Iconicity and Gesture Jointly Facilitate Learning of Second Language Signs at First Exposure in Hearing Nonsigners

Karadöller, Dilay Z.; Peeters, David; Manhardt, Francie; Özyürek, Aslı; Ortega, Gerardo

DOI:
10.1111/lang.12636

LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

EMPIRICAL STUDY

Iconicity and Gesture Jointly Facilitate Learning of Second Language Signs at First Exposure in Hearing Nonsigners

Dilay Z. Karadöller,[a,b] David Peeters,[c] Francie Manhardt,[a,d] Aslı Özyürek,[a,d] and Gerardo Ortega[e]

[a]Max Planck Institute for Psycholinguistics [b]Middle East Technical University [c]Tilburg University [d]Radboud University [e]University of Birmingham

Abstract: When learning spoken second language (L2), words overlapping in form and meaning with one's native language (L1) help break into the new language. When nonsigning speakers learn a sign language as L2, such overlaps are absent because of the modality differences (L1: speech, L2: sign). In such cases, nonsigning speakers might use iconic form-meaning mappings in signs or their own gestural experience as gateways into the to-be-acquired sign language. In this study, we investigated how both these phenomena may contribute jointly to the acquisition of sign language vocabulary

---

CRediT author statement: **Dilay Z. Karadöller**: conceptualization; methodology; investigation; data curation; formal analysis; writing–original draft preparation, review, and editing. **David Peeters**: conceptualization; methodology; software; writing–review and editing. **Francie Manhardt**: investigation; data curation; formal analysis; visualization. **Aslı Özyürek**: conceptualization; methodology; funding acquisition; supervision; writing–review, and editing. **Gerardo Ortega**: conceptualization; methodology; data curation; funding acquisition; supervision; writing–original draft preparation, review, and editing.

A one-page Accessible Summary of this article in nontechnical language is freely available in the Supporting Information online and at https://oasis-database.org

Correspondence concerning this article should be addressed to Dilay Z. Karadöller, Middle East Technical University, Dumlupınar Bulvarı 1, Beşeri Bilimler Building B047, dilayk@metu.edu.tr & Gerardo Ortega, University of Birmingham, Frankland Building, Birmingham, UK. Email: g.ortega@bham.ac.uk

Author Twitter information: *@DilayKaradoller @Peeters_David @franciemanhardt @ozyurek_a @ortega_ger*

The handling editor for this manuscript was Aline Godfroid.

DOI: 10.1111/lang.12636

by hearing nonsigners. Participants were presented with three types of signs in the Sign Language of the Netherlands (NGT): arbitrary signs, iconic signs with high or low gesture overlap. Signs that were both iconic and highly overlapping with gestures boosted learning most at first exposure, and this effect remained the day after. Findings highlight the influence of modality-specific attributes supporting the acquisition of a signed lexicon.

Keywords sign language; gesture; iconicity; L2 acquisition

Introduction

Vocabulary acquisition in spoken second language (L2) requires learners to produce novel lexical units using the sounds of the target language (Escudero & Boersma, 2004; Nation, 2001). Research in the field of L2 acquisition agrees that some important variables that influence L2 vocabulary learning are the phonological repertoire of learners' first language (L1; Jiang, 2004; Schwartz & Sprouse, 1996), the conventions of the L2 writing system (Bassetti, 2008), and the typological similarities between L1 and L2 (Tolentino & Tokowicz, 2014). However, these variables may have limited impact in vocabulary acquisition in learners whose L1 is a spoken language and who go on to acquire a sign language as their L2. What are the variables that modulate L2 sign learning in hearing adults? Given the differences in modality between their spoken L1 (oral-aural) and the signed L2 (manual-visual), it would be tempting to believe that hearing L2 sign learners cannot rely on any existing resources in their L1 that may assist vocabulary learning. However, recent evidence shows that despite these differences, L2 sign learners can rely on at least two important resources during early stages of learning.

The first resource is the perceived iconicity in the to-be-acquired signs. Iconicity has been defined as the perceived "resemblance-based mapping between aspects of form and meaning" (Dingemanse et al., 2020, p. 2; see also Dingemanse et al., 2015; Ortega, 2017; Perniss et al., 2010). It is a prevalent feature of all sign languages (Pietrandrea, 2002; Taub, 2001), and its presence has been shown to facilitate sign learning in hearing adults (Baus et al., 2012; Campbell et al., 1992; Hofweber et al., 2022, 2023). The second resource is gesture. Recent studies have shown that learners' gestural repertoire influences L2 sign learning at the lexical and morphological levels at the earliest stages (Chen Pichler, 2011; Janke & Marshall, 2017; Marshall & Morgan, 2015; Ortega et al., 2019, 2020). Although it is fairly well-documented how iconicity and gesture contribute individually to hearing adults learning signs, it has not yet been explored in the same experimental paradigm how these two resources may interplay and modulate the acquisition of signs' phonological constituents (e.g., handshape, location, and movement) in lexical forms during sign language learning. In order to fill this gap, the current study explores how hearing nonsigners' existing gestural repertoire and the perceived iconic link between form and meaning in the signs may contribute jointly towards the acquisition of a signed lexicon and their phonological components in nonsigning hearing adults.

Background Literature

Sign Languages and Lexical Development in Acquiring a Sign Language as L2

The sign languages of deaf communities are linguistic systems that display the same levels of organisation as spoken languages (Sandler & Lillo-Martin, 2006). Despite their organisational similarities with speech, sign languages present unique features that could make their acquisition significantly different from traditional spoken L2 learning. Spoken face-to-face communication partially relies on the transmission of visual signals (e.g., lip movements, facial expressions, cospeech hand gestures) that accompany concurrent speech; in sign language, however, there is an overall prevalence of the visible bodily articulators for linguistic expression. Unlike L2 learning of spoken languages where speech and text are the common medium of instruction, signers use their visible bodily articulators (hands, arms, lips, eyes, face, and torso) to express lexical and grammatical structures on the signer's body or on the physical space between interlocutors. How hearing nonsigners master the use of their visible bodily articulators to express lexical structures in visual space is a challenge that has only recently been investigated in the emerging field of L2 sign acquisition (see Schönström, 2021, for a review). Some of the studies in this domain have investigated the acquisition of mouthings (Mesch & Schönström, 2021), narrative coherence (Frederiksen & Mayberry, 2018), and perspective-taking (Gulamani et al., 2020). However, the area that has received the most attention is how hearing sign language learners acquire lexical signs on the basis of the signs' constituents.

Similar to spoken languages, sign languages indeed exhibit the property of phonology whereby seemingly meaningless building blocks (i.e., handshape, location, and movement) combine together to create meaningful lexical signs (Brentari, 1999; van der Kooij, 2002). These phonological constituents are important as the substitution of any of these may result in a completely different sign. For example, in British Sign Language (BSL) the sign for "insurance" consists of a closed fist with an extended pinky finger that moves from left to right of the chest (Vinson et al., 2008). Its minimal pair, the sign for

Figure 1 Examples of the signs "insurance" and "morning" in BSL (snapshots from videos of signs obtained from Vinson et al., 2008).

"morning," consists of the same location, movement, and orientation but differs in the configuration of the hand (see Figure 1). It is therefore critical that learners pay close attention to signs' constituents to acquire a manual lexicon.

A scant number of studies have explored the acquisition of a manual phonological system in hearing adults learning a sign language as a second language, at both the beginner and intermediate levels. In one earlier study (Bochner et al., 2011), hearing participants were shown two consecutive sentences in American Sign Language (ASL), and their task was to determine whether the sentences were the same or different. In filler trials, the sentences were exactly the same, but in the experimental trials, the sentences included minimal pairs that differed in only one phonological constituent. The authors found that signs that were more difficult to discriminate were those including a minimal pair that differed in movement, followed by signs differing in handshape, and the signs with a different location were the easiest to discriminate. This is one of the first studies to suggest that not all phonological constituents of signs are perceived with equal ease and that they may be mastered at different stages by hearing L2 learners.

In another study, Ortega and Morgan (2015a) looked at phonology acquisition in a group of hearing learners of BSL. In a sign repetition task, participants were asked to imitate a set of signs as accurately as possible. They were tested right before starting their classes in BSL and once again 11 weeks later, which was after 16 hours of total instruction. It was shown that the most accurately produced parameter was location, followed by orientation, then movement, and handshape was the hardest to execute. The authors argued that location and orientation are perceptually salient and therefore are easy to recall. Movement is a fast and ephemeral parameter, which is hard for participants to perceive. The high proportion of errors in handshape were attributed to learners not having developed the motor dexterity to execute hand configurations accurately.

In sum, this body of work suggests that the phonological constituents of signs— handshape, location, and movement—are learned differentially by hearing nonsigners. However, recent research has also shown that the iconicity of signs and their degree of overlap with hearing learners' gestural repertoires are additional resources specific to the visual modality that also influence sign learning.

Role of Iconicity in Learning Sign Language as L2

Whereas past research confined iconicity to the margins of language (see Dingemanse, 2018), in recent years there has been a dramatic increase in studies convincingly demonstrating that iconicity is an incident feature of all modalities of language (i.e., speech, gesture, and sign; Dingemanse et al., 2015; Perniss et al., 2010; Perniss & Vigliocco, 2014). In the realm of sign languages, many scholars have investigated the prevalence of iconic manual structures from different perspectives with some researchers suggesting that iconicity is at the core of sign languages' lexicons (Cuxac, 1999a, 1999b). Indeed, there are reports arguing that up to two thirds of a sign language lexicon can be traced back to an iconic origin (Pietrandrea, 2002) and importantly, that some phonological constituents are not entirely meaningless but rather are semantically loaded (Armstrong & Wilcox, 2009; Occhino, 2017). For example, in NGT, in the sign for "to cut", the two fingers represent the blades of a pair of scissors, and the opening and closing of the fingers represents the cutting motion (Figure 2a). The configuration of the hands of the sign for "butterfly" represents its wings, and the movement reflects their flapping (Figure 2b). These iconic examples show not only that the sign as a whole may represent a referent iconically but also that each phonological constituent may have semantic content individually. In contrast, arbitrary signs, such as the sign for "doctor", which is executed with the tapping of two bent fingers on the chin, lack an evident link with the concept that they represent; and similarly, their constituents are seemingly meaningless on their own (see Figure 2c). In sum, iconic and arbitrary signs coexist in the manual lexicons of all signed languages and constitute the items that L2 learners have to acquire. In addition, some phonological constituents are seemingly meaningless components of signs, but others appear to contribute with specific form-meaning mappings.

A question that has captured the attention of linguists and psychologists alike is whether the perceived iconic resemblance between form and meaning assists in the acquisition of lexical items. A myriad of studies have shown that, in the spoken modality, iconicity assists in many aspects of language learning (see Laing, 2019, for a review). In the context of L2 sign learning, several studies consistently reported a facilitative effect of iconicity on sign learning (Ortega, 2017). More specifically, iconicity has been shown to help hearing nonsigners in a range of tasks including free recall, forced choice recognition, back-forward translation, and picture naming (Baus et al., 2012; Campbell et al., 1992; Lieberth & Gamble, 1991). For instance, a recent study found that nonsigners were better at recognizing iconic signs over arbitrary signs in the Swedish Sign Language, even when signs were presented in the familiar context of a naturalistic weather forecast video (Hofweber et al., 2022, 2023). These sign language studies add to the growing body of evidence suggesting that iconicity is a key resource that facilitates vocabulary development in hearing adults by exploiting the iconic link between a form and its corresponding conceptual representation.

Although there is abundant evidence showing that iconicity assists language learning, some have raised warning signals that its effect is more modest than previously assumed, that it does not generalize to all aspects of language learning (Nielsen & Dingemanse, 2021), or that it is heavily task-dependent and accessible only when participants are asked to actively seek for form-meaning associations (Van Hoey et al., 2023). Indeed, in the context of learning a sign language as L2, it has been suggested that iconicity assists in learning the semantic aspects of signs, but it might actually hinder the acquisition of a manual phonological system (see Ortega, 2017, for a critical review). In a set of experiments looking at acquisition of sign phonology (Ortega & Morgan, 2015a, 2015b), beginner hearing learners who enrolled in the first level of BSL classes were asked to imitate a set of iconic and arbitrary signs as accurately as possible. Signs across conditions were balanced for phonological complexity and were presented only once. After coding for accuracy in four phonological parameters, it was consistently found that iconic signs were actually articulated less accurately than arbitrary signs. The negative effect of iconicity in the execution of the phonological constituents of signs was found when participants were presented the signs both with and without their meaning (Ortega & Morgan, 2015b). The authors interpreted these results as iconicity giving direct access to the meaning of signs, which motivated participants to execute a manual form with the same iconic motivation but without its exact phonological constituents. In contrast, arbitrary signs, which lack clear form-meaning mappings, were more accurately executed because participants had to pay close attention to produce them as they are novel manual forms. The authors also raise the possibility that learners' gestures, which lack conventionalized internal structure (i.e., phonology), may nevertheless interfere in the learning and execution of the phonological constituents of signs (Ortega & Morgan, 2015a). In the following section, we will review some studies that indeed report that gestures, and more specifically, iconic gestures, are recruited by hearing nonsigners at the earliest stages of sign learning.

The Role of Gesture in Learning Sign Language as L2

Previous research claimed that signs and gestures are manual communicative systems that fall at opposite ends of a spectrum with both being fundamentally different in form and function (e.g., McNeill, 1992). However, more recent research comparing both types of manual communication has posited that signs and gestures are bodily actions that may share further forms and functions than previously attested (de Vos, 2015; Padden et al., 2013; Quinto-Pozos & Parrill, 2015). Some have even suggested that gestures and signs are variations of the same communicative phenomenon (Kendon, 2008; Müller, 2018). The attested similarities between gestures and signs make it possible to expect that many signs may overlap in form and meaning with the gestures of hearing nonsigners due to the general affordances of the manual-visual modality (Perniss et al., 2015). Indeed, research has found that when hearing nonsigners are asked to create spontaneous silent gestures for a set of concepts, they tend to produce very similar gestural forms with each other (Ortega & Özyürek, 2019). Importantly, these produced gestures overlap in form and meaning in varying degrees with conventionalized signs. As such, these spontaneous, yet systematic, gestures constitute a semiotic resource that may be readily available to support sign learning.

A small but growing number of studies have demonstrated that iconic gestures may indeed assist different aspects of learning sign language as L2 (e.g., Chen Pichler, 2011; Janke & Marshall, 2017; Marshall & Morgan, 2015; Ortega et al., 2020; Ortega & Özyürek, 2013). Ortega and Ozyürek (2013) asked hearing nonsigners to imitate as accurately as possible a set of BSL signs (e.g., the sign for "deer"). These same participants were called back six months later, this time to take part in a gesture generation task where they had to come up with a spontaneous silent gesture for the same concepts shown in the previous task (e.g., a gesture for the concept "deer"). They found that, in many instances, nonsigning participants produced spontaneous gestures with the same iconic instantiation as signs (e.g., the antlers of a deer) and that in many cases there were overlapping forms between signs and gestures (e.g., using the head for the deer antlers or an extended index finger for the spinning blades of a helicopter). They conclude that there is cross-modal influence from gesture to sign and that learners exploit their gestural reportoire at the earliest stages of sign learning.

In a follow-up study (Ortega et al., 2019), participants were shown NGT signs that overlapped to different degrees with the form of gestures and asked to guess their meaning at first. Participants then received the meaning of the signs and were asked to rate the signs for their degree of iconicity. Participants were significantly more accurate at guessing the meaning of signs and gave higher iconicity ratings when there was high overlap than when there was no overlap between gesture and sign.

In a related study, Ortega and colleagues (2020) used event-related potentials (ERPs) to further investigate the role of gesture in learning sign language as L2. Participants were presented with signs (matched in degree of iconicity) that had an either high or low degree of overlap with previously collected gestures while their electrophysiological brain activity was recorded. At first exposure, signs with low overlap with gestural forms elicited a more positive inflection in the P300 component versus signs with high overlap with gestural forms, which may index the novelty or unexpectedness of a stimulus (Polich, 2009; Van Petten & Luka, 2012). When ERPs were subsequently recorded a second time after a sign learning phase, differences in P300 amplitude between signs in the high and low overlap condition disappeared. The authors interpret these results as evidence that at first exposure, the brain relies on the degree of similarity between signs and participants' gestures. Signs that do not match the form of gestures violate participants' expectations and must therefore be interpreted as "novel". As such, upon first exposure to a sign language, hearing nonsigners seem to fall back on their gestural repertoire to make form-meaning associations, which can be considered the first stage of acquiring an L2 lexicon (Nation, 2001).

The Present Study

Contrary to a large body of work suggesting that iconicity assists lexical learning (see Ortega, 2017), recent studies warn that its effect may be more moderate than previously assumed (Nielsen & Dingemanse, 2021), that it may be task-dependent (Van Hoey et al., 2023), and that it may actually hinder accurate lexical production (Ortega & Morgan, 2015b). As such, the extent to which the degree of iconicity in signs may facilitate learning in comparison to arbitrary signs remains an empirical question. Furthermore, although it is well-documented that both iconicity and gesture may contribute individually to sign learning in hearing adults, it has not yet been explored in the same experimental paradigm whether these two resources jointly boost or modulate the acquisition of signs. In this study, we bring together for the first time iconicity, arbitrariness, and similarity to gestures to explore how these resources modulate the acquisition of lexical signs and their phonological constituents in beginner hearing learners of a sign language. We test the potential influence of these different resources on learning performance both at first exposure and then again one day later. By comparing the interaction of gesture and iconicity, we go beyond the findings reported in earlier studies (e.g., Ortega et al., 2020) and provide a snapshot of sign learning at first exposure and sign consolidation and recall after a 24-hour period.

To do so, we carried out a sign learning experiment where hearing nonsigners were presented with three types of signs in NGT: arbitrary signs, iconic signs with high overlap with gestures, and iconic signs with low overlap with gestures. Participants were instructed to view the signs and their Dutch translations once (familiarisation phase) and then were asked to watch and reproduce the signs three times (learning phase). This second stage ensured that participants would get sufficient exposure to the to-be-acquired signs and would in principle be able to learn them. Participants were then shown the Dutch translation of each sign and were required to produce the NGT sign themselves. They were tested twice on their accuracy of sign production: once after the learning phase (immediate recall) and once again the day after (delayed recall). Assessing sign language learning over two consecutive days provides valuable insights as to whether the influence of iconicity, and/or overlap with gesture, aid learning and consolidation of lexical signs over a 24-hour period.

On the basis of previous research (Baus et al., 2012; Campbell et al., 1992; Ortega, 2017), we predicted that arbitrary signs will be the least accurately acquired whilst signs high in iconicity and high in overlap with gesture (Ortega et al., 2019, 2020) will be the most accurately learned. Regarding the effect of time on recall, there is some evidence that iconic signs are easier to consolidate and recall than arbitrary signs (Lieberth & Gamble, 1991). This would suggest that both types of iconic signs will be more accurately recalled 24 hours after initial exposure than arbitrary signs. However, if signs that overlap with gestures may map onto learners' existing gestural schemas in memory (Ortega et al., 2019) and signs that do not strongly overlap with learners' gestures do not, one could predict that signs that overlap with gestures will actually be most accurately recalled on the second day of testing, as they more easily map onto existing knowledge. Finally, there is evidence showing that location is the parameter easiest to learn in both production (Ortega & Morgan, 2015b) and comprehension (Bochner et al., 2011) in hearing adults. It is therefore reasonable to predict that this parameter should be the most accurately executed compared to the other phonological constituents.

Method

Participants

Thirty-four hearing nonsigners (27 females; mean age = 22.88; age range: 18–30 years) took part in our sign learning study. All participants were born in the Netherlands and had Dutch as their primary language. None of them reported knowing NGT or any other sign language prior to the start of the experiment. Sample size was based on the previous study in the field that theoretically and methodologically most resembled the present study and observed robust effects (i.e., Ortega et al., 2020), with a small increase in sample size as a function of the small decrease in the number of items per condition in the current study.

Stimuli

Stimuli consisted of a total of 96 signs from NGT that were classified into three groups of 32 signs each: (1) iconic signs with high overlap with gestures; (2) iconic signs with low overlap with gestures; (3) arbitrary signs with no overlap with gestures. These categories were determined by consulting a native signer of NGT and by using a gesture database (Ortega & Özyürek, 2019) reporting the canonical silent gestures of a community of nonsigning Dutch speakers for a total of 127 concepts that are consistently produced by at least ten participants out of 20. The form of these gestures is described in its four formational features (i.e., hand configuration, movement, orientation, and placement; Bressem, 2013). These features are loosely based on the phonological constituents of signs and allow for a fairly direct comparison between the form of NGT signs and the normed gestures.

In our study, iconic signs with high overlap with gestures consisted of signs that were iconic in nature and shared at least three out of four parameters (i.e., handshape, location, movement, and orientation) with the corresponding gesture from the normed database. For instance, Figure 2a shows that the sign for “to cut” overlaps in all its formational parameters with the gestures typically produced by nonsigning Dutch speakers for this concept.

Iconic signs with low overlap with gestures consisted of signs that were iconic in nature and shared only one or two constituents with the corresponding gesture from the normed database. For instance, Figure 2b shows that the sign for “butterfly” is represented by two open palms intertwined by the thumb

Figure 2 Examples of iconic signs with high (A) and low (B) overlap with gesture, and arbitrary signs with no overlap with gesture (C).

whereas the gesture for the same concept is executed by flapping both arms. Although the sign is iconic, it has low resemblance with speakers' gestures.

Arbitrary signs consisted of signs that did not exhibit any evident, perceived resemblance with the concepts that they represented, and there was no systematic gesture that could assist participants in the task. For instance, the sign for "doctor" described earlier does not resemble any apparent feature of the medical profession, and hearing gesturers did not produce a systematic gesture for this concept (see Figure 2c). A full list of all stimuli can be found on the Open Science Framework (https://osf.io/pt9xc/?view_only=d2d6d1032c3f4567bb98eafd67e5989f) and in Appendix S1 in the online Supporting Information.

The iconicity ratings[1] for signs with high ($M = 5.13$; $SD = 1.02$) and low ($M = 4.42$; $SD = 1.08$) overlap with gestures did not differ in their degree of iconicity, $t_{[high\ vs.\ low]}\ (31) = 0.795$, $p = .214$. However, the iconicity ratings of these two conditions did differ with signs in the arbitrary condition ($M = 2.10$; $SD = 0.50$); $t_{[high\ vs.\ arbitrary]}\ (31) = 15.123$, $p < .0001$; $t_{[low\ vs.\ arbitrary]}\ (31) =$

13.495, $p < .0001$. The signs in the first two conditions have been previously used in a sign learning study (Ortega et al., 2020). Signs were also balanced across the three conditions according to the number of hands involved in their production (one vs. two). Independent samples *t* tests showed that the number of hands involved did not differ (all *p*s > .05) across arbitrary signs ($M = 1.59$; $SD = 0.50$), signs with high overlap with gesture ($M = 1.56$; $SD = 0.50$), and signs with low overlap with gesture ($M = 1.63$; $SD = 0.49$) (also see Appendix S1 in the online Supporting Information).

Each of the 96 NGT signs had a one-word translation equivalent in Dutch (see Appendix S1 in the online Supporting Information). Independent samples *t* tests showed that the length, as measured in number of letters, and the log frequency of these words (as taken from the SUBTLEX-NL database; Keuleers et al., 2010) did not differ (all *p*s > .05) across arbitrary signs (length $M = 6.03$, $SD = 2.35$; log frequency $M = 2.95$, $SD = 0.82$), signs with high overlap with gesture (length $M = 6.28$, $SD = 2.13$; log frequency $M = 2.82$, $SD = 0.79$), and signs with low overlap with gesture (length $M = 6.63$, $SD = 2.25$; log frequency $M = 2.68$, $SD = 0.92$).

Procedure

Participants were tested individually in a quiet room where they were seated in front of a 20-inch Samsung computer monitor. The testing procedure consisted of two sessions conducted on two consecutive days (day 1, day 2). At the beginning of each session, participants were informed about the session structure, and their signed informed consent forms were collected.

The first session involved four blocks (first exposure, learning, second exposure, testing). Details of each session are described in detail below. Each block was preceded by five practice trials which relied on stimuli that were not used in the experimental trials. The second session, which was conducted on the day following the first session, repeated only the last block of the first testing session[2].

1. First exposure (Block 1)[3]: Participants were presented with all the 96 signs (32 per condition) described in the stimuli section in random order. Each trial started with a fixation cross that appeared in the middle of the computer screen for 500 ms. This was followed by a printed word in Dutch (e.g., *vlinder* meaning "butterfly" in Dutch), which remained on the screen for 1000 ms. Next, another fixation cross appeared in the middle of the screen for 500 ms, followed by the NGT sign equivalent to the Dutch word (e.g., the sign for "butterfly") in a video (14 × 8 cm). After the sign had

played in full, the next trial began. Participants were instructed to pay close attention to the words and signs and did not receive any other instructions or tasks in this block.

2. Learning (Block 2): Participants were presented with the videos of the same signs from Block 1 and were asked to imitate the presented signs as accurately as possible. Each trial started with a fixation cross that appeared in the middle of the screen for 500 ms before the presentation of the video of a sign. The Dutch translation of the sign (e.g., the Dutch word for "butterfly") was presented under the corresponding video for the duration of the video. Next, a blank screen appeared for 3000 ms, allowing participants to repeat the sign. This procedure was repeated three subsequent times for each sign within a trial, after which the next trial started. This block was video recorded from the side top angle for future coding for accuracy.
3. Second exposure (Block 3): This block was identical to Block 1 except that the signs were presented in a different random order.
4. Testing (Block 4): This block assessed participants' ability to remember the 96 signs. Each trial started with a fixation cross that appeared in the middle of the screen for 200 ms. This was followed by a blank screen that appeared for 200 ms. Next, a printed word appeared for 6000 ms. Printed words were the Dutch translations of the 96 signs presented throughout the experiment. Participants were instructed to produce the corresponding NGT sign for the printed word or say "pass" when they failed to remember the sign. No feedback was provided. Each trial ended after 6000 ms, after which the next trial would begin. This block was video recorded from the side top angle to allow for coding for accuracy.

Coding and Reliability

Once all data were collected, they were annotated and coded using ELAN, a free annotation tool (http://tla.mpi.nl/tools/tla-tools/elan/) for multimedia resources developed by the Max Planck Institute for Psycholinguistics, the Language Archive, Nijmegen, The Netherlands (Sloetjes & Wittenburg, 2018). Once all participants' renditions were annotated, articulation accuracy of each rendition was coded for three parameters: handshape, location, and movement. Note that for the purpose of this study, we consider that orientation is not a separate parameter but a feature of the parameter handshape (Sandler, 1989). Each parameter was subjected to a binary coding, where 1 represented an accurate rendition and 0 represented an inaccurate rendition. The annotation and coding were done by a trained research assistant, who was a hearing nonnative signer, blind to the aim of the study. In order to ensure reliability, another trained research assistant, who was a deaf native signer, independently coded 20% of the data (i.e., all data from randomly selected seven participants) for the articulation accuracy for the parameters handshape, location, and movement. There was a substantial agreement between the coders for the accuracy of handshape (81.81%), location (88.87%), and movement (84.35%). All disagreements were discussed to reach a 100% agreement.

In order to determine the accuracy for each parameter, we took into consideration the following criteria.

1. Handshape: the characterization of the configuration of the hand with regard to the selected fingers, their aperture, and their configuration (van der Kooij, 2002).
2. Location: the place of articulation of the sign. This can be neutral space (i.e., the area in front and above the signer), body-anchored (i.e., a location on the body), or the nondominant hand (i.e., the supporting hand in signing).
3. Movement: a change of position of the main articulator (i.e., the hand), a change in the position of the digits, or a change in the orientation of the articulator. Movement can be classified as path (if the main articulator moves across space), internal (if there is movement of the wrist or the fingers), or it may have both path and internal movement.

In the interest of readability, we refer the reader to the coding scheme described by Ortega and Morgan (2015a) for a full description of the coding criteria.

Results

Data were analysed using generalized linear mixed-effects modelling (glmer) and linear mixed-effects modelling (lmer) with random intercepts for participants and items. This mixed-effects approach allowed us to take into account the random variability due to having different participants and different items. All models reported below were fit with lme4 package (version 1.1-21; Bates et al., 2014) in R (R Core Team, 2018). We used the package LmerTest (version 3.1.-2; Kuznetsova et al., 2017) to retrieve *p* values and the package emmeans (version 1.5.0; Length, 2019; Searle et al., 1980) to interpret significant effects. The analysis script can be found on the Open Science Framework (https://osf.io/pt9xc/?view_only=d2d6d1032c3f4567bb98eafd67e5989f).

First, we investigated whether participants' overall articulation accuracy was predicted by condition (i.e., iconic signs with high overlap with gestures, iconic signs with low overlap with gestures, and arbitrary signs with no

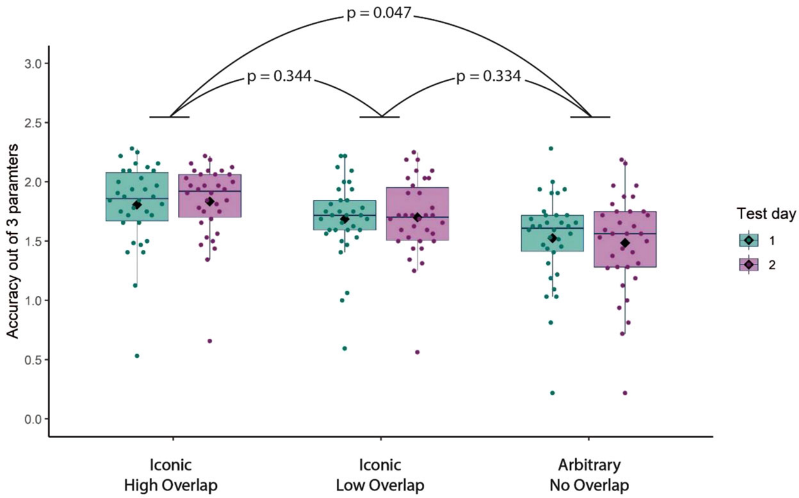

Figure 3 Overall accuracy scores across conditions and test day. Coloured dots represent the average overall accuracy score for each individual participant. Black diamond shapes represent the mean per condition.

overlap with gestures) and test day (i.e., day 1, day 2) or an interaction between condition and test day. To do so, for every participant, we calculated an accuracy score for each item by summing the accuracy of the three parameters. Scores thus ranged from 0 (no parameters were correct, or the participant did not articulate the sign and skipped the particular trial) to 3 (all parameters were correct). We used an lmer model to test for the main effects of condition and test day, and their interaction on the overall accuracy score at the item level (see Figure 3). The fixed effect of condition was then analysed with Helmert contrasts, as this approach allowed us to statistically contrast our conditions of theoretical interest within the same model. That is, at the first level of the statistical model, we compared iconicity with arbitrariness by contrasting the average of the two iconic conditions (i.e., the iconic high overlap with gesture and the iconic low overlap with gesture conditions) to the arbitrary signs. At the second level, we compared the two iconic conditions directly. To do so, at the first level, iconic signs with high overlap with gestures were coded as $-1/3$, iconic signs with low overlap with gestures were coded as $-1/3$, and arbitrary signs with no overlap with gesture were coded as $+2/3$. At the second level, iconic signs with high overlap with gestures were coded as $+1/2$, iconic signs with low overlap with gestures were coded as $-1/2$, and arbitrary signs with no overlap were coded as 0. The fixed effect of test day was analysed with centered contrasts ($-1/2$, $+1/2$).

Table 1 presents the fixed effect estimates from the lmer model for overall articulation accuracy. The model revealed a fixed effect of condition only for the first level comparison, $\beta = 0.05$, $SE = 0.12$, $p = .032$. As such, arbitrary signs, $M = 1.51$, 95% CI [1.17, 1.85], were articulated less accurately than the average of iconic signs with high and low overlap with gestures, $M = 1.76$, 95% CI [1.44, 2.08]. The model did not reveal a significant effect of condition for the second level comparison, $\beta = 0.13$, $SE = 0.13$, $p = .344$, indicating that iconic signs with high overlap with gestures, $M = 1.82$, 95% CI [1.51, 2.13], versus low overlap with gestures, $M = 1.69$, 95% CI [1.36, 2.03] did not lead to statistically significant differences in articulation accuracy. No other main and interaction terms were significant, indicating that participants' articulation accuracy did not differ significantly across the two testing days. In other words, the results from the first day were replicated on the second day of testing.

To follow up on the main effect of condition, we compared articulation accuracy across all pairs of conditions directly. Results showed that arbitrary signs with no overlap with gestures were articulated significantly less accurately than iconic signs with high overlap with gestures, $\beta = 0.32$, $SE = 0.13$, $p = .047$, but not significantly less accurately than iconic signs with low overlap with gestures, $\beta = 0.19$, $SE = 0.13$, $p = .334$. That is, the main effect of condition (on the first-level comparison) on the overall articulation accuracy was mainly driven by the difference between arbitrary signs with no overlap with gestures and iconic signs with high overlap with gestures. As such, the combination of iconicity and high overlap with gesture led to the observed learning benefits—only iconicity (as in the comparison of the arbitrary signs with the iconic signs with low overlap with gestures) or only overlap with gestures (as in the comparison of the iconic signs with high vs. low overlap with gestures) did not suffice for a robust learning advantage.

Next, we investigated the effect of condition separately for the articulation accuracy of each parameter. This analysis was collapsed across test day, as it did not significantly explain variation in the data in the main analysis presented above. Trials on which no attempt was made to produce a sign were not included in this analysis, as these trials by definition did not allow for contrasting the different parameters in terms of accuracy. For each parameter (handshape, location, movement), we hence used separate glmer models to test the main effect of condition on the binary values for accuracy (1 = accurate, 0 = inaccurate). In all of the models, the fixed effect of condition was entered

Table 1 Fixed effect estimates from the lmer model for overall accuracy. Bold typeface indicates a significant effect

Fixed Effect	β	*SE*	*df*	95%CI	*t*	*p*	R^2
							.17
(Intercept)	1.674	0.077	9.49	[1.52, 1.83]	21.669	<.001	
Condition $_{\text{[Level 1: arbitrary vs iconic]}}$	0.052	0.116	9.57	[−0.481, −0.02]	−2.182	**.032**	
Condition $_{\text{[Level 2: iconic high vs iconic low]}}$	0.127	0.134	9.57	[−0.14, 0.39]	0.951	.349	
TestDay $_{\text{[day1 vs day2]}}$	−1.124	0.019	6.40	[−0.04, 0.04]	−0.039	.970	
Condition $_{\text{[Level 1: arbitrary vs iconic]}}$ ∗ Test Day $_{\text{[day1 vs day2]}}$	<0.001	0.040	6.40	[−0.02, 0.14]	1.556	.120	
Condition $_{\text{[Level 2: iconic high vs iconic low]}}$ ∗ TestDay $_{\text{[day1 vs day2]}}$	0.005	0.047	6.40	[−0.11, 0.08]	−0.331	.741	

Note. Model in *R*: Overall_Accuracy ∼ Condition ∗ TestDay + (1|Participant) + (1|Item) with Helmert contrasts identical to the analyses reported above. See Figure 4 and Table 2 for the results of separate parameters.

For handshape, the model did not reveal a significant effect for condition for both comparisons, Level 1: $\beta = -0{,}35$, $SE = 0.33$, $p = .279$; Level 2: $\beta = 0.25$, $SE = 0.37$, $p = .503$. That is, neither the comparison between arbitrary signs with no overlap with gestures, $M = 0.33$, 95% CI [0.17, 0.49], versus the average of iconic signs with high and low overlap with gestures, $M = 0.39$, 95% CI [0.23, 0.56], nor the comparison between iconic signs with high overlap with gestures, $M = 0.41$, 95% CI [0.24, 0.57], versus iconic signs with low overlap with gestures, $M = 0.37$, 95% CI [0.21, 0.54], led to statistically significant differences in the accuracy of the produced handshapes.

For location, the model revealed a significant effect of condition for the first-level comparison, $\beta = -0.76$, $SE = 0.38$, $p = .047$, indicating that in terms of location, arbitrary signs with no overlap with gestures, $M = 0.80$, 95% CI [0.66, 0.93], were remembered less accurately than iconic signs, $M = 0.85$, 95% CI [0.74, 0.97], averaged across degree of iconicity (high and low overlap). However, the model did not reveal a significant effect of condition for the second-level comparison, $\beta = 0.71$, $SE = 0.45$, $p = .109$: Iconic signs with high overlap with gestures, $M = 0.89$, 95% CI [0.79, 1.00], versus iconic signs with low overlap with gestures, $M = 0.82$, 95% CI [0.68, 0.95], did not lead to statistically significant differences in articulating the location of the sign accurately. To follow up on the main effect, we further compared location accuracy across the pairs of conditions. Results showed that, although overall performance was high across conditions for location, arbitrary signs with no overlap with gesture, $M = 0.80$, 95% CI [0.66, 0.93], were articulated significantly worse than iconic signs with high overlap with gestures, $M = 0.89$, 95% CI [0.79, 1.00], $\beta = 1.11$, $SE = 0.44$, $p = .032$, on this parameter. However, the comparison of arbitrary signs with no overlap with gestures, $M = 0.80$, 95% CI [0.66, 0.93], versus iconic signs with low overlap with gestures, $M = 0.82$, 95% CI [0.68, 0.95], did not lead to statistically significant differences in articulation accuracy, $\beta = 0.40$, $SE = 0.44$, $p = .634$. That is, the main effect of condition for the first-level comparison for the accuracy of location was mainly driven by the difference between arbitrary signs with no overlap with gestures and iconic signs with high overlap with gestures.

For movement, the model did not reveal a significant effect for condition for either comparison, Level 1: $\beta = -0.49$, SE = 0.35, p = .166; Level 2: $\beta = 0.10$, SE = 0.41, p = .814. That is, neither the comparison across arbitrary signs with no overlap with gestures, $M = 0.46$, 95% CI [0.29, 0.62], versus

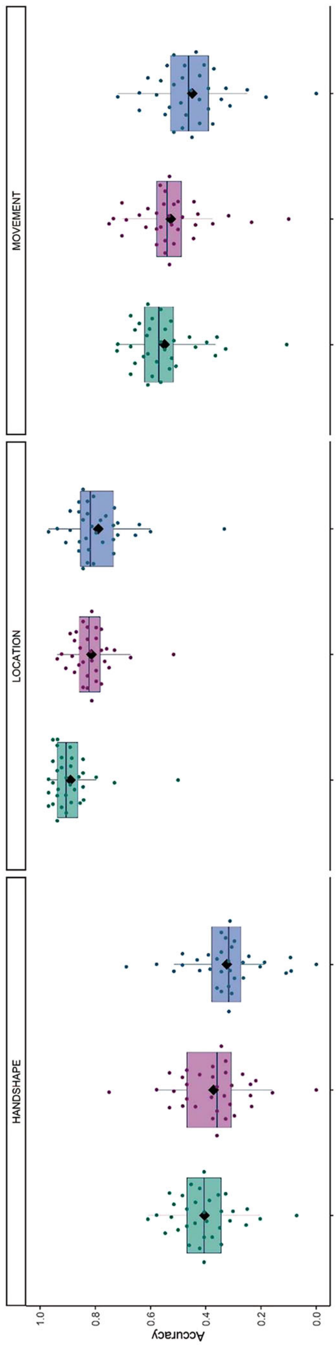

Figure 4 Articulation accuracy for each phonological parameter across conditions. Coloured dots represent the average accuracy score for each individual participant on a given parameter split by condition. Black diamond shapes represent the mean per condition.

Table 2 Fixed effect estimates from the glmer model for handshape, location, and movement. Bold typeface indicates a significant effect

Fixed Effect	β	*SE*	95%CI	*z*	*p*	R^2
Handshape						.005
(Intercept)	−0.838	0.202	[−1.24, −0.44]	−4.156	<.001	
Condition [Level 1: arbitrary vs iconic]	−0.352	0.325	[−0.99, 0.29]	−1.083	.279	
Condition [Level 2: iconic high vs iconic low]	0.251	0.374	[−0.49, 0.99]	0.670	.503	
Location						.016
(Intercept)	2.443	0.219	[2.02, 2.89]	11.178	<.001	
Condition [Level 1: arbitrary vs iconic]	−0.755	0.380	[−1.52, −0.00]	−1.986	**.047**	
Condition [Level 2: iconic high vs iconic low]	0.712	0.445	[−0.17, 1.61]	1.601	.109	
Movement						.009
(Intercept)	0.027	0.216	[−0.40, 0.46]	0.125	.901	
Condition [Level 1: arbitrary vs iconic]	−0.486	0.351	[−1.18, 0.21]	−1.384	.166	
Condition [Level 2: iconic high vs iconic low]	0.096	0.406	[−0.71, 0.90]	0.235	.814	

Note: Models in R:

Handshape ~ Condition + (1|Participant) + (1|Item)
Location ~ Condition + (1|Participant) + (1|Item)
Movement ~ Condition + (1|Participant) + (1|Item)

the average of iconic signs with high and low overlap with gestures, $M = 0.54$, 95% CI [0.37, 0.71], nor the comparison across iconic signs with high overlap with gestures, $M = 0.55$, 95% CI [0.38, 0.72], versus low overlap with gestures, $M = 0.53$, 95% CI [0.36, 0.70], led to statistically significant differences in articulating the movement of the sign accurately.

In sum, these follow-up analyses for the separate phonological constituents indicate that the overall learning effect observed above (Table 1 and Figure 3) must have been mainly driven by participants' performance differences between the arbitrary signs with no overlap with gesture versus the iconic signs with high overlap with gesture in the accuracy of reproducing the location of the to-be-acquired signs (see Figure 4 and Table 2).

Discussion

In this study, we investigated the potential joint influence of iconicity and gesture in sign vocabulary learning at early exposure by hearing nonsigners. Using a sign learning paradigm, we presented hearing nonsigning participants with a relatively large selection of NGT signs, each of which corresponded to one of three conditions: arbitrary signs with no overlap with gestures, iconic signs with high overlap with gestures, and iconic signs with low overlap with gestures. On day 1, participants were presented with the signs for the first time in a randomized order. They then viewed the signs three times and were asked to repeat them to promote learning. After this stage, we tested sign learning in two consecutive phases: once immediately after participants' repetition of signs and once again the day after. In these testing phases, participants were presented with the Dutch translation of the signs in isolation and were asked to produce the signs from memory. Sign learning was operationalized through the composite score of articulation accuracy of the phonological constituents handshape, location, and movement.

Our results revealed that iconicity and gestural overlap jointly led to better recall performance. Arbitrary signs were acquired less accurately compared only to iconic signs with high overlap with gestures. This effect, in addition, seemed mainly driven by the parameter location. Indeed, when we analyzed each parameter separately, statistically we found that only for location arbitrary signs with no overlap with gesture were learned less accurately than iconic signs with high overlap with gestures. The parameters handshape and movement revealed no such effect, although showing similar patterns of results numerically. In addition, we observed no effect of testing phase. Together, these data hence suggest that iconicity together with overlap with gesture facilitate sign learning at first exposure to a sign language.

In general, the present findings align well with a large body of work in both the spoken (Lockwood et al., 2016) and sign modality (Baus et al., 2012; Campbell et al., 1992; Lieberth & Gamble, 1991). The absence of any evident perceived resemblance between a lexical label and the concept that it represents appears to be quite challenging to learners as they are unable to latch onto any prior knowledge to scaffold a new word form in the target L2. In the context of sign learning, arbitrary signs pose an additional challenge to learners as they may be inclined to overinterpret the sign's structure and seek a visual motivation that is not present. Evidence suggests that sign-naïve adults are biased towards assuming that signs denote some form of action or that they capture visual features of the referent (Ortega et al., 2019), but in the case of arbitrary signs, this bias does not facilitate access to the correct meaning of the sign. It is possible that one of the reasons why arbitrary signs are quite challenging for this type of learners is that they have to override these biases to assign the signs their correct meaning.

In this study we also aimed to disentangle further the effect of iconicity and gesture in sign language learning. Previous studies have reported the positive effect of these two variables separately (Ortega et al., 2019, 2020), but until now no study has manipulated them in the same experimental paradigm. Further, recent evidence suggested that the effect of iconicity in lexical learning may not be as strong and prevalent as previously assumed (Nielsen & Dingemanse, 2021; Ortega & Morgan, 2015b; Van Hoey et al., 2023). The general picture that emerges from the present study is that iconicity and gestural overlap combined lead to better learning and recall. In our results, we can see that the arbitrary and low overlap conditions differ only in iconicity, and there we find no significant learning benefit. The two iconic conditions differ in gestural overlap only and did not elicit a learning difference. In contrast, the arbitrary and high overlap conditions differ in both iconicity and gestural overlap, and there we do find an effect. These results thus lend further credence to the consistent finding that arbitrary signs are harder to acquire when compared with iconic signs. We show for the first time that a combination of iconic form-meaning mapping and overlap with available manual schemas (i.e., one's gestural repertoire) may help hearing learners to correctly learn and produce lexical signs.

Interestingly, we did not observe a significant difference in learning performance between iconic signs with low overlap with gestures and arbitrary signs. One possible explanation is that the current stimulus list composition made it difficult for participants to consistently rely on obvious form-meaning mappings. For instance, the sign for "laptop" depicts the rectangular shape of a portable computer, so this shape-based resemblance may have confirmed participants' expectations of the form of the sign. Such form-meaning links were however absent in signs like "doctor," which is executed with two taps on the chin. After not seeing an iconic link between such a sign and its meaning on a given trial, learners may have (at least partially) abandoned the strategy of searching for direct form-meaning mappings. In other words, relying on iconicity was not always a successful strategy. In contrast, when the iconic link between form and meaning was strengthened by overlap with gestures, participants' learning performance did benefit. In such cases of iconic signs that have a high overlap with gestures, the iconic link between form and meaning arguably mapped onto an existing gestural schema in learners' long term memory. This line of reasoning thus suggests that mapping an iconic L2 sign onto an existing L1 gestural schema facilitates learning compared to situations where the iconic link does not yet exist in the learner's memory and first has to be recognized and established. As such, different signs may afford different learning strategies to make form-meaning associations and thus impact sign learning. In any case, our results make it clear that future studies comparing the acquisition of iconic and arbitrary signs should control for gestural overlap.

Another finding that merits attention is that location appears to be the main driver for arbitrary signs being executed less accurately than iconic signs with high overlap with gestures, and at the same time location is the parameter that is most accurately acquired overall (see Figure 4). Previous studies have looked at how each individual phonological parameter contributes to sign learning, with location indeed being the most accurately produced and discriminated (Bochner et al., 2011; Ortega & Morgan, 2015a). The picture that emerges in our study is that location is indeed an important parameter that may be most responsible for differences in learning across conditions, although we do note that results for all three parameters numerically showed a similar pattern. We tentatively propose that some parameters (e.g., location) may present a learning advantage, as they may be more semantically loaded than others. Some scholars have indeed argued against the notion of combinatorial units (dual patterning) in sign languages because the constituents of signs are rarely entirely meaningless (Armstrong & Wilcox, 2009). In many sign languages, for instance, most signs representing mental processes are articulated close to the head, and signs referring to food consumption are executed around the mouth. This is in stark contrast with arbitrary signs, whose location has no apparent motivation vis-à-vis the concepts they represent (e.g., there is no clear reason why the sign for "doctor" in NGT is executed on the chin). A tentative conclusion for the relatively high accuracy in acquiring the signs' correct location is hence that different phonological parameters may contribute differentially to the meaning of signs. Future studies should test this possibility in sign learning and consider not only iconicity and gestural overlap of signs but also the semantic load of each parameter. In addition, longitudinal and more strongly powered follow-up studies may reveal whether the numerical trends for the parameters handshape and movement are borne out. As such, they will shed more light on the complex interaction between iconicity, gestural overlap, and the phonological constituents of signs.

Interestingly, participants displayed the same accuracy in sign learning at both testing phases. If signs with overlap with gestures had a strong facilitative effect in sign learning, it would have been possible to see that these signs are less prone to being forgotten and would have been better recalled on day 2 than signs in the other two conditions. The mechanism that could explain these results is that signs with overlap with gestures map onto existing schemas in long term memory and are therefore potentially consolidated more easily for longer, as they connect to well-established existing knowledge. For signs with low overlap with gestures, that is not the case. Since they cannot be easily connected to existing schemas in memory, they may be lost more easily. In our study, potentially due to the high saliency of items to be learned, a one-day retention interval may have been too short to capture a possible difference in retention and recall across conditions. Future longitudinal work may further test the potential role of existing schemas in facilitating the long-term consolidation of sign language vocabulary.

Given the prevalence of iconicity in all modalities of language and the increased interest for its effect on language learning, it would be beneficial for the field of Second Language Acquisition (SLA) in general to explore in more detail the role of iconicity in L2 learning. There is a small but growing number of studies that have found positive effects of perceived iconicity in vocabulary learning beyond learning signs. Deconinck et al. (2015) found that adult Dutch speakers learning English were more successful at learning words with clear sound-symbolic associations than words that did not have such iconic links. With regards to iconic co-speech gesture, Kelly et al. (2008) found that Japanese words are more easily learned when they are presented with a semantically related iconic gesture (e.g., *nomu* “drink” in Japanese presented with a drinking gesture) than when they are presented in speech only. There is also a growing body of psycholinguistic evidence showing that iconic forms are easier to learn and process (see Lockwood et al., 2016). For instance, onomatopoeic words are processed more easily than words that lack an apparent resemblance between form and meaning, arguably precisely because they partially map onto existing environmental sound representations in the language user's long term memory (Peeters, 2016). A fertile avenue for future exploration is to exploit learners' existing iconic repertoires across all modalities of language (i.e., sound symbolism, iconic gestures, iconic signs) and investigate how these may independently or jointly influence lexical L2 learning. The results of the present study showcase how indeed L2 sign learning can "shed light on important issues in mainstream SLA/bilingualism" (Gullberg, 2022, p. 231).

Limitations and Future Directions

As the reader will have noticed, our experimental design did not orthogonally manipulate signs' overlap with gestures (high vs. low) and form-meaning mapping (arbitrary vs. iconic), although in theory a 2×2 factorial design would have been a possibility. Indeed, our experiment lacked a condition that involved to-be-acquired signs that are not iconic but do have a substantial overlap with learners' gestures, as such sign-gesture combinations were difficult to find. These do however exist in the form of emblems, such as the sign and gesture for the concept "good": a closed fist with thumbs up. Therefore, a remaining issue that can be tested in future studies is the extent to which such noniconic signs that resemble gestures facilitate learning. Based on our results, in spite of their low degree of iconicity, we predict that such signs will be relatively easily learned because of their similarity with people's gestures, and therefore, the well-established gestural schemas can be mapped in learners' memory. However, this potential facilitatory effect might still be lower than the facilitation observed for iconic signs that have a high overlap with gestures, as the present study suggests that having two facilitating features (i.e., gestural overlap and iconicity) rather than one (i.e., gestural overlap or iconicity) may benefit the learner the most.

In general, our study was limited, as it tested a relatively homogeneous group of highly educated adult learners. The extent to which age affects the speed and accuracy of acquiring different types of signs in other hearing populations remains an open question. For instance, older children with extensive world knowledge may benefit more from gestural overlap and iconicity present in signs than younger children with relatively restricted world knowledge (e.g., Newport & Meier, 1985; see also Ortega, 2017, for discussion). Alternatively, hearing parents of deaf children have extensive experience with iconicity and gestural communication, and this experience could be used to enhance and stimulate signing skills more rapidly (Mitchell & Karchmer, 2004). This would result in caregivers having automatic access to a plethora of gestural communicative resources to support one-to-one communication without interpreters' assistance. Interestingly, we observed our effect of gestural overlap and iconicity in Dutch learners that are based in a so-called "low gesture culture". Our findings may be enhanced in learners from spoken communities in "high gesture cultures," such as Turkey, which receive and produce more gestural signals in their everyday spoken interactions to begin with, compared to the Netherlands (see Azar et al., 2020, for a comparison). The relatively low effect sizes observed in the current study may indeed be expected to increase as a function of gesture being a more important communicative tool in learners' everyday life. In sum, we look forward to any future replications of the present study in populations that differ in age and cultural background.

Conclusion

Hearing nonsigners have a rich toolkit of semiotic resources for making accurate form-meaning associations when they are exposed to signs never seen before, even when there are no overlapping conventionalized and lexicalized forms between learners' L1 and L2. Iconicity and gestures are two important resources that give nonsigners an advantageous headstart in their first experience with a sign language. Our findings show that the combination of iconicity and gestures helps hearing learners during their first steps of acquiring an L2 sign language lexicon and correctly producing the newly learnt lexical signs. The important influence of both resources at the same time might also explain why some previous studies, which typically did not control for gestural overlap when selecting stimulus materials, found an advantage of iconicity and some others did not. Broadly speaking, our findings confirm that pedagogical settings may take both the learners' gestural repertoire and the presence of iconicity as early gateways to successful sign language learning.

Final revised version accepted 23 October 2023

Acknowledgments

This study was possible thanks to a VENI grant awarded to Gerardo Ortega (275-890-21) and a VICI grant awarded to Aslı Özyürek (277-70-013) by the Dutch Science Foundation. We are also indebted to Max Planck Institute for Psycholinguistics for allowing the use of their premises for data collection and Max Planck Gesselschaft. This study received ethical approval from the Ethics Assessment Committee (EAC) of the Faculty of Arts of Radboud University (MvB14U.015319). We would like to express our gratitude to Carla Pol, Yeşim Özüer, Bianca Thomsen, and Birgit Knudsen for their support in data collection and coding of the signed videos. We are grateful to Tom Uittenbogert for his assistance to select and film the stimulus materials.

Open access funding enabled and organized by Projekt DEAL.

Open Research Badges

This article has earned Open Data and Open Materials badges for making publicly available the digitally-shareable data and the components of the research methods needed to reproduce the reported procedure and results. All data and materials that the authors have used and have the right to share are available at https://osf.io/pt9xc/?view_only=d2d6d1032c3f4567bb98eafd67e5989f. All proprietary materials have been precisely identified in the manuscript.

Notes

1 Ratings for the NGT signs come from Ortega et al. (2019). In that study, Dutch hearing nonsigners were presented with a randomized list of NGT signs and were asked, first, to guess the meaning of the sign. After they typed their response, they were given the correct translation and were asked to rate on a seven-point Likert scale (*1–7*) to what extent the sign conveyed the meaning of the word.

2 We did not ask participants to self-report their quality of sleep, as in our within-participant design we would not predict it to differentially affect performance in the different experimental conditions.

3 During Blocks 1 and 3, participants' electroencephalogram was continuously recorded. These data are beyond the scope of the current paper and will be reported in a separate manuscript.

References

Armstrong, D. F., & Wilcox, S. E. (2009). Gesture and the nature of semantic phonology. *Sign Language Studies*, *9*(4), 410–416. https://doi.org/10.1353/sls.0.0021

Azar, Z., Özyürek, A., & Backus, A. (2020). Turkish-Dutch bilinguals maintain language-specific reference tracking strategies in elicited narratives. *International Journal of Bilingualism*, *24*(2), 376–409. https://doi.org/10.1177/1367006919838375

Bassetti, B. (2008). Orthographic input and second language phonology. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 191–206). Multilingual Matters. https://doi.org/10.21832/9781847691118-013

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baus, C., Carreiras, M., & Emmorey, K. (2012). When does iconicity in sign language matter? *Language and Cognitive Processes*, *28*(3), 261–271. https://doi.org/10.1016/j.cognition.2008.05.012

Bochner, J. H., Christie, K., Hauser, P. C., & Searls, J. M. (2011). When is a difference really different? Learners' discrimination of linguistic contrasts in American Sign Language. *Language Learning*, *61*(4), 1302–1327. https://doi.org/10.1111/j.1467-9922.2011.00671.x

Brentari, D. (1999). *A prosodic model of sign language phonology*. MIT Press. https://doi.org/10.7551/mitpress/5644.003.0006

Bressem, J. (2013). A linguistic perspective on the notation of form features in gesture. In C. Muller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Body language communication: An international handbook on multimodality in human interaction* (pp. 1079–1098). De Gruyter Mouton. https://doi.org/10.1515/9783110261318.1079

Campbell, R., Martin, P., & White, T. (1992). Forced choice recognition of sign in novice learners of British Sign Language. *Applied Psycholinguistics*, *13*(2), 185–201. https://doi.org/10.1093/applin/13.2.185

Chen Pichler, D. (2011). Sources of handshape error in first-time signers of ASL. In G. Mathur & D. J. Napoli (Eds.), *Deaf around the world: The impact of language* (pp. 96–126). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199732548.003.0005

Cuxac, C. (1999a). French sign language: proposition of a structural explanation by iconicity. *Gesture-Based Communication in Human-Computer*, *1739*(1), 165–184. https://doi.org/10.1007/3-540-46616-9_16

Cuxac, C. (1999b). The expression of spatial relations and the spatialization of semantic relations in French Sign Languag. In C. Fuchs & S. Robert (Eds.), *Language diversity and cognitive representations* (pp. 123–142). John Benjamins. https://doi.org/10.1075/hcp.3

Deconinck, J., Boers, F., & Eyckmans, J. (2015). "Does the form of this word fit its meaning?" The effect of learner-generated mapping elaborations on L2 word recall. *Language Teaching Research*, *21*(1), 1–23. https://doi.org/10.1177/1362168815614048

de Vos, C. (2015). The Kata Kolok pointing system: Morphemization and syntactic integration. *Topics in Cognitive Science*, *7*(1), 150–168. https://doi.org/10.1111/tops.12124

Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: a journal of general linguistics*, *3*(1), https://doi.org/10.5334/gjgl.444.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615. https://doi.org/10.1016/j.tics.2015.07.013

Dingemanse, M., Perlman, M., & Perniss, P. (2020). Construals of iconicity: Experimental approaches to form-meaning resemblances in language. *Language and Cognition*, *12*(1), 1–14. https://doi.org/10.1017/langcog.2019.48

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*, 551–585. https://doi.org/10.1017/s0272263104040021

Frederiksen, A. T., & Mayberry, R. I. (2018). Reference tracking in early stages of different modality L2 acquisition: Limited over-explicitness in novice ASL signers' referring expressions. *Second Language Research*, *35*(2), 252–283. https://doi.org/10.1177/0267658317750220

Gulamani, S., Marshall, C. R., & Morgan, G. (2020). The challenges of viewpoint-taking when learning a sign language: Data from the "frog story" in British Sign Language. *Second Language Research*, *38*(1), 55–87. https://doi.org/10.1177/0267658320906855

Gullberg, M. (2022). Why the SLA of sign languages matters to general SLA research. *Language, Interaction and Acquisition*, *13*, 231–253. https://doi.org/10.1075/lia.22022.gul

Hofweber, J. E., Aumonier, L., Janke, V., Gullberg, M., & Marshall, C. (2022). Breaking into language in a new modality: The role of input and individual differences in recognising signs. *Frontiers in Psychology*, *13*, Article 895880. https://doi.org/10.3389/fpsyg.2022.895880

Hofweber, J. E., Aumonier, L., Janke, V., Gullberg, M., & Marshall, C. (2023). Which aspects of visual motivation aid the implicit learning of signs at first exposure? *Language Learning*, 1–31. https://doi.org/10.1111/lang.12587

Janke, V., & Marshall, C. R. (2017). Using the hands to represent objects in space: Gesture as a substrate for signed language acquisition. *Frontiers in Psychology*, 8. https://doi.org/10.3389/fpsyg.2017.02007

Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *The Modern Language Journal*, *88*(4), 416–432. https://doi.org/10.1111/j.0026-7902.2004.00238.x

Kelly, S. D., Manning, S. M., & Rodak, S. (2008). Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology and education. *Language and Linguistics Compass*, *2*(4), 569–588. https://doi.org/10.1111/j.1749-818X.2008.00067.x

Kendon, A. (2008). Some reflections on the relationship between 'gesture' and 'sign.' *Gesture*, *8*(3), 348–366. https://doi.org/10.1075/gest.8.3.05ken

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*, 643–650. https://doi.org/10.3758/BRM.42.3.643

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Laing, C. (2019). A role for onomatopoeia in early language : Evidence from phonological development. *Language and Cognition*, *11*(2), 1–15. https://doi.org/10.1017/langcog.2018.23

Length, R. (2019). emmeans: Estimated marginal means, aka least-squares means. Retrieved from https://CRAN.R-project.org/package=emmeans

Lieberth, A. K., & Gamble, M. E. (1991). The role of iconicity in sign language learning by hearing adults. *Journal of Communication Disorders*, *24*(2), 89–99. https://doi.org/10.1016/0021-9924(91)90013-9

Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-symbolism boosts novel word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. *39*(1), 2–13. https://doi.org/10.1037/a0028642

Marshall, C. R., & Morgan, G. (2015). From gesture to sign language: Conventionalization of classifier constructions by adult hearing learners of British Sign Language. *Topics in Cognitive Science*, *7*(1), 61–80. https://doi.org/10.1111/tops.12118

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press. https://doi.org/10.2307/1576015

Mesch, J., & Schönström, K. (2021). Use and acquisition of mouth actions in L2 sign language learners. *Sign Language & Linguistics*, *24*(1), 36–62. https://doi.org/10.1075/sll.19003.mes

Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, *4*(2), 138–163. https://doi.org/10.1353/sls.2004.0005

Müller, C. (2018). Gesture and sign: Cataclysmic break or dynamic relations? *Frontiers in Psychology*, *9*, 1651. https://doi.org/10.3389/fpsyg.2018.01651

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524759

Newport, E. L., & Meier, R. P. (1985). The acquisition of American Sign Language. In D. Slobin (Ed.), *The cross-linguistic study of language acquisition* (pp. 881–938). Lawrence Erlbaum Associates. https://doi.org/10.1017/S0272263100211054

Nielsen, A. K. S., & Dingemanse, M. (2021). Iconicity in word learning and beyond: A critical review. *Language and Speech*. *64*(1), 52–72. https://doi.org/10.1177/0023830920914339

Occhino, C. (2017). An introduction to embodied cognitive phonology: Claw-5 handshape distribution in ASL and Libras. *Complutense Journal of English Studies*. *25*, 69–103. https://doi.org/10.5209/cjes.57198

Ortega, G. (2017). Iconicity and sign lexical acquisition: A review. *Frontiers in Psychology*, *8*, 1280. https://doi.org/10.3389/fpsyg.2017.01280

Ortega, G., & Morgan, G. (2015a). Input processing at first exposure to a sign language. *Second Language Research*, *19*(10), 443–463. https://doi.org/10.1177/0267658315576822

Ortega, G., & Morgan, G. (2015b). Phonological development in hearing learners of a sign language: The influence of phonological parameters, sign complexity, and iconicity. *Language Learning*, *65*(3), 660–688. https://doi.org/10.1111/lang.12123

Ortega, G., & Özyürek, A. (2013). Gesture-sign interface in hearing non-signers'first exposure to sign. *Tilburg Gesture Meeting*, 1–5.

Ortega, G., & Özyürek, A. (2019). Systematic mappings between semantic categories and types of iconic representations in the manual modality: A normed database of silent gesture. *Behavior Research Methods*, *52*(1), 51–67. https://doi.org/10.3758/s13428-019-01204-6

Ortega, G., Özyürek, A., & Peeters, D. (2020). Iconic gestures serve as manual cognates in hearing second language learners of a sign language: An ERP study. *Journal of Experimental Psychology: Learning Memory and Cognition*, *46*(3), 403–415. https://doi.org/10.1037/xlm0000729

Ortega, G., Schiefner, A., & Özyürek, A. (2019). Hearing non-signers use their gestures to predict iconic form-meaning mappings at first exposure to signs. *Cognition*, *191*, Article 103996. https://doi.org/10.1016/j.cognition.2019.06.008

Padden, C., Meir, I., Hwang, S.-O., Lepic, R., Seegers, S., & Sampson, T. (2013). Patterned iconicity in sign language lexicons. *Gesture*, *13*(3), 287–305. https://doi.org/10.1075/gest.13.3.03pad

Peeters, D. (2016). Processing consequences of onomatopoeic iconicity in spoken language comprehension. In *38th Annual Meeting of the Cognitive Science Society (Cogsci 2016)* (pp. 1637–1642). Cognitive Science Society.

Perniss, P., Özyürek, A., & Morgan, G. (2015). The influence of the visual modality on language structure and conventionalization: Insights from sign language and gesture. *Topics in Cognitive Science*, *7*, 2–11. https://doi.org/10.1111/tops.12127

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, *1*, 227. https://doi.org/10.3389/fpsyg.2010.00227

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity : From a world of experience to the experience of language. Philosophical transactions of the Royal Society of London. *Series B, Biological Sciences*, *369*, Article 20130300. https://doi.org/10.1098/rstb.2014.0179

Pietrandrea, P. (2002). Iconicity and arbitrariness in Italian Sign Language. *Sign Language Studies*, *2*(3), 296–321. https://doi.org/10.1353/sls.2002.0012

Polich, J. (2009). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Quinto-Pozos, D., & Parrill, F. (2015). Signers and co-speech gesturers adopt similar strategies for portraying viewpoint in narratives. *Topics in Cognitive Science*, *7*, 12–35. https://doi.org/10.1111/tops.12120

R Core Team (2018). *R: A language and environment for statistical computing. R Foundation for Statistical Computing* [Computer software]. https://www.R-project.org/

Sandler, W. (1989). *Phonological representation of the sign: Linearity and nonlinearity in American Sign Language*. Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110250473

Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge University Press. https://doi.org/10.1017/CBO9781139163910

Schönström, K. (2021). Sign languages and second language acquisition research: An introduction. *Journal of the European Second Language Association*, *5*(1), 30–43. https://doi.org/10.22599/jesla.73

Schwartz, B. D., & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, *12*(1), 40–72. https://doi.org/10.1177/026765839601200103

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *The American Statistician*, *34*(4), 216–221. https://doi.org/10.2307/2684063

Sloetjes, H., & Wittenburg, P. (2018). ELAN (version 5.2). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/

Taub, S. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511509629

Tolentino, L. C., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning*, *64*(2), 279–309. https://doi.org/10.1111/lang.12048

van der Kooij, E. (2002). *Phonological categories in sign language of the Netherlands: The role of phonetic implementation and iconicity*. LOT.

Van Hoey, T., Thompson, A. L., Do, Y., & Dingemanse, M. (2023). Iconicity in ideophones: Guessing, memorizing, and reassessing. *Cognitive Science*, *47*(4), Article e13268. https://doi.org/10.1111/cogs.13268

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Vinson, D. P., Cormier, K., Denmark, T., Schembri, A., & Vigliocco, G. (2008). The British Sign Language (BSL) norms for age of acquisition, familiarity, and iconicity. *Behavior Research Methods*, *40*(4), 1079–1087. https://doi.org/10.3758/brm.40.4.1079

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Accessible Summary

Appendix S1. Iconicity Ratings of the Stimulus Materials.