Phonological similarity judgments in ASL

Evidence for maturational constraints on phonetic perception in sign

Matthew L. Hall, Victor S. Ferreira, and Rachel I. Mayberry
University of California, San Diego

We created a novel paradigm to investigate phonological processing in sign and asked how age of acquisition (AoA) may affect it. Participants indicated which of two signs was more phonologically similar to a target, and estimated the strength of the resemblance with a mouse click along a continuous scale. We manipulated AoA by testing deaf native and non-native signers, and hearing L2 signers and sign-naïve participants. Consistent with previous research, judgments by the native and L2 signers reflected similarity based on shared phonological features between signs. By contrast, judgments by the non-native signers and sign-naïve participants were influenced by other (potentially visual or somatosensory) properties of signs that native and L2 signers ignored. These results suggest that early exposure to language helps a learner discern which aspects of a linguistic signal are most likely to matter for language learning, even if that language belongs to a different modality.

Keywords: phonological similarity, critical period, native vs. non-native, second language acquisition

1. Introduction

Like spoken language, sign language has sublexical structure (Battison 1978; Frishberg 1975; Liddell & Johnson 1989; Stokoe 1960, 1978). This linguistic patterning below the level of the word is considered to be phonological in nature (Brentari 1998; Perlmutter 1992; Sandler 1989; Sandler & Lillo-Martin 2006). Like spoken phonology, sign phonology plays an organizing role in language representation and processing, such as in lexical access, memory, and production (Klima & Bellugi 1979; Newkirk, Klima, Pederson & Bellugi 1980; Thompson, Emmorey & Gollan 2005; Thompson, Vinson & Vigliocco 2010; Wilson & Emmorey 1997,

Sign Language & Linguistics 15:1 (2012), **104–127**. DOI 10.1075/sll.15.1.05hal

1998). Much of what we know about phonological processing comes from studies of native learners, individuals whose language experience began from birth, and this is true for both spoken and sign language. Non-native learning, language experience begun at ages past birth, is known to have marked effects on phonological processing in spoken language. Recent research has found that non-native learning affects phonological processing in sign language as well. Unique to sign language are late L1 learners, whose early language experience is importantly different from late L2 learners. Infants who hear normally are immersed in spoken language from birth so that their subsequent acquisition of languages at older ages is L2 acquisition. Infants born deaf can be so isolated from spoken language such that their later acquisition of sign language is late L1 acquisition (Mayberry 2010). Research has found that age of acquisition, AoA, effects are more marked for L1 than L2 learning on language comprehension and grammatical processing (Mayberry 1993; Mayberry, Lock & Kazmi 2002). The question we ask here is whether the same is true for phonological processing in sign.

Sign language research can help adjudicate two competing theoretical explanations of AoA effects on phonological representations. Maturation-based accounts (Lenneberg 1962; Penfield & Roberts 1959) argue that infants are uniquely skilled at performing the perceptual kinds of analyses that lead to native-like phonological acquisition (Werker 1989). By contrast, experience-based accounts (McCandliss, Fiez, Protopapas, Conway & McClelland 2002; Munakata & Pfaffly 2004; Seidenberg & Zevin 2006) propose that the phonology-to-meaning mappings acquired in early life become entrenched through frequent use so that learning the alternate phonological mappings of an L2 at later ages is difficult. Both hypotheses predict that L2 learners of sign language will show AoA effects, but they make different predictions for AoA effects on L1 acquisition. Maturational accounts (Mayberry 2007; Mayberry & Eichen 1991; Newport 1990) predict that AoA effects will be greater for L1 compared to L2 acquisition because infant language experience organizes perceptual space into linguistic space, or categories, necessary for subsequent language acquisition. By contrast, experience-based accounts (Flege 1999, Flege et al. 1999, Seidenberg & Zevin 2006) predict that AoA effects will be reduced for the L1 relative to the L2, because there is no previously acquired phonological system to interfere with later L1 learning. In the present study, we test these predictions by investigating AoA effects on sign phonology using a similarity judgment task.

What little we know about AoA effects on phonological processing in sign language comes largely from studies of categorical perception. An important characteristic of phonological perception in speech is that native speakers cannot discriminate small increments along a psychophysical continuum underlying a phonetic feature and instead perceive tokens from this continuum dichotomously, rather than gradiently. Categorical perception has two classic signatures: (1) a sharp boundary across which people reliably classify ambiguous stimuli as belonging to different categories, and (2) markedly enhanced discrimination of two tokens drawn from opposite sides of that boundary compared to near-chance discrimination of two tokens drawn from the same side of the boundary.

Five studies of categorical perception in American Sign Language, ASL, have found that signers overdiscriminate; that is, they can reliably perceive differences between two tokens drawn from the same side of the category boundary (Baker, Idsardi, Golinkoff & Petitto 2005; Best, Mathur, Miranda & Lillo-Martin 2010; Emmorey, McCullough & Brentari 2003; Morford, Grieve-Smith, MacFarlane, Staley & Waters 2008; Newport 1982). But the degree to which they overdiscriminate depends upon their early language experience. The earliest study (Newport 1982) observed overdiscrimination, but tested only 4 participants and did not report language background. More recently, Emmorey et al. (2003) tested deaf native signers and hearing sign-naïve participants; both groups were found to overdiscriminate, but the groups were not directly compared to one another. Baker et al. (2005) also tested deaf native signers and hearing sign-naïve participants. The naïve participants made more within-category discrimination relative to the native signers. In a fourth study, Morford et al. (2008) tested deaf native and non-native signers, and hearing L2 learners. Both the non-native and L2 signers showed greater within-category discrimination compared to the native signers. Finally, Best et al. (2010) tested deaf native and non-native signers, along with hearing L2 and naïve participants. They also found non-native signers to make more within-category discrimination compared to the native signers, but they did not directly compare the discrimination of the native signers with that of either the hearing L2 or naïve participants. The finding across these studies suggest that native signers, like native speakers, are less sensitive to within-category variation; non-native signers who learned ASL at later ages and naïve participants who have no knowledge of sign language tend to make more within category discriminations. The broad outline of these research trends suggests that early AoA is linked to the learning of phonetic categories in sign phonology in a fashion analogous to phonetic learning in early spoken language acquisition.

AoA effects on sign phonology extend to sign recognition. On a gated sign recognition task, deaf non-native signers required more phonetic information to recognize signs compared to native signers (Morford & Carlson 2011). Consistent with the hypothesis that early sign learning leads to phonetic categories, the non-natives' responses prior to accurate sign recognition were less phonologically organized in comparison to those of both native and hearing L2 signers. The finding that the hearing L2 signers patterned with the deaf native signers, rather than with the non-native signers despite learning to sign at older ages, provides support for the maturational hypothesis (Morford & Carlson 2011; see also Emmorey & Corina 1990). Accuracy and reaction time differences in favor of native signers over non-native signers have also been reported for sign tasks such as primed lexical decision with signs and pseudo-signs, and sign detection/spotting, in British and Spanish Sign Language (Dye & Shih 2006; Carreiras, Gutierrez-Sigut, Baquero & Corina 2008; Orfanidou, Adam, McQueen & Morgan 2009).

AoA effects on sign phonology have also been investigated with similarity judgment tasks. Hildebrandt & Corina (2002) asked native signers, non-native signers, and sign-naïve participants to indicate which of four pseudo-sign options was most phonologically similar to a given target pseudo-sign. Native and non-native signers showed subtle differences in both of their experiments. When the alternative options were comprised of minimal pairs, the non-native signers patterned with sign-naïve participants, rather than with native signers. However, when the alternative options were less strongly related to the target, each participant group behaved differently. Thus, the pattern is not entirely clear and calls for additional investigation.

To gain further insights into AoA effects on sign phonological processing, we conducted two experiments using a phonological similarity judgment task. We developed a novel experimental paradigm using a continuous measure of phonological similarity, which may be more sensitive than the forced-choice measures used in the past. To manipulate AoA, we tested signers with a range of ASL experience allowing us to determine whether the maturational or experience-based account best predicts AoA effects on the perception of sign phonology. Our hypothesis is that early language experience may help a learner decipher which categories are likely to be (un)important for language learning, even if the language belongs to a different modality. If so, we should expect deaf native and hearing L2 signers to pattern together because they both acquired language in infancy, albeit different languages in different modalities. We should also expect deaf non-native signers to pattern differently due to their more restricted language experience in infancy prior to their ASL acquisition at older ages. Therefore, Experiment 1 compares deaf native, deaf non-native, and hearing L2 signers on a novel phonological similarity judgment task. Experiment 2 tests sign-naïve participants on a modified version of the same task.

2. Experiment 1

2.1 Method

2.1.1 *Participants*

We tested 36 participants who fell into one of the following three categories. Deaf native signers (n = 10) were born profoundly deaf and exposed to sign language from birth. One native participant had a deaf older brother and parents who signed; the remainder had at least one deaf signing parent. Deaf non-native signers (n = 10) were prelingually deaf but exposed to sign language after infancy and had varied AoA. In order to focus on first-language acquisition effects, individuals who reported being skilled communicators in speech/lipreading or writing prior to ASL acquisition were not recruited. Thus, the non-native participants represent the population of signers who acquire ASL as a late **first** language. Hearing L2 signers (n = 16) were included if they had been using ASL for at least 5 years outside the classroom. They were strongly dominant in English, but fairly proficient in ASL. Participant characteristics are given in Table 1. Participants were paid for their participation.

Table 1. Participants' language background and demographics, based on self-report.

Population	**N**	**Mean Age**	**# Deaf parents**	**AoA (ASL)**	**Years ASL**	**Education Level** [a]
Deaf Native	9	30	1.56	0.67	29	4.0
Deaf Non-Native	9	36	0	8.13	28	2.89
Hearing L2	16	29	0	16.31	13	4.25

[a] Education level was operationalized as follows: 1 = high school or GED; 2 = entered college but no degree; 3 = Associate's degree; 4 = Bachelor's degree; 5 = entered grad school but no degree; 6 = Masters degree; 7 = doctorate or terminal degree.

2.1.2 *Materials*

The materials for this study were part of a larger study of syntactic priming in ASL (Hall, Ferreira & Mayberry 2010), which required all items to be pictureable objects. Due to this constraint, our goal was to find pairs of phonologically related signs without attempting to balance which sign parameters were shared. There were two sets of items composed of a set of 24 target pictures, and 48 ASL signs. Each target picture was paired with two signs: one that was related to the picture, and one that was not. The difference between the lists was that in the "phonological" set the target and the related sign were phonological minimal pairs and semantically unrelated (e.g., target picture — BIRD; related sign — NEWSPAPER; unrelated sign –ROCK). In the "semantic" set the target and the related sign were members of the same semantic category, but we intended both alternative signs to be **phonologically** unrelated (e.g., target picture — FLUTE; related sign — VIOLIN; unrelated sign — HOT-DOG). The phonologically unrelated pairs shared no more than 1 phonological feature (handshape, location, or movement). However, these may not be the only features to which all signers are sensitive; in fact, the need to empirically verify these pairings is what motivated the present study.

Although the stimulus sets contrasted in phonological and semantic relations, we only asked participants to make phonological judgments. We expected to find rating differences between the related and unrelated items in the phonological set, but not in the semantic set. English translations of the materials are provided in Appendix A.

2.1.3 *Design & procedure*

Participants rated phonological similarity as part of a larger study that lasted 90 minutes. Prior to this task, they performed a block of ASL picture naming, an ASL sign comprehension task, and a syntactic priming task that involved perceiving and producing simple noun phrases and shape descriptions. Data from those conditions are reported elsewhere (Hall, Ferreira & Mayberry 2010).

The phonological similarity task was introduced by a video in which a deaf native signer explained the task in ASL. The signer explained that the participants' task was first to look at a target picture, think of its corresponding sign, and then press the space bar. Upon the key press, two videos appeared on the screen in sequence, each one showing a different ASL sign. Neither video was the actual sign for the picture. Because naïve participants, like naïve speakers, would not understand what we meant by "phonological similarity", the video explained that the participants' task was to determine if either of the signs "looked like" or "felt like" the sign for the target picture. Although the task is ostensibly visual only, we included the "felt like" instruction because recent evidence suggests that somatosensory information may impact sign processing as much as or more than visual information for some tasks (Emmorey, Bosworth & Kraljic 2009; Emmorey, Korpics & Petronio 2009). To clarify what these instructions meant, the signer then provided an example where the target picture was milk, and the two signs were ORANGE (a minimal pair differing in location) and WATER (a semantically related sign). The signer then demonstrated how the signs for MILK and ORANGE "looked like" or "felt like" each other. The participants were instructed to ignore relationships based on meaning, such as MILK and WATER and how to make their response, which was to use a mouse to click inside a rating box that was on the laptop screen beneath the two lexical sign choices (see Figure 1). At the center of the box was an anchor point of "0". The numbers increased from 1 to 3 in both directions. The participants were told that if they thought that the sign on the right looked/felt more like the sign for the target picture than the one on the left, they should click somewhere to the right of 0 (and vice versa). The stronger the perceived similarity, the farther the participant should click in that direction. If the participant thought that both signs were equally related or equally unrelated to the sign for the target picture, they were instructed to click on or near 0. Participants

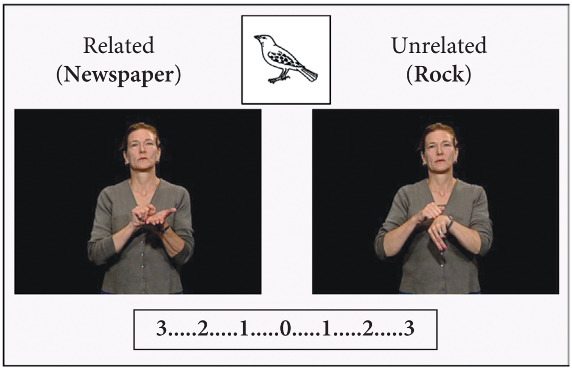

Figure 1. Schematic illustration of the paradigm. The trial pictured is from the "Phonological" set; participants click inside the rating box in the direction of the sign they rate as being more phonologically similar to the target sign. Here, participants are expected to click toward "NEWSPAPER". (English labels are for illustrative purposes only.)

were also told that the numbers 0 to 3 were only present as guides, and that they should feel free to click in between the numbers anywhere within the rating box.

After explaining how the rating box worked, the participants saw two more example trials: one resembling a trial from the phonological set and one resembling a trial from the semantic set; the example items did not appear in the experimental materials. If participants clicked in the expected range (in the direction of the related item for the phonological trial, and near 0 for the semantic trial), they were given positive feedback. If not, they were given differential feedback as a function of the kind of error they made. Finally, participants were told to ask the experimenter (the first author, who is a fluent signer) if they had any questions. When the participant understood the task to the experimenter's satisfaction, the 48 test trials began. Items were presented in a fixed random order, which was displayed to half the participants in reverse order. Half of the related items in each set were presented on the left and half on the right.

2.2 Results

The first question concerned how native-like the non-native and L2 signers would be. Therefore, we give the native signers' ratings first and then turn to those of the non-native and L2 signers. We used two measures to compare the groups, item-by-item correlations, and a continuous measure of mean relatedness.[1] For both measures, mouse clicks in the direction of the related sign (be it phonological or

1. This paradigm also allows analysis of categorical choices; in the present data, they yielded the same results as continuous measures and will not be discussed further.

semantic) were operationalized as the positive direction. The direction magnitude was measured in raw pixel counts to which we applied a linear transformation creating a scale from -100 to 100, where 100 represents clicking as far as possible in the direction of the related picture, and -100 represents clicking as far as possible in the direction of the unrelated picture. A value of 0 represents clicking near the 0 point, that is, a judgment that the two signs were equally related or equally unrelated.[2] Due to computer error, data from one native signer were not recorded. In addition, data from one non-native signer revealed a misunderstanding of the task, such that the phonological set received low ratings and the semantic set received high ratings. This participant's data were removed prior to the analyses, leaving data from 9 native, 9 non-native, and 16 L2 participants.

2.2.1 *Deaf native signers*

Before comparing the groups, a few comments about the native signers' ratings are in order. As shown by the variability along the y-axis of Figure 2, phonological similarity is not an all-or-none phenomenon. Although the phonological items (red dots) would traditionally be considered minimal pairs, the native signers' responses ranged from near the top of the scale (SOCK — STAR, native rating = 96.98) to the moderate/low end (UMBRELLA — WINDOW, native rating = 25.02). Likewise, although we did not intend there to be any phonological relatedness among the semantic items (blue dots), native signers nevertheless perceived some relatedness in some items, to the point that some semantic items (COFFEE — TEA, native rating = 51.58; CAKE — PIE, native rating = 30.87) received higher ratings than some phonological items. Nevertheless, the overwhelming tendency for the native signers was to rank phonological items more highly similar than the semantic items, suggesting that this novel paradigm is viable. The finding that native signers gave variable ratings suggests that they are not using top-down strategies or explicit knowledge about phonological features in ASL. If so, we would expect the results to look more categorical than the observed gradient data.

2.2.2 *Item-by-item correlations*

We computed the mean similarity rating for each item across each group. Here "item" refers to a given target picture and its two response options. The mean rating for each item by native signers is plotted on the y-axis of Figure 2, with the mean rating for the same item by non-native signers plotted on the x-axis of Figure 2a; the data from L2 signers are displayed in Figure 2b. Recall that positive values

2. Although it is not possible to distinguish "0" responses indicating "both signs are related to the target, and equally so" from those indicating "neither sign is at all related to the target", we are confident that given our design, "0" responses indicate that both signs are unrelated to the target.

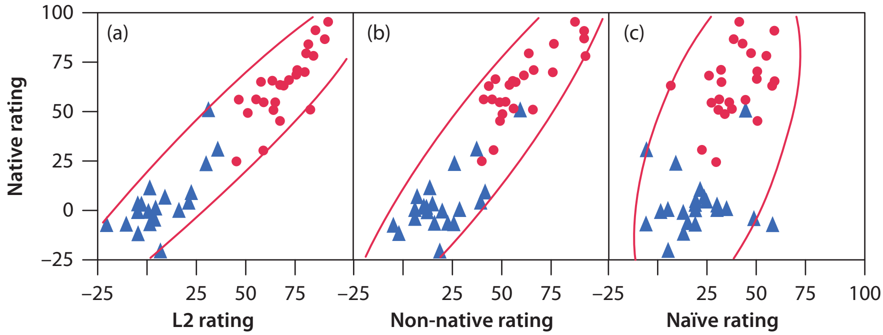

Figure 2. Comparison of mean ratings for each item from each population. Native-like performance would be a straight diagonal line. Red circles represent items from the phonological set; blue triangles represent items from the semantic set. Data from sign-naïve participants (Figure 2c) are discussed in Experiment 2 (Section 3).

indicate clicks in the direction of the related sign, whereas negative values indicate clicks in the direction of the unrelated sign, and the farther from 0, the stronger the perceived relationship. As is evident from the figures, both non-native and L2 signers generally approximate the native signers' ratings. Perfect native-like performance would be evidenced if each item fell exactly on the diagonal described by the line $y = x$, with $r^2 = 1$. Computing deviation from the diagonal is therefore a measure of deviation from native-like performance. We calculated the residuals for each item for each group (the square of the deviation between the native signers' mean rating and the other groups' mean rating) and then compared the size of these residuals with a paired t-test over items. This revealed that the L2 signers were significantly more native-like than the non-native signers [$t(47) = 2.42$, $p < .02$; see Figure 3, left two bars].

2.2.3 *Mean relatedness*

Next we asked whether non-native and L2 signers differed from native signers in the overall amount of similarity they observed in each stimulus set. These means are displayed in Figure 4. L2 signers did not significantly differ from native signers. By contrast, non-native signers reported more phonological similarity within the semantic set than did the natives. These results are obtained in two 2 x 3 mixed ANOVAs over subjects (F_1) and items (F_2), with Stimulus Set (Phonological or Semantic) as a within-subjects and between-items factor, and Group (native, non-native, or L2) as a between-subjects and within-items factor. Results revealed the predicted main effect of Stimulus Set [F_1 (1, 31) = 348.59, $p < .001$; F_2 (1,92) = 188.57, $p < .001$], a main effect of Group by items only [F_1 (2, 31) = .86, $p < .43$; F_2 (2,92) = 4.62, $p < .02$], and a Stimulus Set by Group interaction [F_1 (2, 31) = 5.22, $p < .02$; F_2 (2,92) = 29.29, $p < .001$]. Using planned comparisons to explore

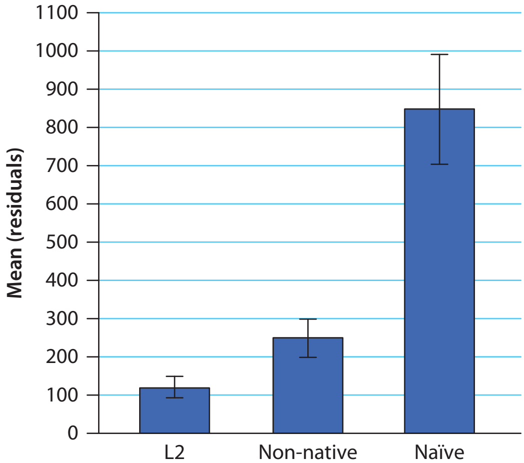

Figure 3. Mean residuals (squared deviation) between native signers and each other population. Lower values represent more native-like performance. Error bars represent standard error of the mean.

this interaction, we found that the L2 signers' ratings did not differ from those of the native signers for either the phonological set [F_1 (1,31) = .54, p = .47; F_2 (1,92) = 2.71, p = .10) or the semantic set [F_1 (1,31) = .10, p = .75; F_2 (1,92) = .50, p = .48)]. There was a trend for non-native signers to give lower ratings than the native signers for the phonological items, but only in the items analysis [F_1 (1, 31) = .82, p = .37; F_2(1,92) = 5.24, p < .05]. Compared to the native signers, the non-native signers gave significantly higher phonological similarity ratings to items in the semantic set [F_1 (1, 31) = 6.79, p < .02; F_2(1,92) = 43.23, p < .001]. As expected, all groups did give significantly higher ratings to items in the phonological set than in the semantic set (all F_1 & F_2 values > 56.0, all p values < .001).

2.3 Discussion

Examining the item-by-item correlations among the different groups revealed that, on the whole, both the deaf non-native and hearing L2 signers gave native-like ratings. However, between the two groups, the non-native signers were consistently less native-like than the L2 signers. When we analyzed participants' ratings as a function of stimulus set, a clear pattern emerged: non-native signers gave higher **phonological** similarity ratings than did the native signers to the sign option **semantically** related to the target picture. Before interpreting these results, it is important to consider alternative explanations for this pattern.

One possibility is that differences in the picture names could be driving the results. There is dialectal variability in sign, as in speech, and if the non-native signers simply used different signs from the native and L2 signers, we would expect their ratings to be different. (Recall that the task required participants to view a target picture and **to think of** their sign for that picture.) Fortunately, the first task that participants performed during the testing session was naming (in ASL) all 96 pictures that were in the study. We were therefore able to repeat the above analyses excluding trials where the participant initially produced anything other than the target sign. Those analyses yielded the same pattern of results; if anything, the difference between non-native and the native and L2 signers was magnified. Therefore, the above results may be conservative. We can thus be confident that the differences among the groups are not driven by dialectal variation.

Another plausible alternative is simply that, although the task was to consider only phonology, the non-native signers were nonetheless influenced by the presence of semantic relationships. If so, this might constitute evidence of early learning effects on inhibitory control, or late learning tendencies toward holistic sign processing perhaps akin to iconic gesture processing. It would not, however, provide unambiguous evidence of AoA effects on phonology itself. To address this question, we conducted experiment 2.

3. Experiment 2

In Experiment 1, deaf non-native signers gave unexpectedly high phonological similarity ratings to pairs of items that were chosen to be semantically and not phonologically related, whereas deaf native and hearing L2 signers did not. If this pattern is somehow driven by the semantic relationships of these items, then the effect should disappear in individuals who have no knowledge of sign language semantics. Alternatively, non-native signers may be attending to some visual property of the stimuli that native and L2 signers ignore. Indeed if some visual and/or somatosensory properties of the stimuli are salient, then it is possible that hearing naïve participants (who have perceptual access to sign language without knowledge of its phonological structure or semantics) might also be sensitive to it. Evidence for this explanation would come from hearing naïve participants showing the same pattern of results as the deaf non-native signers in Experiment 1, especially if they show the effect on the same items. Experiment 2 was a modified version of Experiment 1 with hearing sign-naïve participants.

3.1 Method

3.1.1 *Participants*

We tested 20 hearing participants who reported being monolingual native English speakers with no knowledge of ASL. Participants were undergraduate students at UC San Diego and received course credit for their participation.

3.1.2 *Materials*

Because the participants could not generate an ASL sign corresponding to the target picture, we replaced the target picture with its corresponding ASL sign. This was the only difference from Experiment one and allowed the task to be performed by the naïve participants.

3.1.3 *Design & procedure*

The procedure was identical to Experiment 1, except that the ASL instruction video was replaced by instructions in written English.

3.2 Results

As in Experiment 1, we performed both the item-by-item correlations as well as the overall mean relatedness judgments for each stimulus set.

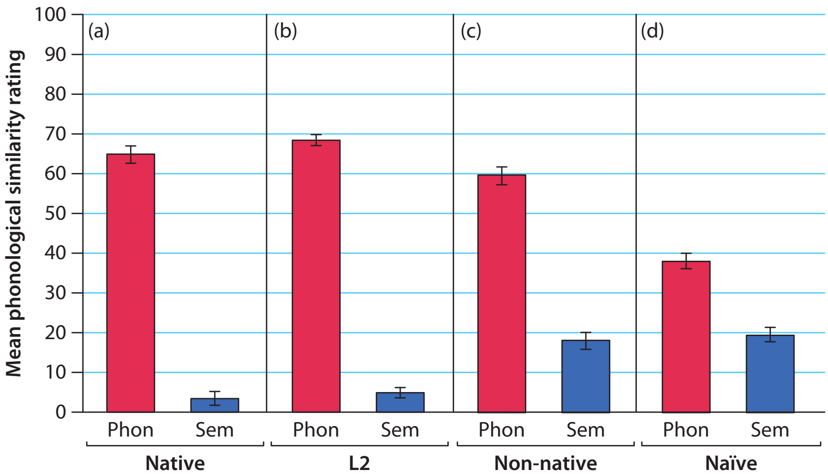

Figure 4. Mean ratings from each population by item set: phonological or semantic. Error bars represent standard error of the mean.

3.2.1 *Item-by-item correlations*

The x-axis of Figure 2c plots the mean rating given by the naïve participants for each item. Although perceptual information affords some general approximation of phonological relatedness in sign language, it is clear that the hearing naïve group is far less native-like than either the deaf non-native or hearing L2 groups. This is confirmed by comparing the residuals of the naïve group against those of the non-native and L2 groups. A one-way repeated measures ANOVA with item as a random effect was significant [$F(1,94) = 18.02$, $p < .001$]. Planned contrasts revealed that the naïve group differed from the non-native group [$F(1,94) = 21.44$, $p < .001$] and from the L2 group: [$F(1,94) = 31.63$, $p < .001$], as shown in Figure 3.

3.2.2 *Mean relatedness judgments*

The crucial question of Experiment 2 was whether the naïve participants would give higher-than-expected ratings of **phonological** similarity to items that were **semantically** related. As shown in Figure 4d, this was in fact the case. A 2 x 2 mixed ANOVA with Stimulus Set (Phonological or Semantic) as a within-subjects factor, and Group (native or naïve) as a between-subjects factor yielded the expected main effect of Stimulus Set [$F_1(1, 27) = 67.03$, $p < .001$; $F_2(1,46) = 127.99$, $p < .001$], and no main effect of Group [F_1 $(1, 27) = 2.38$, $p = .13$; F_2 $(1,46) = 3.97$, $p = .05$]. Importantly, there was a significant Stimulus Set x Group interaction [F_1 $(1, 27) = 19.21$, $p < .001$; F_2 $(1,46) = 59.41$, $p < .001$]. As expected, both groups gave significantly higher ratings to items in the phonological set than in the semantic set (all F_1 & F_2 values > 11.0, all p values < .005).

Planned contrasts revealed that the naïve participants showed two response proclivities: they gave lower similarity ratings than the native signers to the phonological items [F_1 $(1, 27) = 15.21$, $p < .001$; F_2 $(1,46) = 47.04$, $p < .001$]. Importantly for the present question, they also gave higher similarity ratings than the native signers to the semantic items [F_1 $(1,27) = 5.29$, $p < .03$; F_2 $(1,46) = 16.33$, $p < .001$]. Although both of these patterns were also observed in the non-native signers, we focus mainly on this latter effect in the discussion below.

3.2.3 *Are sign semantics transparent?*

Our rationale for testing hearing naïve participants was based on the assumption that they would be insensitive to the meaning of the signs. However, some of the pairs in the semantic set were iconic (e.g. VIOLIN — FLUTE). Is it possible that the naïve participants were sensitive to these items such that their increased similarity ratings reflect a semantic influence, rather than sensitivity to some visual-perceptual factor? If so, then the higher ratings in the semantic set should be driven by trials where the semantic relationship is guessable by non-signers. To test this, we recruited two more groups of naïve participants and asked them to

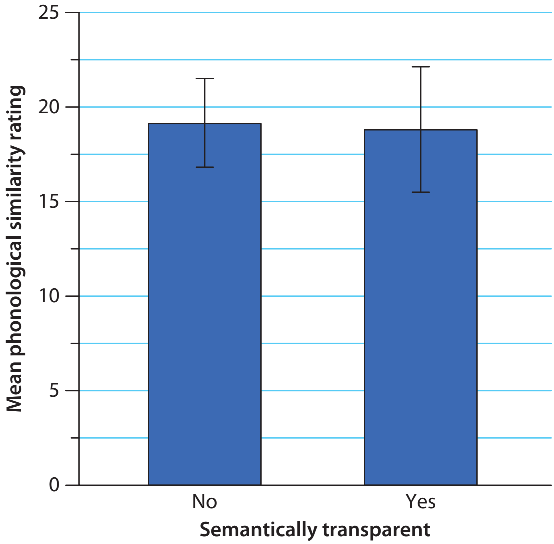

Figure 5. Mean rating from sign-naïve participants for items in the Semantic set as a function of whether or not the semantic relationship between the target and related sign was transparent. Error bars represent standard error of the mean.

guess the meanings of the ASL stimuli. For one group, each sign was paired with four possible pictures and they had to select which one they thought was the correct meaning. The other group saw a sign and were instructed to guess its meaning without any other cues. Performance across these two tasks was used to generate a composite score reflecting the iconicity of the items. We then re-categorized semantic trials based on whether the semantic relationship could be inferred by non-signers. Trials where such a relationship could be easily inferred were considered "transparent"; trials where the relationship could not be easily inferred were considered "non-transparent". If the higher overall ratings in our naïve data were driven by semantics, then phonological similarity ratings should be higher for the transparent trials, and lower for the non-transparent trials. However, as shown in Figure 5, this was not the case; participants' ratings of **phonological** similarity were the same regardless of whether the semantic relationship was evident [$t(19) = .14$, $p = .89$]. In other words, semantic transparency cannot explain why even sign-naïve participants gave higher-than-expected ratings of **phonological** similarity to items that were **semantically** related, whether they knew it or not.

3.2.4 *Are the same items prompting similarity ratings from non-native and naïve participants?*

If latent influence from semantic relatedness is not the source of this effect, what is? One possibility is that, by chance, the semantically related pairs share some visual-perceptual property that native and L2 signers ignore, but which non-native signers and naïve participants are sensitive to. This account would be supported if the non-native signers and naïve participants gave higher-than-native ratings to the same semantic items, and if both the non-native signers and naïve participants gave lower-than-native ratings to items in the phonological set. If the phonological systems of non-native signers and naïve participants deviate from the phonological systems of native signers in similar ways, then there should be a correlation in the size and direction of these deviations for each item and the correlation should be strongest in these groups. By contrast, the L2 signers' deviations from native-like performance should correlate weakly or not at all with the deviations for the non-native and naïve groups.

For example, consider the item set where the target is VIOLIN, and the options are FLUTE (semantically related) and HOT-DOG (semantically unrelated). If the native signers behaved as expected, their mean **phonological** similarity rating for this item should be near 0. To continue the example, suppose that the native signers' normalized mean response for this item was 5 (indicating that as a group, they tended to click on a pixel that was 5% of the maximum possible distance in the direction of the related sign). However, if VIOLIN and FLUTE share some visual property to which the non-native signers and naïve participants are sensitive, then their responses to this item should deviate in the positive direction by a similar amount. Thus, if we found that the non-native signers gave a mean rating of 11 and the naïve participants gave a mean rating of 13, we would plot positive 6 for the non-native (11–5 = 6) and positive 8 for the naïve participants (13–5 = 8). Performing this calculation across all items could reveal significant item-by-item correlations between the ratings of the non-native and L2 signers, but only if these participants are sensitive to the same visual perceptual properties for a given item.

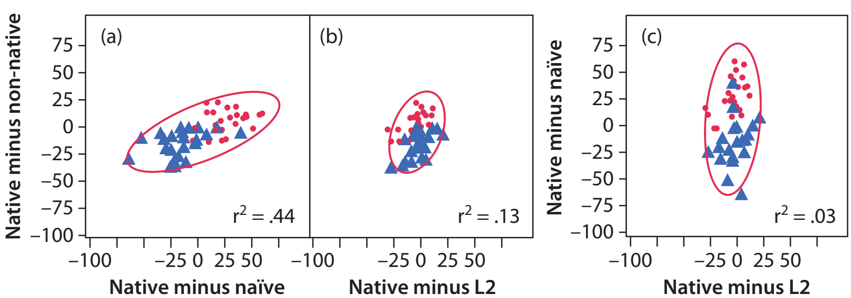

Figure 6. Correlations of difference scores. Strong correlations indicate that different populations deviate from native signers in similar ways, suggesting that their phonological systems are sensitive to the same kinds of information. Weak or absent correlations suggest that deviations from native-like performance are not systematic.

As predicted, there was a strong and significant correlation between the non-native signers' deviations from native-like ratings and the naïve participants' deviations ($r^2 = .44$, $p < .001$). The same comparison between the non-native and L2 signers revealed a much weaker but still significant correlation ($r^2 = .14$, $p < .01$). Consistent with the results from Experiment 1, there was no significant correlation between the ratings of the L2 signers and naïve participants ($r^2 = .03$, $p = .26$). These correlations are shown in Figure 6.

3.3 Discussion

If the naïve participants had given low similarity ratings to the semantic set items, we could have attributed the ratings of the non-native signers to a failure to ignore semantic features in a phonological task. However, because the naïve participants also gave systematically higher-than-native ratings to items in the semantic set, and because these tended to be the same items, we tentatively conclude that there are some, to-be-identified, visual-kinesthetic properties of the items that both the non-native signers and naïve participants were sensitive to, but the native and L2 signers were not. These results also suggest that there were phonological features that the native and hearing L2 signers were sensitive to, but the non-native and naïve participants were not.

4. General discussion

We used a novel paradigm to investigate AoA effects on phonological similarity judgments in ASL. This paradigm asks participants to choose which of two lexical signs is more phonologically similar to a target, and asks them to indicate the perceived strength of this relatedness. One finding resulting from this novel measure was that the mean item ratings fell along the entire range from 0 to 100. This suggests, unsurprisingly, that not all minimal pairs are created equal. Although this paradigm is amenable to investigating fine details of which feature bundles are perceived as being more similar than others, this was not our main focus here. Our goal was to ascertain whether and how AoA affects phonological similarity judgments of ASL signs and which of two accounts of AoA effects on phonological processing best predicted our results. To this end, we tested deaf native signers, who learned ASL from birth, deaf non-native signers whose language experience in infancy was restricted and who learned ASL at older ages, hearing L2 signers who had unrestricted language experience in infancy and learned ASL at older ages, and hearing naïve participants who had no prior knowledge of ASL.

Using the native signers' ratings as the baseline, our analysis of item-by-item correlations revealed that the most native-like performance on this task came from the L2 signers, with non-native signers coming in a close second. The naïve participants, though somewhat successful, were less native-like than either of the other two groups (Figure 3). When examining mean ratings from all the groups as a function of whether the items came from the phonological set or the semantic set, a similar pattern emerged. We found no differences between the native and L2 signers, but we did find differences between the native and non-native signers. Specifically, the non-native signers gave slightly lower ratings to items in the phonological set than did the native signers. The non-native signers also gave markedly higher similarity ratings to items in the semantic set than did the native signers, implying that they perceived the semantically related items as more phonologically similar to the target than the unrelated alternative items. We ruled out the possibility that dialectical variation among the signers was driving these results. In Experiment 2, we demonstrated that semantic relatedness is also unlikely to be driving these results because participants naïve to ASL also gave high similarity ratings to these items even when the semantic relationship could not be known (Figure 5). Thus, we found evidence consistent with the explanation that the non-native signers and naïve participants were sensitive to certain phonetic, visual, or somatosesory features that both the native and L2 signers were not sensitive to. This was shown by the finding that the non-native signers and naïve participants tended to give higher-than-native ratings to the same items to a larger extent than did the L2 signers (Figure 6).

The exact phonetic, visual, or somatosensory features that are causing these differences remain elusive. Table 2 presents the items where native and non-native signers showed the greatest discrepancy in their ratings, in both directions. The items where non-native signers indicated substantially **more** phonological similarity than natives are presented in Table 2A. The majority of these items (7/10) do not share any formal sign parameters (handshape, location, movement). The remaining 3 include pairs with shared location (TOAST — BREAD), handshape (BEAR — WOLF), and movement (POTATO — ROCK, which was from the phonological set). These were distributed throughout the top 10 with no evident trend. Classifying the pairs in Table 2A according to Battison's (1978) sign types provides little additional insight. Although 3 of the top 10 signs did belong to the same "type" category (CANDLE — MATCH: Type 4; FLUTE — VIOLIN: Type 2; FOOTBALL — BAT: Type 2), the remaining 7 were mismatches of diverse types. Moreover, these results are very similar in Table 2B, which lists the items where non-native signers indicated substantially **less** phonological similarity than natives (4 type matches, 6 type mismatches). Thus, the current data do not suggest any obvious candidate features that might be driving the differences we observed. Fortunately, the novel paradigm we have presented here may be easily used to explore these questions using more controlled stimuli. For example, using carefully controlled non-signs (as in Hildebrandt & Corina 2002) would remove semantics altogether while allowing parametric variation in sign features. Unlike Hildebrandt and Corina (2002), the present paradigm yields more fine-grained information about perceived **degree** of similarity. This greater sensitivity will be useful in tracking down subtle effects like those reported here.

Table 2. The below tables present the items that showed the greatest discrepancies between native and non-native signers. Table 2A shows the items where non-native signers indicated **more** phonological similarity than natives did. Table 2B shows the items were non-natives signers indicated **less** phonological similarity than natives did.

Table 2a.

Target Item	Related Item	Stimulus Class	Native minus Non-native
PRIEST	NUN	Semantic	-16.57
TOAST	BREAD	Semantic	-15.29
CANDLE	MATCH	Semantic	-14.21
BROOM	VACUUM	Semantic	-13.52
RING	NECKLACE	Semantic	-12.55
BEAR	WOLF	Semantic	-12.39
FLUTE	VIOLIN	Semantic	-9.75
FOOTBALL	BAT	Semantic	-8.16
WITCH	GHOST	Semantic	-7.86
POTATO	ROCK	Phonological	-6.59

Table 2b.

Target Item	Related Item	Stimulus Class	Native minus Non-native
EAGLE	DOLL	Phonological	10.23
KEY	ONION	Phonological	10.03
RAINBOW	FENCE	Phonological	8.06
DOCTOR	SOAP	Phonological	7.79
COMB	RAKE	Phonological	5.69
EGG	KNIFE	Phonological	5.46
BIRD	NEWSPAPER	Phonological	5.28
LAWNMOWER	MOTORCYCLE	Phonological	4.77
SOCK	STAR	Phonological	4.74
SCISSORS	LOBSTER	Phonological	4.71

One consistent finding across previous research investigating categorical perception in sign is that native signers are the least sensitive to within-category phonetic variation, while naïve participants are the most sensitive to it. Baker et al. (2005), Best et al. (2010), and Morford et al. (2008) all found that sign-naïve participants were more likely than other participants to discriminate between two stimuli that both came from the same side of a category boundary (e.g. two different '5' handshapes). These results accord well with our explanation that the naïve participants were sensitive to perceptually-based properties of the present stimuli to which the native signers were not. And, like Morford et al. (2008), we found that non-native signers patterned with the naïve participants; however, whereas they found that L2 signers discriminated more than native signers in their discrimination task, we found that native and L2 signers patterned together in our phonological similarity judgment task.

These results have implications for psycholinguistic theories of phonetic processing, and competing accounts for AoA effects on language development. Experience-based accounts of AoA effects predict that a delayed language onset should not impair successful acquisition, given enough input and practice. By contrast, maturation-based accounts predict that language delay will disrupt subsequent language acquisition, including phonology. The present results, in concert with those reviewed above, are best predicted by maturation-based accounts of AoA effects. It appears that even decades of experience with one's dominant language do not fully compensate for a lack of language input during infancy and early childhood. Furthermore, the present data suggest that **having** language exposure from birth can facilitate native-like acquisition of a second phonological system by L2 learners. Further work is necessary to determine whether the same would hold true for deaf native signers who are acquiring a second sign language, or whether the first and second language must belong to separate modalities. Addressing that question would inform whether the limits on L2 phonological processing are due to domain-general constraints on language learning, or to direct competition from L1. This, in turn, could inform current debates about whether having sign language skills is likely to be helpful, harmful, or neutral for deaf children who are attempting to acquire a spoken phonology through cochlear implants. Finally, additional research is required to identify the precise nature of these age of acquisition effects on phonological development and phonetic processing both within and across modalities.

5. Conclusions

We have presented results from a novel paradigm for collecting phonological similarity judgments in ASL. We found that, although ratings from hearing L2 signers and from deaf non-native signers closely approximated ratings from deaf native signers, there were revealing differences between native and non-native signers. Specifically, non-native signers saw phonological similarity among the semantic items, whereas the native signers either did not perceive it, or judged any perceived similarity to be irrelevant to the task. We found the same pattern among naïve participants, and it was equally strong regardless of whether or not the items were semantically transparent. We propose that, consistent with observations from studies of categorical perception in sign language, non-native signers and naïve participants are sensitive to as-yet-unidentified visual or somatosensory properties of signs that native and L2 signers overlook. Early language experience helps learners discern which linguistic categories are important for language learning, and the product of this early language learning appears to transcend sensory-motor modality.

Acknowledgments

We gratefully acknowledge Marla Hatrak, Brittany Hernandez, and Kristyn Claypool for their help in the design and coding of the study, as well as Deaf Community Services of San Diego for providing space for testing. We are also indebted to many members of San Diego's Deaf community for their participation. Funding for this research was provided by NIH grant HD051030, NIH grant 5T32DC000041–19, and by the Rita L. Atkinson Graduate Fellowship.

References

Baker, Stephanie A., William J. Idsardi, Roberta Michnick Golinkoff & Laura-Ann Petitto. 2005. The perception of handshapes in American Sign Language. *Memory & Cognition* 33. 887–904.

Battison, Robin. 1978. *Lexical borrowing in American Sign Language.* Silver Spring, MD: Linstok Press.

Best, Catherine T., Gaurav Mathur, Karen A. Miranda & Diane Lillo-Martin. 2010. Effects of sign language experience on categorical perception of dynamic ASL pseudosigns. *Attention, Perception, and Psychophysics* 72. 747–762.

Brentari, Diane. 1998. *A prosodic model of sign language phonology.* Cambridge, MA: MIT Press.

Carreiras, Manuel, Eva Gutiérrez-Sigut, Silvia Baquero & David Corina. 2008. Lexical processing in Spanish Sign Language (LSE). *Journal of Memory and Language* 58. 100–122.

Corina, David P. & Ursula K. Hildebrandt. 2002. Psycholinguistic investigations of phonological structure in ASL. In Richard P. Meier, Kearsy Cormier & David Quinto-Pozos (eds.), *Modality and structure in signed and spoken languages*, 88–110. Cambridge: Cambridge University Press.

Dye, Matthew W.G. & Shui-I Shih. 2006. Phonological priming in British Sign Language. In Louis M. Goldstein, Douglas H. Whalen & Catherine T. Best (eds.), *Laboratory Phonology* 8, 241–264. Berlin: Mouton de Gruyter.

Emmorey, Karen & David Corina. 1990. Lexical recognition in American Sign Language: Effects of phonetic structure and morphology. *Perceptual and Motor Skills* 71. 1227–1252.

Emmorey, Karen, Stephen McCullough & Diane Brentari. 2003. Categorical perception in American Sign Language. *Language and Cognitive Processes* 18. 21–45.

Emmorey, Karen, Franco Korpics & Karen Petronio. 2009. The use of visual feedback during signing: Evidence from signers with impaired vision. *Journal of Deaf Studies and Deaf Education* 14. 99–104.

Emmorey, Karen, Rain Bosworth & Tanya Kraljic. 2009. Visual feedback and self-monitoring of sign language. *Journal of Memory and Language* 61. 398–411.

Flege, James Emil. 1999. Age of learning and second-language speech. In David Birdsong (ed.), *Second language acquisition and the critical period hypothesis*, 101–131. Mahwah, NJ: Lawrence Erlbaum.

Flege, James, Emil, Grace H. Yeni-Komshian & Serena Liu. 1999. Age constraints on second language acquisition. *Journal of Memory and Language* 41. 78–104.

Frishberg, Nancy. 1975. Arbitrariness and iconicity: Historical change in American Sign Language. *Language* 51. 696–719.

Hall, Matthew L., Victor S. Ferreira & Rachel I. Mayberry. 2010. Syntactic priming in American Sign Language. Talk presented at the *10th Theoretical Issues in Sign Language Research Conference (TISLR)*, October 2010.

Hildebrandt, Ursula & David Corina. 2002. Phonological similarity in American Sign Language. *Language and Cognitive Processes* 17. 593–612.

Lenneberg, Eric. 1962. *Biological foundations of language.* New York: Wiley.

Liddell, Scott K. & Robert E. Johnson. 1989. American Sign Language: The phonological base. *Sign Language Studies* 64. 195–277.

Klima, Edward & Ursula Bellugi. 1979. *The signs of language.* Cambridge, MA: Harvard University Press.

Mayberry, Rachel I. 1993. First-language acquisition after childhood differs from second-language acquisition: the case of American Sign Language. *Journal of Speech and Hearing Research* 36. 1258–1270.

Mayberry, Rachel I. 2007. When timing is everything: Age of first-language acquisition effects on second-language learning. *Applied Psycholinguistics* 28. 537–549.

Mayberry, Rachel I. 2010. Early language acquisition and adult language ability: What sign language reveals about the critical period for language. In Marc Marschark & Patricia E. Spencer (eds.), *The Oxford handbook of Deaf studies, language, and education, Volume 2*, 281–291. Oxford: Oxford University Press.

Mayberry, Rachel I. & Ellen B. Eichen. 1991. The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language* 30. 486–512.

Mayberry, Rachel I., Elizabeth Lock & Hena Kazmi. 2002. Linguistic ability and early language exposure. *Nature* 417. 38.

McCandliss, Bruce D., Julie A. Fiez, Athanassios Protopapas, Mary Conway & James L. McClelland. 2002. Success and failure in teaching the [r]–[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, & Behavioral Neuroscience* 2. 89–108.

Morford, Jill P. & Martina L. Carlson. 2011. Sign perception and recognition in non-native signers of ASL. *Language Learning and Development* 7. 149–168.

Morford, Jill P., Angus B. Grieve-Smith, James MacFarlane, Joshua Staley & Gabriel Waters. 2008. Effects of sign language experience on the perception of American Sign Language. *Cognition* 109. 41–53.

Munakata, Yuko & Jason Pfaffly. 2004. Hebbian learning and development. *Developmental Science* 7. 141–148.

Newkirk, Don, Edward S. Klima, Carlene C. Pedersen & Ursula Bellugi. 1980. Linguistic evidence from slips of the hand. In Victoria A. Fromkin (ed.), *Errors in linguistic performance: slips of the tongue, ear, pen, and hand*, 165–197. New York: Academic Press.

Newport, Elissa L. 1982. Task specificity in language learning? Evidence from speech perception and American Sign Language. In Eric Wanner & Lila R. Gleitman (eds.), *Language acquisition: The state of the art*, 450–520. New York: Cambridge University Press.

Newport, Elissa L. 1990. Maturational constraints on language learning. *Cognitive Science* 14. 11–28.

Orfanidou, Eleni, Robert Adam, James M. McQueen & Gary Morgan. 2009. Making sense of nonsense in British Sign Language (BSL): The contribution of different phonological parameters to sign recognition. *Memory & Cognition* 37. 302–315.

Penfield, Wilder & Lamar Roberts. 1959. *Speech and brain mechanisms*. New York: Atheneum.

Perlmutter, David M. 1992. Sonority and syllable structure in American Sign Language. *Linguistic Inquiry* 23. 407–442.

Sandler, Wendy. 1989. *Phonological representation of the sign: Linearity and nonlinearity in American Sign Language*. Dordrecht: Foris.

Sandler, Wendy & Diane Lillo-Martin. 2006. *Sign language and linguistic universals.* Cambridge: Cambridge University Press.

Seidenberg, Mark S. & Jason D. Zevin. 2006. Connectionist models in developmental cognitive neuroscience: Critical periods and the paradox of success. In Yuko Manakata & Mark Johnson (eds.), *Processes of change in brain and cognitive development — attention & performance XXI*, 585–612. Oxford: Oxford University Press.

Stokoe, William C. 1960. Sign language structure: An outline of the visual communication systems of the American Deaf. *Studies in Linguistics Occasional Papers* 8. Buffalo: University of Buffalo Press [Re-issued 2005, *Journal of Deaf Studies and Deaf Education* 10(1), 3–37].

Stokoe, William C. 1978. *Sign language structure: the first linguistic analysis of American Sign Language.* Silver Spring, MD: Linstok Press.

Thompson, Robin L., David P. Vinson & Gabriella Vigliocco. 2010. The link between form and meaning in British Sign Language: Evidence of iconicity for phonological decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 36. 1017–1027.

Thompson, Robin, Karen Emmorey & Tamar H. Gollan. 2005. "Tip of the fingers" experiences by Deaf signers: Insights into the organization of a sign-based lexicon. *Psychological Science* 16. 858–860.

Werker, Janet F. 1989. Becoming a native listener. *American Scientist* 77. 54–59.

Wilson, Margaret & Karen Emmorey. 1997. A visuospatial "phonological loop" in working memory: Evidence from American Sign Language. *Memory & Cognition* 25. 313–320.

Wilson, Margaret & Karen Emmorey. 1998. A "word length effect" for sign language: Further evidence for the role of sign language in structuring working memory. *Memory & Cognition* 26. 584–590.

Appendix A.

Stimulus Class	Target	Related	Unrelated
Phonological	BIRD	NEWSPAPER	ROCK
Phonological	BOOK	BOAT	BUTTER
Phonological	CHAIR	TRAIN	BICYCLE
Phonological	CHEESE	PAPER	SOAP
Phonological	CHURCH	CHOCOLATE	WINDOW
Phonological	COMB	RAKE	MOON
Phonological	DOCTOR	SOAP	CHOCOLATE
Phonological	EAGLE	DOLL	TYPEWRITER
Phonological	EGG	KNIFE	FLAG
Phonological	FISH	FLAG	KNIFE
Phonological	GLASSES	MOON	RAKE
Phonological	KEY	ONION	LOBSTER
Phonological	LAWNMOWER	MOTORCYCLE	ONION
Phonological	NOSE	MOUSE	STAR
Phonological	NURSE	BUTTER	BOAT
Phonological	OWL	BINOCULARS	SHOWER
Phonological	PIANO	TYPEWRITER	MOUSE
Phonological	POTATO	ROCK	NEWSPAPER
Phonological	RAINBOW	FENCE	MOTORCYCLE
Phonological	SCISSORS	LOBSTER	FENCE
Phonological	SHOE	BICYCLE	TRAIN
Phonological	SOCK	STAR	BINOCULARS
Phonological	SUN	SHOWER	DOLL
Phonological	UMBRELLA	WINDOW	PAPER
Semantic	AIRPLANE	HELICOPTER	BAT (BASEBALL)
Semantic	BEAR	WOLF	LETTUCE
Semantic	BROOM	VACCUUM	GHOST
Semantic	BUTTERFLY	BUG	SNAKE
Semantic	CAKE	PIE	MATCH
Semantic	CANDLE	MATCH	PIE
Semantic	COFFEE	TEA	BUG
Semantic	COW	PIG	BOY
Semantic	FLUTE	VIOLIN	HOTDOG
Semantic	FOOTBALL	BAT (BASEBALL)	FORK
Semantic	GIRAFFE	ELEPHANT	VACCUUM
Semantic	GIRL	BOY	PIG
Semantic	HAMBURGER	HOTDOG	VIOLIN
Semantic	HAMMER	SCREWDRIVER	COAT

Semantic	HAT	COAT	SCREWDRIVER
Semantic	PRIEST	NUN	TEA
Semantic	RING	NECKLACE	FROG
Semantic	SPIDER	SNAKE	NUN
Semantic	SPOON	FORK	ELEPHANT
Semantic	TOAST	BREAD	FLOWER
Semantic	TOMATO	LETTUCE	WOLF
Semantic	TREE	FLOWER	BREAD
Semantic	TURTLE	FROG	NECKLACE
Semantic	WITCH	GHOST	HELICOPTER

Corresponding author's address

Matthew Hall
Department of Psychology, Mail Code 0109
University of California, San Diego
9500 Gilman Dr.
La Jolla, CA 92093-0109
U.S.A.

matthall@ucsd.edu