. Boston: Faber & Faber.

BRYSBAERT, M., & GHYSELINCK, M. (2006). The effect of age of acquisition: Partly frequency-related, partly frequency-independent. *Visual Cognition*, **13**, 992-1011.

CARREIRAS, M., GUTIÉRREZ-SIGUT, E., BAQUERO, S., & CORINA, D. (2008). Lexical processing in Spanish Sign Language (LSE). *Journal of Memory & Language*, **58**, 100-122.

CARROLL, J. B., & WHITE, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, **25**, 85-95.

CHIARELLO, C., LIU, S., SHEARS, C., & KACINIK, N. (2002). Differential asymmetries for recognizing nouns and verbs: Where are they? *Neuropsychology*, **16**, 35-48.

CORTESE, M. J., & FUGETT, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, **36**, 384-387.

CORTESE, M. J., & KHANNA, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Quarterly Journal of Experimental Psychology*, **60**, 1072-1082.

CORTESE, M. J., & KHANNA, M. M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behavior Research Methods*, **40**, 791-794.

CREE, G. S., & MCRAE, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, **132**, 163-201.

EMMOREY, K. (2002). *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Erlbaum.

EMMOREY, K., & HERZIG, M. (2003). Categorical versus gradient properties of classifier constructions in ASL. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 221-246). Mahwah, NJ: Erlbaum.

GAYNOR, A., WOOLFE, T., & WOLL, B. (in press). Early lexical development in British Sign Language. *Journal of Deaf Studies & Deaf Education*.

GERNSBACHER, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, **113**, 256-281.

GHYSELINCK, M., CUSTERS, R., & BRYSBAERT, M. (2004). The effect of age of acquisition in visual word processing: Further evidence for the semantic hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 550-554.

GILHOOLY, K. J., & GILHOOLY, M. L. (1980). The validity of age-of-acquisition ratings. *British Journal of Psychology*, **71**, 105-110.

GILHOOLY, K. J., & LOGIE, R. H. (1980). Age of acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, **12**, 395-427.

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

JUHASZ, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, **131**, 684-712.

KOUSTA, S.-T., VINSON, D. P., & VIGLIOCCO, G. (2007). *Believe it or not, abstract words have a processing advantage over concrete words.*

Presented at the 48th Annual Meeting of the Psychonomic Society, Long Beach, CA.

Lieberth, A. K., & Gamble, M. E. B. (1991). The role of iconicity in sign language learning by hearing adults. *Journal of Communication Disorders*, **24**, 89-99.

McKee, D., & Kennedy, G. D. (2006). The distribution of signs in New Zealand Sign Language. *Sign Language Studies*, **6**, 372-390.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, **37**, 547-559.

Meier, R. P., Mauk, C. E., Cheek, A., & Moreland, C. J. (2008). The form of children's early signs: Iconic or motoric determinants? *Language Learning & Development*, **4**, 63-98.

Microbooks (2005). *The advanced dictionary of British Sign Language* [CD-ROM]. Sunbury-on-Thames, U.K.: Microbooks.

Morford, J., & MacFarlane, J. (2003). Frequency characteristics of American Sign Language. *Sign Language Studies*, **3**, 213-225.

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology*, **50A**, 528-559.

Orlansky, M. D., & Bonvillian, J. D. (1984). The role of iconicity in early sign language acquisition. *Journal of Speech & Hearing Disorders*, **49**, 287-292.

Schembri, A., Jones, C., & Burnham, D. (2005). Comparing action gestures and classifier verbs of motion: Evidence from Australian Sign Language, Taiwan Sign Language, and nonsigners' gestures without speech. *Journal of Deaf Studies & Deaf Education*, **10**, 272-290.

Slobin, D. I. (1996). From "thought and language" to "thinking to speaking." In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). Cambridge: Cambridge University Press.

Slobin, D. I., Hoiting, N., Kuntze, K., Lindert, R., Weinberg, A., Pyers, J., et al. (2003). A cognitive/functional perspective on the acquisition of "classifiers." In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 271-296). Mahwah, NJ: Erlbaum.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 174-215.

Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods*, **38**, 598-605.

Sutton-Spence, R., & Day, L. (2001). Mouthings and mouth gestures in British Sign Language (BSL). In P. Boyes Braem & R. Sutton-Spence (Eds.), *The hands are the head of the mouth: The mouth as articulator in sign languages* (pp. 69-85). Hamburg: Signum.

Thompson, R., Emmorey, K., & Gollan, T. (2005). Tip-of-the-fingers experiences by ASL signers: Insights into the organization of a sign-based lexicon. *Psychological Science*, **16**, 856-860.

Vigliocco, G., Vinson, D. P., Woolfe, T., Dye, M. W., & Woll, B. (2005). Words, signs and imagery: When the language makes the difference. *Proceedings of the Royal Society B*, **272**, 1859-1863.

Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, **40**, 183-190.

NOTES

1. We use the term *spoken languages* to refer to auditory/oral languages, whether in spoken or written form, in contrast to signed languages, which are visual/corporal and without writing systems.

2. Only one major study of lexical frequency has been undertaken, based on a corpus of 100,000 lexical items collected from 50 h of video-taped conversation and other text types in New Zealand Sign Language (McKee & Kennedy, 2006; see also Morford & MacFarlane, 2003, for a smaller scale study of frequency in American Sign Language). Familiarity norms have also recently been collected for a set of signs in Spanish

Sign Language (LSE; Carreiras, Gutiérrez-Sigut, Baquero, & Corina, 2008).

3. Two participants listed SSE (Sign Supported English, the use of BSL lexical signs but produced following English word order, rather than BSL syntax), rather than BSL, as their preferred language, but these participants did not statistically differ from the others on the rating tasks reported here.

4. Participants were allowed to indicate multiple sources, so the sum exceeds the number of participants.

5. Basingstoke is a small town in south England.

6. Many of the signs rated as less familiar were listed as regional signs (Brien, 1992).

7. The numeric values given here do not necessarily estimate the actual age at which a child learning BSL from infancy may acquire a given sign, due to the varying linguistic profiles of our participants. Instead, they should be considered relative measures only, with smaller values indicating that that sign is typically acquired earlier than signs with larger values.

8. This is true even including the 15 lowest rated items, which we excluded from analysis.

9. In our data, 17 of the 20 signs rated as earliest acquired were also iconic (i.e., had iconicity ratings above 4.0). However, we did not attempt to select the signs that are likely to be the very first acquired by children learning BSL from infancy, so this does not provide strong evidence against the findings of Orlansky and Bonvillian (1984).

10. Comprehension and production proportions were extremely highly correlated with each other ($r = +.923, p < .01$).

ARCHIVED MATERIALS

The following materials may be accessed through the Psychonomic Society's Archive of Norms, Stimuli, and Data, www.psychonomic.org/archive.

To access these files or links, search the archive for this article using the journal (*Behavior Research Methods*), the first author's name (Vinson), and the publication year (2008).

File: BSL_norming.txt provides a summary of ratings for each item. It is a tab-delimited text file with column labels in the first row.

Column:

label: Rough English translation of the sign. Corresponds to the file name of each BSL video clip (see below). Please note that file names and the accompanying labels should be considered as identifiers rather than complete translations of the BSL signs.

AoA_mean: Average age of acquisition (AoA) rating for that sign.

AoA_#: Number of participants (of 20) who gave an AoA rating for that sign.

AoA_DK: Number of participants who responded *don't know* (they did not recognize that sign).

AoA_sd: Standard deviation of AoA ratings for that sign.

AoA_known: Proportion of raters who knew a given sign (i.e., responded other than *don't know*).

Icon_mean: Average iconicity rating for that sign.

Icon_#: Number of participants who gave an iconicity rating for that sign.

Icon_DK: Number of participants who responded *don't know*.

Icon_sd: Standard deviation of iconicity ratings for that sign.

Icon_known: Proportion of raters who knew a given sign.

Famil_mean: Average familiarity rating for that sign.

Famil_sd: Standard deviation of familiarity ratings for that sign.

Famil_#: Number of participants who rated that sign.

File: BSL_norming.xls contains the above information in Microsoft Excel format.

Files: *.MOV

Each BSL lexical sign described in the data files above has an accompanying video clip in Quicktime (.MOV) format. The file name matches the first column of data files BSL_norming.

Author's e-mail address: d.vinson@ucl.ac.uk

APPENDIX
Instructions for the Rating Tasks

Instructions for Age of Acquisition Ratings (English Version)

For this task we want to know about when you learnt different BSL signs.

Please indicate how old you were when you learnt each of the signs on the list. By learning a sign we mean the age when you would have understood that sign if somebody had used it in front of you, EVEN IF YOU DID NOT use it yourself at the time.

You will see different age categories from birth to adulthood. Please pick the age when you believe you understood each sign. For example if you see the sign MILK, you may have understood this sign at an early age so you would select an early age band 0–2 or 3–4. If you see the sign PSYCHOLOGY, you may have understood this at a much later age, such as 13–14 or 15–16.

An approximate age is good enough for this rating. If you do not know the meaning of a sign, select I DO NOT KNOW THIS SIGN.

Watch each sign and begin rating them at your own speed. Work fairly quickly but do not be careless in your ratings. The important thing is for you to be as accurate as possible.

Feel free to use the entire range of ages, from 0–2 to 17+; at the same time, do not be concerned about how often you select a particular age group, as long as you are honest in your ratings.

Instructions for Iconicity Ratings (English Version)

For this task we want to know how iconic you think some BSL signs are. First we will explain what we mean by iconic.

Some signs are considered to be iconic; the sign somehow looks like what it means. One sign generally considered to be very iconic is DRINK, which looks like a person holding a cup and bringing it to their mouth. You would be able to guess this sign's meaning even if you did not know BSL. Other signs are not iconic at all; for example the sign BROTHER does not look like a brother. Signs can be motivated in other ways (i.e. there can be a reason why the sign looks the way it does). The sign FATHER clearly is clearly linked to the fingerspelled letter F, but it is not iconic because the sign does not look like what it means (it does not look like a father).

For each sign that you will see, rate on a scale of 1 to 7 how iconic you think the sign is, with 1 as not iconic at all and 7 as highly iconic. For example DRINK is very iconic; it looks just like drinking from a cup, this would be a 7. BROTHER is not at all iconic and would be a 1. Signs that are intermediate in iconicity, of course, should be rated appropriately between the two extremes, for example the sign MAN may have a rating of 3 or 4.

Feel free to use the entire range of numbers, from 1 to 7; at the same time, don't be concerned about how often you use a particular number as long as you are honest in your ratings.

Remember, do not rate fingerspelled signs as iconic just because they resemble the letters of the alphabet used in an English word—try to think how the fingerspelling does or does not look like what it means.

If you do not know the meaning of a sign, choose I DO NOT KNOW THE MEANING OF THIS SIGN. Work fairly quickly but do not be careless in your ratings, the important thing is for you to be as accurate as possible.

Instructions for Familiarity Ratings (English Version)

For this task we want to find out how often you see certain BSL signs. You will be given a list of signs and you are to rate each one as to the number of times that you have seen others in the community using it by simply choosing a number on the given 1 to 7 scale.

In this scale, 1 represents NEVER, that is, you have never seen the sign in your life; the number 2 represents RARELY, that is you have seen the sign at least once before, but only rarely and so on until 7, which represents VERY OFTEN, that is, you have seen the sign nearly every day of your life.

For example you may see others use the sign MILK very often so you would select 6 or 7, there may be some rare signs which you do not see very often e.g., the sign NICARAGUA, so you would select 2 or 3.

Do not be bothered if you are not sure of the meaning of some of the signs. Simply rate each one as to the number of times you have seen others use it regardless of its meaning.

Watch each sign and begin rating them at your own speed. Work fairly quickly but do not be careless in your ratings, the important thing is for you to be as accurate as possible.

Some of the signs in this task may be very rare, so you are not expected to have seen all of them. Just make the best estimates you can. Feel free to use the entire range of numbers, from 1 to 7; at the same time, do not be concerned about how often you use a particular number, as long as you are honest in your ratings.