Iconic Gestures Serve as Manual Cognates in Hearing Second Language Learners of a Sign Language: An ERP Study

Gerardo Ortega  
University of Birmingham

Aslı Özyürek  
Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

David Peeters  
Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands, and Tilburg University

When learning a second spoken language, cognates, words overlapping in form and meaning with one’s native language, help breaking into the language one wishes to acquire. But what happens when the to-be-acquired second language is a sign language? We tested whether hearing nonsigners rely on their gestural repertoire at first exposure to a sign language. Participants saw iconic signs with high and low overlap with the form of iconic gestures while electrophysiological brain activity was recorded. Upon first exposure, signs with low overlap with gestures elicited enhanced positive amplitude in the P3a component compared to signs with high overlap. This effect disappeared after a training session. We conclude that nonsigners generate expectations about the form of iconic signs never seen before based on their implicit knowledge of gestures, even without having to produce them. Learners thus draw from any available semiotic resources when acquiring a second language, and not only from their linguistic experience.

Keywords: sign language, gesture, iconicity, ERPs, second language acquisition

Native speakers of English will not have difficulty understanding Dutch words like hotel or ocean because their translation equivalents have similar or even identical forms in English and other languages. Such cognates, words that overlap in form and meaning between one’s first language and a second language, give immediate access to the meaning of words never seen before (Hall, 2002). But what happens when the target language is a sign language, the manual–visual languages of Deaf communities? The modality differences between speech (aural–oral) and sign language, the manual–visual languages of Deaf communities? The modality differences between speech (aural–oral) and sign (manual–visual) do not allow hearing adults to match the spoken words they know with the structure of to-be-acquired signs. As a result, one could assume that this population cannot alleviate some of the burden to establish form-meaning associations between the target sign and a word from their native language.

However, people do have at their disposal a repertoire of gestures that are commonly used in face-to-face interaction (Kendon, 2004; McNeill, 1992). Silent gestures in particular, those produced when spoken language is not possible or allowed, are a unique communicative tool that conveys rich visual information in a single hand configuration (Goldin-Meadow & Brentari, 2017). Silent gestures are different from cospeech gestures in that they do not seem to be heavily influenced by speakers’ language Goldin-Meadow & Brentari, 2017. To some degree, they exhibit systematic forms within a community of speakers (Ortega & Özyürek, 2019; Van Nispen, Van De Sandt-Koenderman, & Krahmer, 2017). They do not have a linguistic mental representation akin to that of signs in sign languages (i.e., they do not consist of sub-lexical constituents), but they may have some form of mental representation that maps onto existing schemas (Kita, Alibali, & Chu, 2017; Labeye, Oker, Badard, & Versace, 2008; Van Nispen et al., 2017).

In certain cases, gestures may overlap in form with signs because of similar iconic mappings of the concepts they represent (Kendon, 2008; Müller, 2018; Perniss, Özyürek, & Morgan, 2015; Wilcox, 2004). For instance, hearing nonsigners depicting a helicopter in silent gesture may come up with manual forms with a strong resemblance to the conventional sign HELICOPTER used by Deaf people in some sign languages (Figure 1A). It is an intriguing, but currently untested, question, whether hearing nonsigning adults implicitly exploit their repertoire of iconic gestures at first exposure to a sign language. This possibility would extend previous research by showing that gesture assists not only in the...
acquisition of a second spoken language (Kelly, McDevitt, & Esch, 2009), but also in the acquisition of a sign language as a second language. Importantly, it would suggest that learners resort not only to their mother tongue at the earliest stages of second language learning, but also to other nonlinguistic semiotic tools to support vocabulary learning. Clearly, there are significant differences between sign languages and iconic gestures. Sign languages are real linguistic systems with the same level of organization as spoken languages (Sandler & Lillo-Martin, 2006), and Deaf signers process them through the decomposition of signs’ sublexical constituents (Carreiras, Gutiérrez-Sigut, Baquero, & Corina, 2008). In contrast, iconic gestures cannot be regarded as a linguistic system per se because they are holistic units of form-meaning mappings that are spontaneously generated (McNeill, 1992). Nevertheless, both iconic gestures and signs are restricted by the same physical constraints to express a concept iconically in the manual modality (Kendon, 2008; Müller, 2016; Perniss et al., 2015). That is, the body shapes the extent to which signs and gestures can create manual forms that resemble an intended referent. Both iconic gestures and iconic signs seem to originate from the selection of salient features of a concept, the schematization of such features, and their representation with the body (Taub, 2001; Van Nispen et al., 2017). In addition, some have argued that up to two thirds of the lexicon of some sign languages has an iconic motivation (Pietrandrea, 2002). Thus, it is not surprising that iconic signs and iconic gestures may overlap in form and meaning for many concepts.

Inspired by the possible overlap in the form of some silent iconic gestures and conventionalized signs, the current study investigates whether sign-naïve hearing adults exploit their gestures to access the meaning of signs they have never seen before. Electroencephalography (EEG) was used as an online neurophysiological measure of cognitive processes involved at first exposure to a second language in the context of learning. Crucially, this method is taken as a direct measure of online processing at the earliest stages of exposure to a new language, a point in time where behavioral measures might not yet show any effects (Osterhout et al., 2008). In an event-related potential (ERP) experiment, we presented hearing Dutch nonsigners with signs in Sign Language of the Netherlands (NGT). Based on silent gestures produced by a separate group of Dutch speakers, two types of iconic signs were distinguished. 

**Signs with high overlap with gesture**

- to descend
- to-DESCEND
- helicopter
- HELICOPTER

**Signs with low overlap with gesture**

- toothbrush
- TOOTHBRUSH
- butterfly
- BUTTERFLY

*Figure 1.* Systematic silent gestures from Dutch nonsigners and their sign equivalent in Sign Language of the Netherlands (NGT). Panel A shows that hearing nonsigners and Deaf signers produced remarkably similar manual forms for the same concept (i.e., sign–gesture high overlap). Panel B shows that—while nonsigners consistently produce the same gesture for some concepts—these concepts have a different form in sign (i.e., sign–gesture low overlap). These images are used with permission. See the online article for the color version of this figure.
matched in their degree of iconicity so as to ensure that any effect of gesture was not confounded by a potential effect of iconicity.

We predicted that if nonsigning hearing adults exploit their gestural knowledge at first exposure to a sign language, differences in brain activity should be observed as a function of the degree of overlap between their silent gestures and the newly encountered signs—even when participants are not explicitly asked to produce gestures. Any difference in brain activity would be informative in suggesting that gesture gives access to the meaning of signs at first exposure. To learn more about the specific mechanisms underlying the perception and acquisition of signs by novice learners at early exposure to a sign language, we specifically focused on two well-known ERP components: the P300 (P3a) and the N400.

It is well established that stimulus novelty causes enhanced P300 amplitude (Friedman, Cycowicz, & Gaeta, 2001; Polich, 2007). Signs with low overlap with gesture will be novel to our participants, whereas signs with high overlap with gesture may map onto existing gestural schemas. Particularly, modulations of participants, whereas signs with high overlap with gesture may be expected for signs with low overlap with gesture, as these novel stimuli cannot be predicted by our participants at first exposure based on existing schemas (Friedman et al., 2001; Van Petten & Luka, 2012). Therefore, enhanced P3a amplitude for low-overlap (vs. high-overlap) signs would reflect activation of existing gestural schemas for high-overlap signs.

Additionally, we tested for potential sensitivity of the amplitude of the N400 component to overlap between sign and gesture, as this may be taken to reflect different, relevant processes. First, N400 amplitude to individual lexical items in second language has been linked to processing ease, for instance in the context of second-language processing when comparing spoken language cognates to matched control words (Midgley, Holcomb, & Grainger, 2011; Peeters, Dijkstra, & Grainger, 2013). In line with these findings, we hypothesized that if signs with high overlap with gesture are processed more easily compared to signs with low overlap with gesture, reduced amplitude of the N400 component should be observed for the high-overlap condition compared to the low-overlap condition. Second, earlier work has linked N400 amplitude to semantic integration (e.g., van Berkum, Haartman, & Grainger, 2004; Devlin, Eriksen, & Eriksen, 1988). It might be easier to integrate an observed high overlap sign (vs. a low overlap sign) with the corresponding preceding word in our paradigm, because of the availability of a gestural schema for the signs with high overlap with gesture. Third, previous work has linked N400 amplitude to prediction (e.g., Szewczyk & Schriefers, 2018). In our paradigm, based on their gestural repertoire, participants may predict the form of an upcoming sign after having perceived the preceding word. If they would do so, a disconfirmed prediction in the low overlap condition could be reflected in enhanced N400 amplitude. These final two interpretations of N400 amplitude may be less relevant in the context of the current study given that we presented lexical items outside a sentence or discourse context.

We further predicted that these two potential effects may attenuate or even disappear after sign learning once all signs, regardless of their gestural overlap, become tightly linked to their corresponding meaning, as both P3a (Friedman et al., 2001) and N400 amplitude (Osterhout, McLaughlin, Pitkänen, French, & Molinaro, 2006) may reduce with learning.

### Stimuli Selection

The stimuli selection consisted of a two-stage procedure that involved i) collecting a set of iconic gestures that could be generalized across Dutch participants (silent gesture task). These gestures are a subset of a published database of 109 silent gestures (Ortega & Özyürek, 2019). Having collected these gestures, it was possible to ii) carry out a comparison between the form of each systematic gesture with its NGT sign equivalent (gesture-sign cross-comparison). This allowed to have two sets of iconic signs that had high and low resemblance with the iconic gestures collected in step (i).

**Silent gesture task.** Participants of this part of the study consisted of 20 adults (mean age: 27 years; age range: 21–46 years, 10 females), born in the Netherlands and with Dutch as their single native language (none of these participants took part in the later ERP experiment). They were seated in front of a laptop and were instructed to spontaneously come up with a gesture that conveyed the same meaning as a single word (n = 272) presented in written form on the screen. Participants were not allowed to speak or point at any object in the room during the production of gestures, but they could say “pass!” when they could not come up with a gesture. Each trial started with a fixation cross in the middle of the screen (500 ms), followed by a single word in Dutch (4000 ms) during which they had to come up with their gestural rendition. After the 4000 ms had lapsed, the next trial began. This strict timing encouraged participants to come up with their most intuitive response.

Participants’ renditions were coded using the linguistic annotator ELAN Version 4.9.1 (Sloetjes & Wittenburg, 2018). Each gesture or sequence of gestures consisted of a preparation phase, a stroke, and a (partial/full) retraction (Kita, van Rijin, & van der Hulst, 1997). The form of each gesture was further annotated according to an existing coding scheme that describes their forms without relying on written descriptions (Bressem, 2013). This notation system is applied to gestures’ more salient structural features, which are loosely based on the four phonological parameters described for sign languages (Brentari, 1999; van der Kooij, 2002). These features are the configuration of the hand, its orientation, the direction of the movement of the main articulator (i.e., the hand/s), and the location where the gesture takes place. Speed and quality of the movement are additional features considered in this notation system but were not applied in the current study.

Systematicity in gestural productions was operationalized as gestures that at least across 50% of participants (n = 10) shared minimally three out of its four features (i.e., handshape, orientation, movement, and location; Bressem, 2013). If less than 10 participants produced a gesture that had sufficient overlap according to our criteria, then it was considered that the concept did not elicit a systematic gesture and was not included in the collection of systematic gestures. For example, for the concept “butterfly” (vlinder), 11 participants flapped their arms as if personifying the butterfly.

---

1. This study received ethical approval from the Ethics Assessment Committee (EAC) of the Faculty of Arts of Radboud University (ref: MvB14U.015319). We are also indebted to three anonymous reviewers whose suggestions helped us improve the manuscript.
insect themselves, so this rendition was considered a systematic gesture (Figure 1A). In contrast, the concept “to cook” (oken) elicited a wide array of gestural forms that were not homogeneous with at least 10 of the 20 participants. Therefore, this concept was considered not to elicit a systematic gesture and was not included in the set of systematic gestures.

**Gesture-sign cross-comparison.** A Deaf native signer of NGT was recruited as a consultant to record the same 272 concepts used in the silent gesture task in NGT. This Deaf consultant has used NGT all his life, is a qualified sign language teacher, and has been an active member of the Deaf community in the Netherlands. After signing consent forms, he was asked to produce the citation form of each concept with neutral face and without any mouthing so as to avoid giving hints about the meaning of the sign via lip patterns. Once all these signs were recorded, a different group of 20 hearing nonsigning adults (mean age = 21.8 years; age range: 19–32, 14 female) were asked to rate these signs for their degree of meaning transparency (i.e., iconicity ratings). Participants were asked to rate the degree of form-meaning mapping on a 7-point Likert scale while they viewed the sign along with its translation (1: low iconicity, 7: high iconicity). None of these raters took part in the EEG experiment or in the silent gesture task.

In order to establish the degree of form similarity between gestures and signs, we carried out a comparison between the four main features of the systematic gestures from the silent gesture task and the four components of conventionalized NGT signs (i.e., hand configuration, orientation, movement, and location). Two categories were created. **Signs with high gestural overlap** consisted of signs that overlap in at least three out of four constituents. For instance, the NGT sign TO-BREAK falls in this category because all its sublexical constituents overlap with the four features of the elicited systematic gesture. **Signs with low gestural overlap** are signs that differ in two or more of its constituents with the corresponding elicited systematic gesture. The sign BUTTERFLY falls in this category because there is no overlap between sign and gesture in any of the constituents except for the handshape (i.e., extended palm).

In order to ensure that it was the overlap with gesture and not the degree of iconic form-meaning mapping behind any possible effect, we selected signs so that the final set of signs was balanced for degree of iconicity across conditions (high overlap: n = 36, mean rating: 4.77, sd = 1.32; low overlap: n = 36, mean: 4.76, sd = 1.12; t(35) = .032, p = .974). There were 17 one-handed signs in the high overlap condition (19 two-handed signs) and 14 one-handed signs in the low overlap condition (22 two-handed signs). Furthermore, the duration of the videos of the signs did not differ across condition (high overlap: mean duration = 2423 ms, sd = 454.97; low overlap: mean duration = 2611 ms, sd = 637.21; t(35) = -1.417, p = .165). The Dutch words presented prior to the signs were controlled for length, frequency, and concreteness. See Appendices A and B for a complete list of the attributes of all stimulus materials.

**Event-Related Potential Experiment**

**Participants.** Twenty-nine right-handed participants (mean age 22 years, range: 19–29 years, 19 females) participated in the ERP experiment. All participants were Dutch, studying in Nijmegen, and Dutch was their single native language. None of these participants took part in the silent gesture task or in the iconicity ratings task, and they reported not having any experience with any sign language. EEG data from one participant was not analyzed due to a large number of EEG artifacts visible during the recording session. Data from four participants was excluded from the ERP analysis due to a large number of artifacts that had to be removed during the preprocessing stage. In sum, data from 24 participants (mean age 20 years, range 19–29 years, 14 females) entered the ERP analyses. Data from all 29 participants were included in the behavioral analyses.

**Procedure.** After providing informed consent, participants were instructed that they would take part in a four-block sign learning experiment. Each block was preceded by 5 practice trials, using stimuli that were not used in the experimental trials.

1. **First exposure (block 1):** The aim of this block was to measure ERPs prior to any sign language learning experience to determine whether the brain signal was sensitive to signs’ similarities with gestures at first exposure to sign language. Participants were seated in front of a 20-in. Samsung computer monitor on which the stimulus materials (36 trials per condition) were shown. Distance between participants and the screen was 100 cm. Each trial consisted of a fixation cross in the middle of the screen (500 ms), which was followed by a printed word in Dutch (e.g., vlinder, butterfly) that remained on the screen for 1000 ms. After this time had lapsed, another fixation cross appeared in the middle of the screen (500 ms) followed by the NGT sign equivalent of the Dutch word (e.g., the sign BUTTERFLY) in a video (14 × 8 cm). After the sign had played in full, the next trial began. ERPs were time-locked to video onset. In addition, the sign onset, defined as the instance when the hand reached its location in the first fully formed handshape (Crasborn et al., 2015), was determined by the first author using the frame-by-frame feature of the linguistic annotator software ELAN Version 4.9.1 (Sloetjes & Wittenburg, 2018). On average, the sign onset was 460.8 ms after video onset. Signs were presented in randomized order. Participants were instructed to pay close attention to the words and signs but were not required to perform any task during the presentation of the stimuli.

2. **Learning phase (block 2):** Participants were told they were going to be taught the same signs from the first block. Each trial consisted of a fixation cross (500 ms), followed by the video of a sign with a word in Dutch (the translation of the sign) presented under the corresponding video for the duration of the video. This was then followed by a 3000-ms blank screen. This trial was repeated three times for each sign, and after each single presentation of the sign participants were required to imitate it as accurately as possible so as to encourage learning. Once the sign had been presented and imitated sequentially three times, the next trial with a different sign began. Sign repetitions were video recorded and no ERPs were measured.

3. **Postlearning exposure (block 3):** The aim of this block was to determine whether there was a significant dif-
ference in brain responses after participants had received relatively extensive training with the signs. The structure of each trial was the same as in first exposure (block 1), but the same signs were presented in a different randomized order.

4. Testing phase (block 4): Participants’ ability to retain the signs was assessed in this block. In each trial, a fixation cross was presented in the middle of the screen for 200 ms, followed by a blank screen (200 ms), followed by a printed word (6000 ms), which was the Dutch translation of one of the signs presented throughout the experiment. Participants were instructed to produce the NGT sign equivalent while each of the 72 concepts were randomly presented on the screen. There was no feedback, and participants could say “pass” to indicate that they could not remember the form of the sign. We were interested in getting intuitive responses, so we imposed a strict timing, and after the 6000 ms lapsed the next trial began. Sign productions were video recorded and no ERPs were measured.

This four-block design allowed for a manipulation of gestural overlap (high overlap vs. low overlap between the presented signs and the participants’ gestural repertoire) and learning (block 1 vs. block 3).

Electrophysiological Recording and Analysis (Block 1 and Block 3)

Participants’ EEGs were recorded continuously from 59 active electrodes (Brain Products, Munich, Germany) held in place on the scalp by an elastic cap (Neuroscan, Singen, Germany). In addition to the 59 scalp sites (see Figures 2 and 3 for equidistant electrode montage), three external electrodes were attached to record participants’ electrooculogram (EOG), one below the left eye (to monitor for vertical eye movement/blinks), and two on the lateral canthi next to the left and right eye (to monitor for horizontal eye movements). One additional electrode was placed over the left mastoid bone and one over the right mastoid bone. Electrode impedances were kept below 5 kΩ. The continuous EEG was recorded with a sampling rate of 500 Hz, a low cut-off filter of 0.01 Hz, and a high cut-off filter of 200 Hz. All electrode sites were referenced online to the electrode placed over the left mastoid and re-referenced offline to the average of the electrodes placed over left and right mastoids.

Preprocessing and ERP analyses were carried out in Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). Raw EEG data was low-pass filtered offline at 40 Hz. Epochs from 100 ms preceding video onset to 1400 ms after picture onset were selected. The 100-ms

Figure 2. Grand average waveforms time-locked to video-onset comparing high overlap to low overlap trials in the first block. P300 and N400 time windows were calculated from sign onset, that is, the offset of the sign preparation phase. The topographic plot shows the widespread corresponding voltage difference between the two conditions between 700 and 800 ms after video onset. See the online article for the color version of this figure.
pre-video period served as a baseline. Trials containing ocular or muscular artifacts were not taken into consideration in the averaging process. Data from two left posterior electrodes were not included in the analyses due to malfunctioning of the electrodes during data collection. The number of rejected trials did not differ significantly across conditions (remaining trials: block 1 high overlap = 620; low overlap = 601; block 3 high overlap = 572; low overlap = 595).

Event-related potential data were analyzed using cluster-based permutation tests (Maris & Oostenveld, 2007) on two epochs of interest: the P300 time-window (700–800 ms after video onset, corresponding to 240–340 ms after the sign onset) and the N400 time window (800–1000 ms after video onset, corresponding to 340–540 ms after the onset of the sign). An additional analysis on the interval between video-onset and the onset of the signs’ meaningful part (0–460 ms after video-onset) revealed no significant differences across conditions in either block (both ps > .287), indicating no differential processing of the initial, nonmeaningful parts of the signs presented in the videos.

The cluster-based, nonparametric, data-driven approach to data analysis has the advantage of controlling for the familywise error rate that arises when an effect of interest is evaluated at multiple time points and electrodes (Maris & Oostenveld, 2007), which has often led to a multiple comparisons problem in electrophysiological data analysis (Maris, 2012). To describe the cluster-based permutation approach briefly, for each data point (electrode by time), a simple dependent-samples t test comparing two conditions was performed. Adjacent data points (spatial or temporal) exceeding an alpha level of .05 were grouped into clusters. For all clusters (both positive and negative), the sum of the t statistics was used in the cluster-level test statistic. A null distribution was then calculated that assumed no difference between conditions (3,000 randomizations, calculating the largest cluster-level statistic for each randomization), after which the actually observed cluster-level statistics were compared against this null distribution. Clusters falling in the highest or lowest 2.5% percentile were considered significant (Bonferroni corrected; a p value < .025 corresponds to a significant effect).

**Sign Imitation (Block 2) and Sign Production Analysis (Block 4)**

In order to obtain a baseline of accuracy in sign production we looked at participants’ sign articulation in block 2 (learning phase). Participants imitated each sign three times during this block, so we investigated their first rendition, which was their first ever attempt to execute the signs seen. We compared this baseline with sign production in block 4 (testing phase), where participants had to produce the sign from memory. Renditions across blocks were off-line coded using the linguistic annotator ELAN Version 4.9.1.
Behavioral Results (Training Phase—Block 2 and Test Phase—Block 4)

In the training phase, participants were equally accurate at imitating high overlap signs ($M = 94.5, sd = .23$) compared to low overlap signs ($M = 94.5, sd = .23$). In the test phase, participants were numerically slightly better in producing high overlap signs ($M = 97.5, sd = .16$) compared to low overlap signs ($M = 96.3, sd = .19$). The binomial logistic regression analysis showed no significant main effect of condition ($p = .14$), a significant main effect of block ($p = .0003$), and no significant interaction effect between condition and block ($p = .19$). Thus, participants were significantly more accurate at producing signs after training compared to imitating signs during training.

Electrophysiological Results (Blocks 1 and 3)

Event-related potentials were time-locked to the onset of the sign videos, to allow for a stable baseline period across conditions. P300 and N400 time windows were calculated on the basis of the onset of the sign.

**P300 time window.** Cluster-based permutation tests comparing the two conditions in the P300 time-window revealed a significant difference ($p = .017$) between the high overlap condition and the low overlap condition for block 1. This difference, reflecting a significantly higher positive amplitude for the low overlap condition compared to the high overlap condition, was observed during the full epoch (700–800 ms) and widespread over the scalp (i.e., observed in 39 out of 57 analyzed electrodes). Figure 2 illustrates this slightly left-lateralized and anteriorly dominant effect. No significant difference between conditions was observed in the same analysis for block 3 ($p > .195$; see Figure 3 for comparison with Figure 2).

**N400 time-window.** Cluster-based permutation tests comparing the two overlap conditions in the N400 time window revealed no significant effects. No statistical differences were observed in this time window for block 1 ($p > .133$) nor for block 3 ($p > .417$).

An additional ERP analysis comparing block 1 (first exposure) to block 3 (postlearning exposure) can be found in Figure C1 in Appendix C.

Discussion

Words that overlap in form and meaning with words in one’s native language (i.e., cognates) help to break into a second language one wishes to acquire (Hall, 2002). But what happens when the to-be-acquired second language is a sign language? Because of the modality differences between speech and sign, one would intuitively assume that there are no such cognates. However, given that iconic signs and iconic gestures may overlap in form and meaning for many concepts due to their shared manual modality, the current study tested whether hearing nonsigners access their knowledge of gestures at first exposure to a sign language. Participants saw iconic signs with high and low overlap with gestures while their electrophysiological brain activity was recorded. We observed that, upon first exposure, signs with low overlap with gestures elicited enhanced positive amplitude in the P300 time window compared to signs with high overlap with gestures. There were no differences between both types of signs in the amplitude of the N400 component. After participants had watched and imitated each sign three times, ERP recordings showed no processing differences in the P300 or N400 time windows. Importantly, participants learned both types of signs (high overlap and low overlap) with equal ease at the end of the experiment. Our results indicate that at first exposure to a sign language, nonsigners activate their gestural knowledge, when generating expectations about the form of signs.

Due to its anterior distribution over the scalp, we interpret the observed effect in the P300 time window as a P3a effect (Friedman et al., 2001; Polich, 2007). As mentioned in the introduction, enhanced amplitude in this component has been consistently linked to stimulus novelty (Friedman et al., 2001; Polich, 2007). At first exposure, signs with low overlap with gesture were novel to our participants, whereas signs with high overlap with gesture will have mapped onto existing gestural schemas. As the low-overlap signs did not map onto participants’ gestural schemas, any prediction based on reading the preceding, corresponding word would have been followed by a disconfirmation. This finding is therefore also in line with earlier work arguing that P300 amplitude may index (dis)confirmed expectations about upcoming stimuli (Van Petten & Luka, 2012).

Prima facie, it is surprising that we did not observe any differences in the N400 time window, given that studies in spoken languages have consistently shown N400 effects for cognates compared to noncognate control words (e.g., Midgley et al., 2011; Peeters et al., 2013). Spoken language research has argued that the cognate status of a word facilitates mapping the encountered word form to its meaning. We note two critical differences between the present study and earlier research reporting cognate N400 effects in the domain of spoken language. First, our sign stimuli in both the high and low overlap condition were highly iconic, whereas spoken language research on cognates typically uses word stimuli that mostly have an arbitrary link between form and meaning. It is an exciting possibility that iconicity may facilitate form-meaning mapping in the acquisition of a second language in sign (Baus, Carreiras, & Emmorey, 2013) and spoken languages (Deconinck, Boers, & Eyckmans, 2017). Second, spoken language research on cognates typically studies bilingual participants who already have quite some knowledge of the foreign language they are tested in (Peeters, Vanlangendonck, Rueschemeyer, & Dijkstra, 2019), whereas our participants had no knowledge of sign language prior to the experiment. As such, future research should investigate directly if learners of a second spoken language also exhibit enhanced
P3a amplitude for control words compared to cognates at first exposure to a foreign spoken language.

Participants were very accurate at producing signs in the behavioral task, and there were no statistical differences as a function of gestural overlap in sign production in the training (block 2) and testing phase (block 4). In the training phase, participants imitated signs immediately after observing them on the screen, so this resulted in high degree of accuracy under our coding scheme. Participants occasionally produced some of the errors that have been reported in the literature, such as inaccurate hand configuration (Ortega & Morgan, 2015) and production of the mirror image of signs (Rosen, 2004), but these renditions were still intelligible. Analysis of the renditions produced during the testing phase showed that, in general, having observed each sign five times during the experiment led to successful sign learning. We did see, however, few instances of gestural interference during sign production. For instance, when attempting to recall the sign BUTTERFLY, one participant produced the gesture documented in the silent gesture task (see Figure 1). That said, this was not a recurrent mistake. Future research could explore gestural interference in sign learning over longer periods of time, for instance by testing participants at a later stage (e.g., a week) after training.

Earlier claims about differences between gestures and signs are currently being reconsidered given the growing evidence showing that both forms of manual communication share more similarities than previously assumed (Kendon, 2008; Müller, 2018; Perniss et al., 2015). The systematicity observed in the iconic silent gestures across participants, as well as their overlap with many signs for the same concept, suggest that in many instances both hearing and Deaf participants employ similar strategies to depict certain concepts iconically, as in the cases of high overlap condition (although only signs are part of a conventionalized lexicon). We suggest that the conceptual representations shared by hearing speakers and Deaf signers, as well as the physical affordances on the manual modality, result in a mutual representation of form features in gesture. For instance, Muller, Cienk, E. Fricke, S. Ladewig, D. McNeill, & S. Tessential (Eds.), Body - Language - Communication: An international handbook on multimodality in human interaction (pp. 1079–1098). Berlin, Germany: De Gruyter Mouton. http://dx.doi.org/10.1515/9783110261318.1079


ICONIC GESTURES ACT AS MANUAL COGNATES IN SIGN LEARNING

Bielefeld, Germany, September 17–19, 1997, proceedings (pp. 23–35). Berlin, Germany: Springer-Verlag.


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


(Appendices follow)
## Appendix A

### Measures of Length, Frequency, and Concreteness for the Dutch Words

<table>
<thead>
<tr>
<th>Dutch</th>
<th>English</th>
<th>Length</th>
<th>Frequency</th>
<th>Concreteness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High overlap</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Knippen</td>
<td>to cut (scissors)</td>
<td>7</td>
<td>9.08</td>
<td>4.60</td>
</tr>
<tr>
<td>2 Oppompen</td>
<td>to pump</td>
<td>8</td>
<td>.32</td>
<td>3.80</td>
</tr>
<tr>
<td>3 Vogel</td>
<td>bird</td>
<td>5</td>
<td>32.27</td>
<td>4.87</td>
</tr>
<tr>
<td>4 Baby</td>
<td>baby</td>
<td>4</td>
<td>151.80</td>
<td>4.67</td>
</tr>
<tr>
<td>5 Telefoon</td>
<td>telephone</td>
<td>8</td>
<td>156.92</td>
<td>4.87</td>
</tr>
<tr>
<td>6 Lepel</td>
<td>spoon</td>
<td>5</td>
<td>5.01</td>
<td>4.93</td>
</tr>
<tr>
<td>7 Handdoek</td>
<td>towel</td>
<td>8</td>
<td>10.50</td>
<td>5.00</td>
</tr>
<tr>
<td>8 Piano</td>
<td>piano</td>
<td>5</td>
<td>14.11</td>
<td>4.93</td>
</tr>
<tr>
<td>9 Auto</td>
<td>car</td>
<td>4</td>
<td>349.11</td>
<td>5.00</td>
</tr>
<tr>
<td>10 Rekenmachine</td>
<td>calculator</td>
<td>12</td>
<td>.57</td>
<td>4.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Low overlap</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 snijden</td>
<td>to cut (knife)</td>
<td>7</td>
<td>20.65</td>
<td>4.13</td>
</tr>
<tr>
<td>2 stelen</td>
<td>to steal</td>
<td>6</td>
<td>59.82</td>
<td>3.33</td>
</tr>
<tr>
<td>3 vliegen</td>
<td>to fly</td>
<td>7</td>
<td>89.69</td>
<td>4.27</td>
</tr>
<tr>
<td>4 olifant</td>
<td>elephant</td>
<td>7</td>
<td>12.01</td>
<td>4.93</td>
</tr>
<tr>
<td>5 adelaar</td>
<td>eagle</td>
<td>7</td>
<td>3.93</td>
<td>4.93</td>
</tr>
<tr>
<td>6 laptop</td>
<td>laptop</td>
<td>6</td>
<td>5.88</td>
<td>4.93</td>
</tr>
<tr>
<td>7 doos</td>
<td>box</td>
<td>4</td>
<td>38.28</td>
<td>4.87</td>
</tr>
<tr>
<td>8 nieten</td>
<td>to staple</td>
<td>6</td>
<td>3.16</td>
<td>3.93</td>
</tr>
<tr>
<td>9 slang</td>
<td>snake</td>
<td>5</td>
<td>21.59</td>
<td>4.87</td>
</tr>
<tr>
<td>10 paraplu</td>
<td>umbrella</td>
<td>7</td>
<td>3.43</td>
<td>4.80</td>
</tr>
<tr>
<td>11 rammelaar</td>
<td>rattle</td>
<td>9</td>
<td>23.67</td>
<td>4.67</td>
</tr>
<tr>
<td>12 kloppen</td>
<td>to knock</td>
<td>7</td>
<td>23.39</td>
<td>4.53</td>
</tr>
<tr>
<td>13 jongleren</td>
<td>to juggle</td>
<td>9</td>
<td>56.89</td>
<td>3.80</td>
</tr>
<tr>
<td>14 botsen</td>
<td>to crash</td>
<td>6</td>
<td>56.67</td>
<td>3.80</td>
</tr>
<tr>
<td>15 vlinder</td>
<td>butterfly</td>
<td>7</td>
<td>6.13</td>
<td>5.00</td>
</tr>
<tr>
<td>16 ruitenwisser</td>
<td>windscreen wiper</td>
<td>12</td>
<td>10.56</td>
<td>4.47</td>
</tr>
<tr>
<td>17 aankleden</td>
<td>to put clothes on</td>
<td>9</td>
<td>9.33</td>
<td>3.53</td>
</tr>
</tbody>
</table>

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

(Appendices continue)
### Appendix A (continued)

<table>
<thead>
<tr>
<th>Dutch</th>
<th>English</th>
<th>Length</th>
<th>Frequency</th>
<th>Concreteness</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>schildpad</td>
<td>turtle</td>
<td>9</td>
<td>44.12</td>
</tr>
<tr>
<td>19</td>
<td>kat</td>
<td>cat</td>
<td>3</td>
<td>52.85</td>
</tr>
<tr>
<td>20</td>
<td>konijn</td>
<td>rabbit</td>
<td>6</td>
<td>18.87</td>
</tr>
<tr>
<td>21</td>
<td>deur</td>
<td>door</td>
<td>4</td>
<td>247.48</td>
</tr>
<tr>
<td>22</td>
<td>fles</td>
<td>bottle</td>
<td>4</td>
<td>45.71</td>
</tr>
<tr>
<td>23</td>
<td>champignons</td>
<td>mushroom</td>
<td>10</td>
<td>22.09</td>
</tr>
<tr>
<td>24</td>
<td>bloem</td>
<td>flower</td>
<td>5</td>
<td>13.49</td>
</tr>
<tr>
<td>25</td>
<td>bed</td>
<td>bed</td>
<td>3</td>
<td>239.93</td>
</tr>
<tr>
<td>26</td>
<td>restaurant</td>
<td>restaurant</td>
<td>10</td>
<td>41.78</td>
</tr>
<tr>
<td>27</td>
<td>kip</td>
<td>chicken</td>
<td>3</td>
<td>37.89</td>
</tr>
<tr>
<td>28</td>
<td>bal</td>
<td>ball</td>
<td>3</td>
<td>80.63</td>
</tr>
<tr>
<td>29</td>
<td>tandenborstel</td>
<td>toothbrush</td>
<td>13</td>
<td>4.16</td>
</tr>
<tr>
<td>30</td>
<td>rollator</td>
<td>zimmer frame</td>
<td>8</td>
<td>.05</td>
</tr>
<tr>
<td>31</td>
<td>rolstoel</td>
<td>wheelchair</td>
<td>8</td>
<td>8.37</td>
</tr>
<tr>
<td>32</td>
<td>pistool</td>
<td>pistol</td>
<td>7</td>
<td>102.63</td>
</tr>
<tr>
<td>33</td>
<td>huilen</td>
<td>to cry</td>
<td>6</td>
<td>54.52</td>
</tr>
<tr>
<td>34</td>
<td>injectoren</td>
<td>to inject</td>
<td>10</td>
<td>1.81</td>
</tr>
<tr>
<td>35</td>
<td>vliegtuig</td>
<td>plane</td>
<td>9</td>
<td>89.92</td>
</tr>
<tr>
<td>36</td>
<td>pinguin</td>
<td>penguin</td>
<td>7</td>
<td>34.88</td>
</tr>
</tbody>
</table>

Mean:

<table>
<thead>
<tr>
<th>Length</th>
<th>Frequency</th>
<th>Concreteness</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.92</td>
<td>43.98</td>
<td>4.56</td>
</tr>
</tbody>
</table>

### Appendix B

**Word-Sign Pairs, Sign Iconicity Ratings, and Number of Hands per Signs**

<table>
<thead>
<tr>
<th>Dutch</th>
<th>NGT sign (gloss)</th>
<th>No. hands</th>
<th>Iconicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>High overlap</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>knippen</td>
<td>1</td>
<td>6.61</td>
</tr>
<tr>
<td>2</td>
<td>oppompen</td>
<td>2</td>
<td>5.16</td>
</tr>
<tr>
<td>3</td>
<td>vogel</td>
<td>1</td>
<td>6.42</td>
</tr>
<tr>
<td>4</td>
<td>baby</td>
<td>2</td>
<td>6.39</td>
</tr>
<tr>
<td>5</td>
<td>telefoon</td>
<td>1</td>
<td>6.22</td>
</tr>
<tr>
<td>6</td>
<td>lepel</td>
<td>1</td>
<td>5.33</td>
</tr>
<tr>
<td>7</td>
<td>handdoek</td>
<td>2</td>
<td>5.11</td>
</tr>
<tr>
<td>8</td>
<td>piano</td>
<td>2</td>
<td>6.05</td>
</tr>
<tr>
<td>9</td>
<td>auto</td>
<td>2</td>
<td>4.74</td>
</tr>
<tr>
<td>10</td>
<td>rekenmachine</td>
<td>2</td>
<td>4.78</td>
</tr>
<tr>
<td>11</td>
<td>kameel</td>
<td>1</td>
<td>5.42</td>
</tr>
<tr>
<td>12</td>
<td>sleutel</td>
<td>1</td>
<td>6.26</td>
</tr>
<tr>
<td>13</td>
<td>wringen</td>
<td>2</td>
<td>6.12</td>
</tr>
<tr>
<td>14</td>
<td>breken</td>
<td>2</td>
<td>6.79</td>
</tr>
<tr>
<td>15</td>
<td>kiwi</td>
<td>2</td>
<td>2.06</td>
</tr>
<tr>
<td>16</td>
<td>melk</td>
<td>2</td>
<td>2.68</td>
</tr>
<tr>
<td>17</td>
<td>omhoog lopen</td>
<td>1</td>
<td>2.72</td>
</tr>
<tr>
<td>18</td>
<td>omlaag lopen</td>
<td>1</td>
<td>2.44</td>
</tr>
<tr>
<td>19</td>
<td>koffer</td>
<td>1</td>
<td>3.61</td>
</tr>
<tr>
<td>20</td>
<td>helikopter</td>
<td>1</td>
<td>4.58</td>
</tr>
<tr>
<td>21</td>
<td>spin</td>
<td>1</td>
<td>4.53</td>
</tr>
<tr>
<td>22</td>
<td>asp</td>
<td>2</td>
<td>4.50</td>
</tr>
<tr>
<td>23</td>
<td>gordijnen</td>
<td>2</td>
<td>4.74</td>
</tr>
<tr>
<td>24</td>
<td>appel</td>
<td>1</td>
<td>2.53</td>
</tr>
<tr>
<td>25</td>
<td>gevangenis</td>
<td>2</td>
<td>3.84</td>
</tr>
<tr>
<td>26</td>
<td>sms' en</td>
<td>2</td>
<td>4.17</td>
</tr>
<tr>
<td>27</td>
<td>uitgummen</td>
<td>1</td>
<td>4.21</td>
</tr>
</tbody>
</table>

(Appendices continue)
<table>
<thead>
<tr>
<th>Dutch</th>
<th>NGT sign (gloss)</th>
<th>No. hands</th>
<th>Iconicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>boor</td>
<td>DRILL</td>
<td>1</td>
</tr>
<tr>
<td>29</td>
<td>deken</td>
<td>BLANKET</td>
<td>2</td>
</tr>
<tr>
<td>30</td>
<td>hert</td>
<td>DEER</td>
<td>2</td>
</tr>
<tr>
<td>31</td>
<td>brug</td>
<td>BRIDGE</td>
<td>1</td>
</tr>
<tr>
<td>32</td>
<td>huis</td>
<td>HOUSE</td>
<td>2</td>
</tr>
<tr>
<td>33</td>
<td>kreeft</td>
<td>LOBSTER</td>
<td>2</td>
</tr>
<tr>
<td>34</td>
<td>slaan</td>
<td>TO-SLAP</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>zwemmen</td>
<td>TO-SWIM</td>
<td>2</td>
</tr>
<tr>
<td>36</td>
<td>fiets</td>
<td>BIKE</td>
<td>1</td>
</tr>
</tbody>
</table>

Mean: 4.77

<table>
<thead>
<tr>
<th>Dutch</th>
<th>Iconicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>5.15</td>
</tr>
<tr>
<td>38</td>
<td>5.23</td>
</tr>
<tr>
<td>39</td>
<td>5.24</td>
</tr>
<tr>
<td>40</td>
<td>5.26</td>
</tr>
</tbody>
</table>

Mean: 5.25

(Appendices continue)
Appendix C
Additional ERP Analysis

An additional ERP analysis was carried out comparing block 1 (first exposure) to block 3 (post-learning exposure). Because we had no specific predictions for this comparison, as it was planned on the basis of reviewers’ suggestions, we carried out an analysis on the entire time window between sign onset (460 ms) and video offset (1400 ms). A cluster-based permutation test (same parameters used as in the analyses described in the main text) comparing the two blocks revealed a significant difference ($p < .001$) between the two blocks. This difference, reflecting a sustained positivity for the signs when presented in block 1 compared to the same signs when presented in block 3, was observed during the full epoch (460–1400 ms) and widespread over the scalp (i.e., observed in 43 out of 57 analyzed electrodes). This difference was statistically independent from the signs’ gestural overlap, i.e., there was no interaction between block (block 1 vs. block 3) and gesture overlap (high overlap vs. low overlap).

Received October 11, 2018
Revision received April 15, 2019
Accepted April 16, 2019

Figure C1. Grand average waveforms time-locked to video onset comparing the event-related potentials (ERPs) elicited in block 1 to those from block 3, collapsed over the two gestural overlap conditions. The topographic plot shows the widespread corresponding voltage difference between the two blocks between sign onset (460 ms) and video offset (1400 ms). Overall, signs in block 1 elicited a sustained positivity compared to signs in block 3. See the online article for the color version of this figure.