

Evidence for visual simulation during sign language processing

Ortega Delgado, Gerardo; Ostarek, Markus

DOI:

[10.1037/xge0001041](https://doi.org/10.1037/xge0001041)

License:

None: All rights reserved

Document Version

Peer reviewed version

Citation for published version (Harvard):

Ortega Delgado, G & Ostarek, M 2021, 'Evidence for visual simulation during sign language processing', *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0001041>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

© American Psychological Association, 2021. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. The final article is available, upon publication, at: <https://doi.org/10.1037/xge0001041>

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Evidence for visual simulation during sign language processing

Abstract

What are the mental processes that allow us to understand the meaning of words? A large body of evidence suggests that when we process speech we engage a process of perceptual simulation whereby sensory-motor states are activated as a source of semantic information. But does the same process take place when words are expressed with the hands and perceived through the eyes? To-date, it is not known whether perceptual simulation is also observed in sign languages, the manual-visual languages of deaf communities. Continuous flash suppression is a method that addresses this question by measuring the effect of language on detection sensitivity to images that are suppressed from awareness. In spoken languages, it has been reported that listening to a word (e.g., 'bottle') activates visual features of an object (e.g., the shape of a bottle) and this in turn facilitates image detection. An interesting but untested question is whether the same process takes place when deaf signers see signs. We found that processing signs boosted the detection of congruent images, making otherwise invisible pictures visible. A boost of visual processing was only observed for signers but not for hearing non-signers, suggesting that the penetration of the visual system through signs requires a fully-fledged manual language. Iconicity did not modulate the effect of signs on detection, neither in signers nor in hearing non-signers. This suggests that visual simulation during language processing occurs regardless of language modality (sign vs speech) or iconicity, pointing to a foundational role of simulation for language comprehension.

Keywords: perceptual simulation, language processing, continuous flash suppression, sign languages, iconicity

Introduction

What are the mental processes that allow us to comprehend the meaning of words? The traditional view that semantic processing relies entirely on abstract, amodal representations is falling out of favour as multiple studies have demonstrated that access to the meanings of words involves sensory processes resembling those during our multimodal experience with the world (Barsalou, 2008; Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012; Ralph, Jeffries, Patterson, & Rogers, 2017). For instance, reading action verbs like ‘kick’ activates brain regions engaged in the execution of such actions (Hauk, Johnsrude, & Pulvermüller, 2004; Shtyrov, Butorina, Nikolaeva, & Stroganova, 2014). Importantly for the present paper, there is evidence that words referring to concrete objects activate visual representations of those objects (Correia et al., 2014; Lewis & Poeppel, 2014) and facilitate the subsequent visual processing of corresponding pictures (Boutonnet & Lupyan, 2015; Lupyan & Ward, 2013; Ostarek & Huettig, 2017). The embodied theory of language makes a strong case that language is not divorced from sensorimotor systems but rather that they are tightly linked and interact when the meaning of a word is accessed.

An important gap in theories of embodiment is whether engagement of sensorimotor systems take place in all modalities of language. An important discovery in the language sciences was that sign languages, the manual communicative systems of deaf communities, are *on a par* with spoken languages. Sign languages are unique in that they exploit the hands and body as primary articulators and deaf signers process language through their eyes (Meier, 2002). While one could expect remarkable differences due to the different modalities (speech: oral-aural; sign: visual-manual), research has shown that both systems have considerable parallels. Speech and sign have similar linguistic

organisation (i.e., sub-lexical constitution, lexicon, syntax, see Sandler & Lillo-Martin, 2006), they are processed by overlapping brain regions (MacSweeney, Waters, Brammer, Woll, & Goswami, 2008), and they follow similar developmental trajectories (Bonvillian, Orlansky, & Novack, 1983). The evidence gathered over the last century leaves no room to question that despite their different channels of expression, sign languages are fully fledged languages in their full right.

Despite their attested similarities there is uncharted territory that may reveal important differences between speech and signs during language processing. Specifically, there is reason to question whether deaf signers show effects of visual simulation during language comprehension similar to those observed in spoken languages. Unlike spoken languages like English or Dutch, sign languages involve the visual processing of a continuous stream of body movements, the bottom-up processing of which is expected to overlap with the processes that would be activated during visual simulation. This may preclude the engagement of the visual system for language comprehension processes. That is, the processing of a sign language (through the visual system) could limit visual simulation. This contrasts not only with spoken language, but also with written language that typically consists of two-dimensional symbols that differ substantially in their visual processing requirements compared to their referents. An alternative possibility is that deaf individuals are able to juggle the parallel demands of bottom-up visual processing and top-down simulation (McCullough et al 2012; Secora & Emmorey, 2015). This feat could be facilitated by compensatory changes in the visual system which result in an increased ability to attend to multiple items in parallel (Bavelier et al., 2000; Dye, Hauser, & Bavelier, 2009).

Another factor that may reveal important differences between speech and sign is the high prevalence of iconicity in sign languages, understood as the direct relationship between form and meaning (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015). Recent developments in the languages sciences have convincingly demonstrated that iconic forms are not a marginal phenomenon but rather that they are an important component in speech and sign (Perniss, Thompson, & Vigliocco, 2010; Perniss & Vigliocco, 2014). Sign languages are unique in that *non-iconic signs* (e.g., COATRACK) co-exist with iconic forms that may represent different perceptual features of the concept they denote (see Figure 1). Some iconic signs may be categorised as *action signs* because they represent an action associated with the referent (e.g., the sign KEY represents the motor action of turning a key). *Perceptual signs* tend to represent the physical shape of the referent (e.g., the sign BUTTERFLY depicts the outline of a butterfly's wings).

Most claims around embodiment have been developed on the basis of spoken/written words with non-iconic links with the concepts they represent and to the best of our knowledge there have not yet been attempts to link iconicity with perceptual simulation. Here we explored the possibility that words that reflect more directly the motor and perceptual features of the referent (i.e., iconic) could have a stronger effect in visual simulation. Support for this prediction comes from multiple studies showing a robust processing advantage of iconic signs over non-iconic ones which has been hypothesized to arise due to more direct links of linguistic forms and the corresponding perceptual experience of the referents (e.g., Vinson, Thompson, Skinner, & Vigliocco, 2015). Thus perceptual signs, which depict the form of the referent (e.g., BUTTERFLY), could facilitate visual processing more than action and arbitrary signs. This prediction rests on

the assumption that simulations activate the same sensory processes that are active during perception. In this scenario, perceptual-iconic signs should activate category-specific visual representations more strongly because both the processing of the visual form of perceptual-iconic signs and the simulation they trigger activate congruent sensory representations. Action-perceptual signs, in contrast, represent a bodily action that is not congruent with the visual representation of the object it represents.

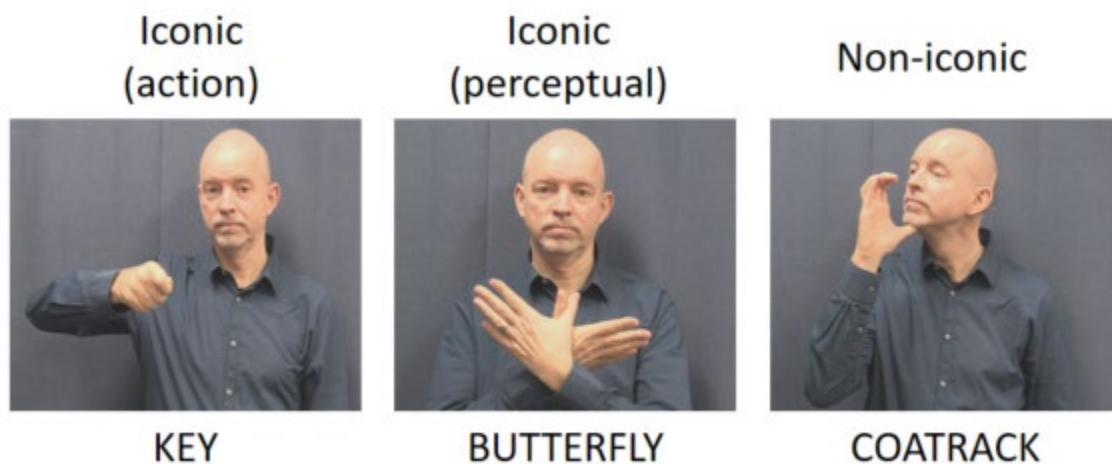


Figure 1: Examples of iconic and non-iconic signs in Sign Language of the Netherlands (NGT). The sign KEY is an *action iconic sign* because it depicts a bodily action associated with the referent (i.e., how the body interacts with a key). The sign BUTTERFLY is a *perceptual iconic sign* because the hands represent the shape of a butterfly's wings. The sign COATRACK can be classed as *non-iconic* because it lacks an evident visual relationship with the referent.

In sum, it is currently not known whether the visual system is recruited for simulation during sign language comprehension in similar ways as has been reported in spoken languages. Further, the presence of signs with different form-meaning mappings with the referent (i.e., iconic vs. non-iconic) may reveal differentiated engagement of the visual system depending on the different types of signs processed. As such, sign languages

are a unique test case to further our understanding of the processes that are common to all human languages (spoken and signed) and those that are shaped by channel of expression (i.e., modality).

In order to test our claims, we used a novel paradigm (Lupyan & Ward, 2013; Ostarek & Huettig, 2017) that measures the effect of words on basic visual detection of pictures that are suppressed from awareness using continuous flash suppression (CFS; Tsuchiya & Koch, 2005). The strength of this paradigm is that detection in CFS depends on how efficiently visual features of suppressed pictures are processed (Stein, Thoma, & Sterzer, 2015). As such, effects of words on detection capabilities are only expected if words activate visual processes involved in the earliest stages of conscious vision (Lupyan & Ward, 2013). Thus, in contrast to previously used congruency paradigms (Zwaan, Stanfield, & Yaxley, 2002), whose results were compatible with modal and amodal theories (Mahon, 2015), the present CFS paradigm can be considered a strong test of visual simulation (Ostarek & Huettig, 2019).

In the present study¹, a group of deaf users of Sign Language of the Netherlands (NGT) were presented with different types of signs (i.e., action iconic, perceptual iconic, non-iconic) after which they had to try to detect pictures suppressed using CFS that were congruent or incongruent with the previously presented sign. We predicted that signs would activate visual simulations and therefore boost the detection of congruent compared to incongruent pictures, as has been demonstrated for spoken languages (Lupyan & Ward, 2013; Ostarek & Huettig, 2017). Given that iconic signs have clear form-meaning

¹ This study received Ethical approval from the Centre for Language Studies Ethics Committee at Radboud University (reference MvB14U.015319).

mappings we expected that they could engage in visual simulation to a higher extent than non-iconic signs. In particular, the bottom-up processing of signs whose form resembles the shape of their referents (perceptual iconic signs) could pre-activate visual representations that are also recruited during visual simulation, and thus would be expected to give a boost to the effect of simulation on detection. However, it is also possible that all types of signs lead to the same degree of activation suggesting that words, iconic and non-iconic alike, have the same evocative power to activate visual representations (Lupyan & Thompson-Schill, 2012). We also tested a group of hearing non-signers on the same paradigm because they could potentially recognise the iconic motivation of some signs (Klima & Bellugi, 1979) or because signs could resemble the iconic gestures used by hearing speakers (Ortega, Ozyurek, & Peeters, 2019).

Method

Participants

Two groups of participants took part in the study. The first consisted of 27 deaf users of Sign Language of the Netherlands (NGT). Nine deaf participants were excluded from the experiment because of their high detection rate (mean > 90%) or excessive false alarm rates (> 50% or FA > mean hit rate). This left a total of 18 deaf NGT signers (of which 11 were native signers who acquired NGT from birth). The second group consisted of 19 hearing people who reported having Dutch as their mother tongue and none of them reported knowledge of any sign language (2 additional participants were excluded due to excessive false alarm rates).

Stimuli

The stimuli were selected from a set of 270 lexical signs from Sign Language of the Netherlands (NGT) which had been previously categorized according to their type of iconicity (action, perceptual, non-iconic) by two deaf research assistants (Figure 1). Statistical analysis of their agreement in the categorization was high (kappa cohen: 0.818, $p < 0.001$, 95% confidence interval [0.670, 0868]). In order to estimate the degree of iconicity of all signs, a group of 10 deaf NGT signers and 10 hearing adults with no knowledge of a sign language, rated all signs in a 7-point scale. Both groups were presented with each sign along its translation (in Dutch) after which they had to choose a value that reflected how well the sign represented the concept (1: low iconicity; 7: high iconicity). None of the participants in the ratings task took part in the actual experiment. Once the ratings from both groups were collected, 12 signs were selected for each condition (action, perceptual, non-iconic). The iconicity ratings by both groups of participants are as follows. Mean action signs: hearing = 6.2, deaf = 6.1; mean perceptual signs: hearing = 5.9, deaf = 6.1; mean non-iconic signs: hearing = 2.0, deaf = 2.0. A 2 (hearing, deaf) x 3 (action, perceptual, non-iconic) analysis of variance revealed that the ratings across the three conditions did not differ across participants $F(1,66) = 0.464, p = 0.50, \eta^2 = 0.01$. There was no interaction, but there was a significant difference in the ratings across conditions $F(2,66) = 138.45, p < 0.001, \eta^2 = 0.88$. Post-hoc analysis after Bonferroni corrections revealed that non-iconic signs differed significantly from action $p < 0.001$; CI = [3.84, 4.72] and perceptual signs $p < 0.001$; CI = [3.58, 4.46]; but action and perceptual signs did not differ from each other $p = 0.244$; CI = [-0.182, 0.696]. This shows that the stimuli in the action and perceptual conditions did not differ in their degree of iconicity but both differed from non-iconic signs. Importantly, deaf and hearing participants did not differ in their

judgements in iconicity ratings in the stimulus materials. Images of the full list of signs, their iconicity ratings, and the length of the videos can be found in the Supplementary Materials and in the Open Access repository at <https://osf.io/7s5uy/>.

Procedure

Before the CFS experiment, signers were asked to produce their favoured sign for the concepts used in the experiment. When the signs did not match the stimulus materials they were excluded from analysis, which was only necessary in 17 cases in total (see Supplementary Materials). After receiving instructions about the detection task, participants were asked to put on custom-made prism goggles (prism dioptre: 10 Δ) and place their head on a chin rest 80 cm from the screen (resolution: 1900x720, refresh rate: 60Hz). There was a separator from the nose towards the center of the screen to ensure that each eye only saw the ipsilateral half of the screen. CFS was achieved by presenting rapidly changing (at a rate of 10 Hz) masks consisting of rectangles of random sizes and colors (similar to (Hesselmann, Hebart, & Malach, 2011) to one eye, and a greyscale and slightly blurry (Gaussian blur, 3 pixel radius) picture to the other (see Figure 2).

Suppression strength was adjusted before the main experiment such that without the influence of preceding signs, pictures would be detected about half of the time, and remain completely invisible the other half. To that end, participants first performed a staircase procedure aimed at determining individual detection thresholds (96 trials), where each hit resulted in a slight reduction of the image contrast and each miss results in a slight increase. This avoids ceiling and floor effects and in most cases leads to hit rates of

approximately 50% in the main experiment. Stimulus presentation and response-logging was done using Presentation Software (Version 16.2, www.neurobs.com).

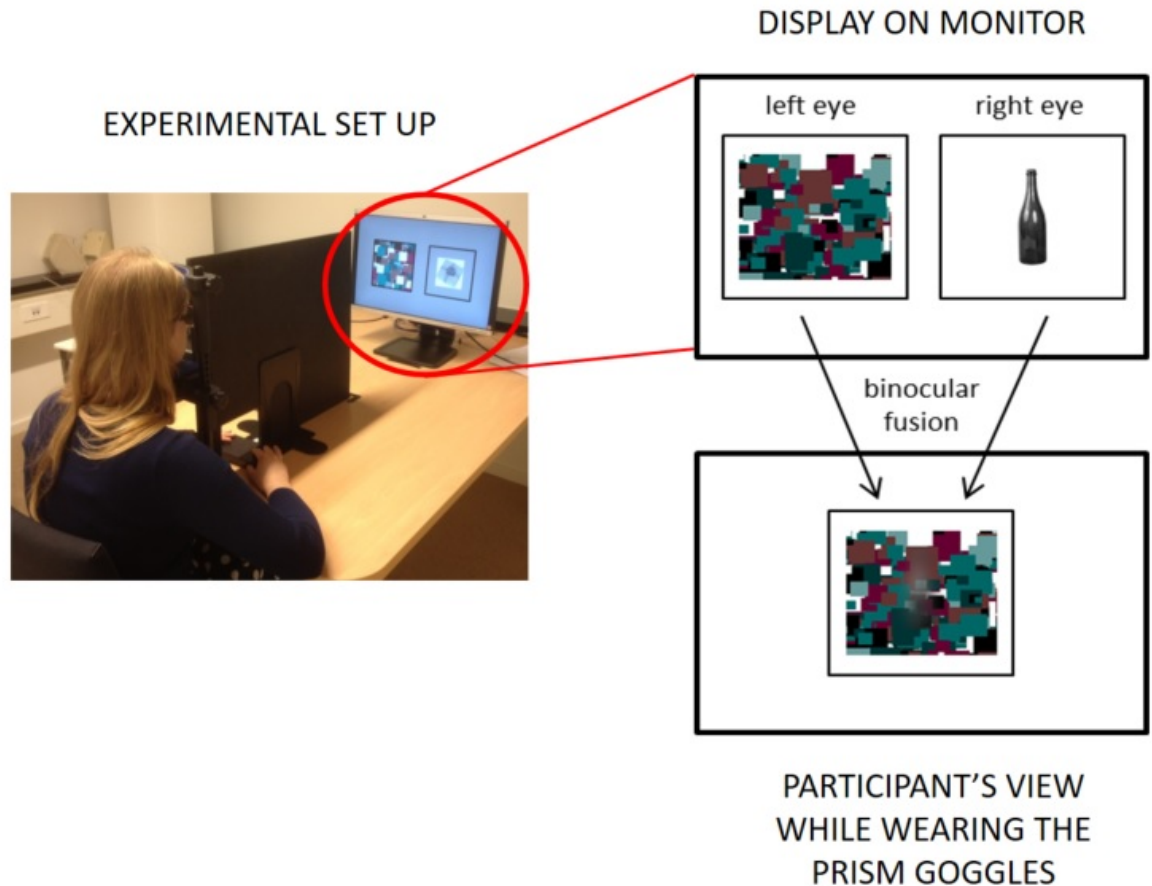


Figure 2. Experimental set up of the Continuous Flash Suppression (CFS) paradigm. Participants sat in front of a computer with their head on a chin rest while wearing custom-made prism goggles. A separator from the nose towards the center of the screen was located on the table to ensure that each eye only saw the ipsilateral half of the screen. The computer screen displayed the visual noise on the left side and on the right side the picture was displayed which was congruent or incongruent with the sign shown before. The intensity of the pictures was adjusted such that detection rates approximated 50% (i.e., pictures remained invisible about half of the time).

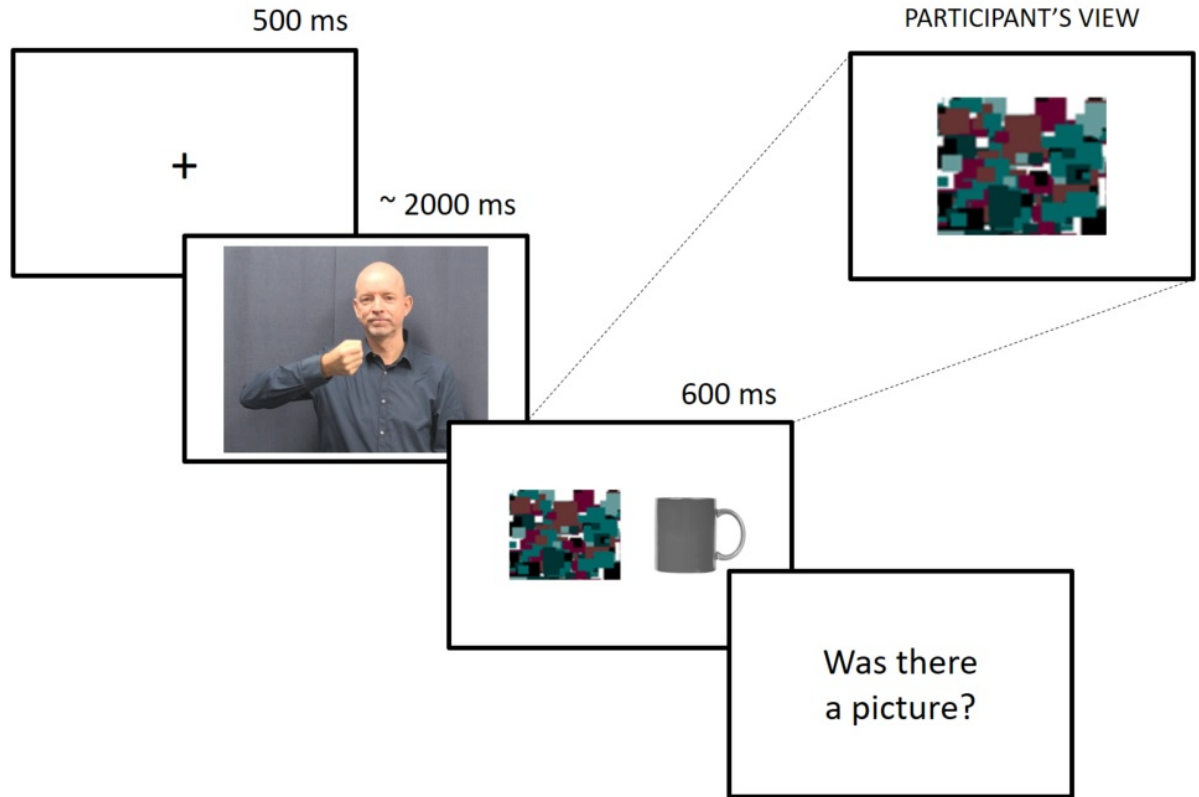


Figure 3. Structure of each experimental trial. A fixation cross was presented for 500 ms, and this was followed by an NGT sign which played automatically in its entirety. The sign was immediately followed by the CFS masks along with a congruent or incongruent picture for 600 ms (note that participants view only the visual nose due to the prims goggles). Participants were then required to press a button to indicate whether they had seen or not a picture.

In the main experiment, participants were presented with short videos of 36 signs (12, action-iconic, 12 shape-iconic, 12 non-iconic) and 36 corresponding pictures. The pictures were taken from the (De Groot, Koelewijn, Huettig, & Olivers, 2016) database. They were repeated eight times across the experiment (total of 288 picture-present trials). In an additional 144 trials no picture was present, resulting in a total of 432 trials. Critically, in half of the picture-present trials signs were congruent vs. incongruent with the pictures. Incongruent images were fixed per participant (e.g., for a given participant, in incongruent

trials where the sign for BALL was presented it would always be followed by a picture of a battery). Trials were presented in random order. Each trial began with a central fixation cross (500 ms). Then, a short video of a sign was displayed. Based on the previous finding from spoken language that effects of words on detection in CFS are specific to the first hundreds of ms after word onset (Ostarek & Huettig, 2017), we displayed the signs only for as long as absolutely necessary to be recognized (preparations and retraction of the sign were not part of the sign videos). There were two versions per sign, which were used in two blocks (the order was counterbalanced across participants). In one version the video was stopped as soon as the movement of retraction of the sign began. The second version lasted an additional 200 ms. As this did not affect any of the results, the data from both versions were collapsed. The signs were immediately followed by the CFS masks (and a picture when applicable) that were displayed for 600 ms. Finally a screen appeared asking the participant to indicate by button press whether there was a picture or not (see Figure 3).

Analysis

We determined exclusion criteria before data collection according to which participants with detection rates higher than 90% or lower than 10% on average, as well as participants with higher false alarms than detection rates or false alarm rates higher than 50% were not used for analysis. This led to the exclusion of 9 signers (2 due to ceiling effects, 7 due to excessive false alarm rates) and 2 non-signers (both due to excessive false alarm rates), suggesting that there are differences in how strongly deaf signers vs. controls are affected by CFS (see Supplemental Materials for further information). One item (clock) was excluded in the control group due to a ceiling effect (> 90% detection).

To analyse detection rates, we used a logistic mixed effects model (as implemented in the R package lme4, (Bates, Mächler, Bolker, & Walker, 2014). The first analysis tested whether the congruency between signs and targets influenced detection rates and whether signers and non-signers differed. To this end, a logistic mixed effects model was run with Congruency and Group as well as their interaction as fixed effects, by-participant and by-item random intercepts, and by-participant and by-item random slopes for the effect of Congruency. To test for effects of iconicity, this factor was added as a fixed effect, as well as random intercepts and by-participant random slopes. Follow-up within-group analyses were run with Congruency and Iconicity as well as their interaction as fixed effects and random by-participant intercepts and slopes for the effect of Congruency and Iconicity and their interaction, as well as by-target item intercepts and slopes for the effect of Congruency. For models including factors with more than two levels, we calculated main and interaction effects using Type II Wald Chi-square tests (see Baayen, Davidson, & Bates, 2008). Planned follow-up comparisons of the effect of Congruency at the three different types of iconicity were done using pairwise tests of estimated marginal means.

Note that the different iconicity conditions refer to the signs associated with the targets, not the signs that participants saw on a given trial. At first sight, this seems counter-intuitive, but it is preferable because this way the targets are exactly the same in the congruent and incongruent conditions. This allowed us to evaluate whether detection of a given picture was influenced by whether a congruent sign was presented and whether the iconicity status of the corresponding sign mattered. The alternative, where performance is compared for a sign with a given iconicity status followed by a) the congruent picture, or b) an incongruent (and hence different) picture, is less desirable because in that case for

each sign different conditions display different pictures for which baseline detection rates likely differ.

Second, we conducted equivalent analyses on d' -prime scores. We calculated d' -prime scores per participant per condition which were then submitted to a repeated-measures 2 (Congruency; congruent vs. incongruent) by 2 (Group; signer vs. non-signers) to analyse the data across groups, and to 2 (Congruency; congruent vs. incongruent) by 3 (Prime Type; action-iconic vs. shape iconic vs. non-iconic) ANOVAs for within-group analyses. Paired-samples t -tests were used to test the effect of Congruency at the three levels of Prime Type.

Results

The main result (Figure 4) is that for deaf signers detection rates and sensitivity were higher in the congruent condition (hit rate: $M = 0.59$, $SE = 0.011$; d' : $M = 1.091$, $SE = 0.17$) compared to the incongruent condition (hit rate: $M = 0.52$, $SE = 0.011$), whereas for non-signers there was virtually no difference between the congruent (hit rate: $M = 0.477$, $SE = 0.011$; d' : $M = 1.197$, $SE = 0.165$) and incongruent condition (hit rate: $M = 0.472$, $SE = 0.011$; d' : $M = 1.174$, $SE = 0.161$). The logistic mixed effects and d' analyses revealed a main effect of Congruency with higher scores in the congruent condition (hit rate: estimate = 0.1, $SE = 0.032$, $z = 3.077$, $p = 0.002$; d' : $F(1, 35) = 8.44$, $p = 0.006$), and crucially this detection boost for congruent pictures was stronger for deaf signers than for hearing non-signers (interaction Congruency*Group for hit rates: estimate = 0.085, $SE = 0.029$, $z = -2.879$, $p = 0.006$; d' : $F(1, 35) = 5.291$, $p = 0.028$). Follow-up per-group analyses confirmed that detection was enhanced for congruent targets in deaf signers (hit rates:

estimate = 0.180, $SE = 0.060$, $z = 3.021$, $p = 0.003$; d' : $F(1,17) = 8.353$, $p = 0.010$, $\eta^2 = 0.329$) but not in hearing controls ($z < 1$).

We next explored whether the Congruency effect in deaf signers was modulated by age of first exposure to a sign language. Only around 5-10% of deaf children learn a sign language natively from their signing caregivers (Mitchell & Karchmer, 2004) with delay to sign exposure having an effect in sign processing (Mayberry, 2007). We ran a model with Congruency and nativeness as well as their interaction as fixed effects, random by-participant intercepts and slopes for the effect of Congruency, and random by-target intercepts and slopes for the effects of congruency and nativeness. There was no evidence that nativeness modulated the Congruency effect ($p > .3$). However, we would like to note that the statistical power for this analysis was insufficient to draw firm conclusions on the effect of nativeness.

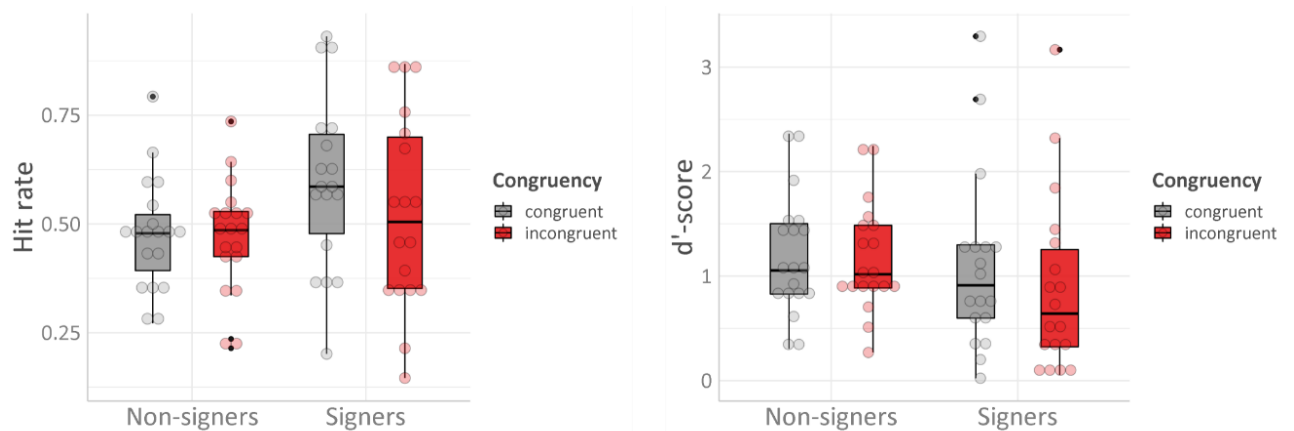


Figure 4. Boxplots of hit rates (left) and d' -scores (right) of signers and non-signers in the congruent and incongruent trials. The solid horizontal lines indicate the medians, the upper and lower ends of the boxes indicate the 75th and 25th percentiles, dots indicate individual participants' means.

We now turn to the role of iconicity. There was some evidence for a main effect of iconicity (hit rates: $\chi^2 = 2.6, p = 0.27, d': F(2, 70) = 13.612, p < 0.001$) and an interaction between iconicity and Group (hit rates: $\chi^2 = 5.42, p = 0.066, d': F(2, 70) = 3.669, p = 0.031$), likely due to slightly better detection of targets associated with non-iconic signs, especially for non-signers (see Fig. 5). Note that these tentative effects reflect baseline detection differences regardless of the congruency between signs and target images. Importantly, there was no evidence that iconicity modulated the congruency effect for either group; the interaction between congruency and iconicity ($p > 0.3$) and the interaction between congruency, group, and iconicity ($p > 0.3$) were not significant. Nevertheless, as this directly pertains to the prediction that the congruency effect would be boosted in the perceptual-iconic condition, we report the corresponding post-hoc analyses. For signers, the interaction of Congruency and Iconicity was not significant (hit rates: $\chi^2(2) = 3.770, p = 0.152; d': F(2,34)=0.977, p = 0.387$). Planned pairwise contrasts indicated Congruency effects in the action-based iconic (hit rates: estimate (log-odds ratio) = 0.485, $SE = 0.149, z = 3.248, p = 0.001; d': t(17)=2.718, p = 0.015$) and the non-iconic condition (hit rates: estimate (log-odds ratio) = 0.425, $SE = 0.155, z = 2.734, p = 0.006; d': t(17)=2.088, p = 0.052$), but not in the perceptual-based iconic condition (hit rates: estimate (log-odds ratio) = 0.178, $SE = 0.149, z = 1.197, p = 0.231; d': t(17) = 1.376, p = 0.187$). Non-signers did not reveal an effect of Congruency for any type of sign (Fig. 5).

Finally, false alarm rates were higher ($\chi^2(1)=5.231, p = 0.022$) for signers (Action-iconic $M = 0.235, SE = 0.017$; Perceptual-iconic $M = 0.223, SE = 0.017$; Non-iconic $M = 0.214, SE = 0.017$) than for non-signers (Action-iconic $M = 0.145, SE = 0.013$; Perceptual-

iconic $M = 0.135$, $SE = 0.013$; Non-iconic $M = 0.125$, $SE = 0.013$). There was no evidence for an effect of Iconicity ($p > 0.3$) or an interaction effect ($p > 0.9$) on false alarm rates.

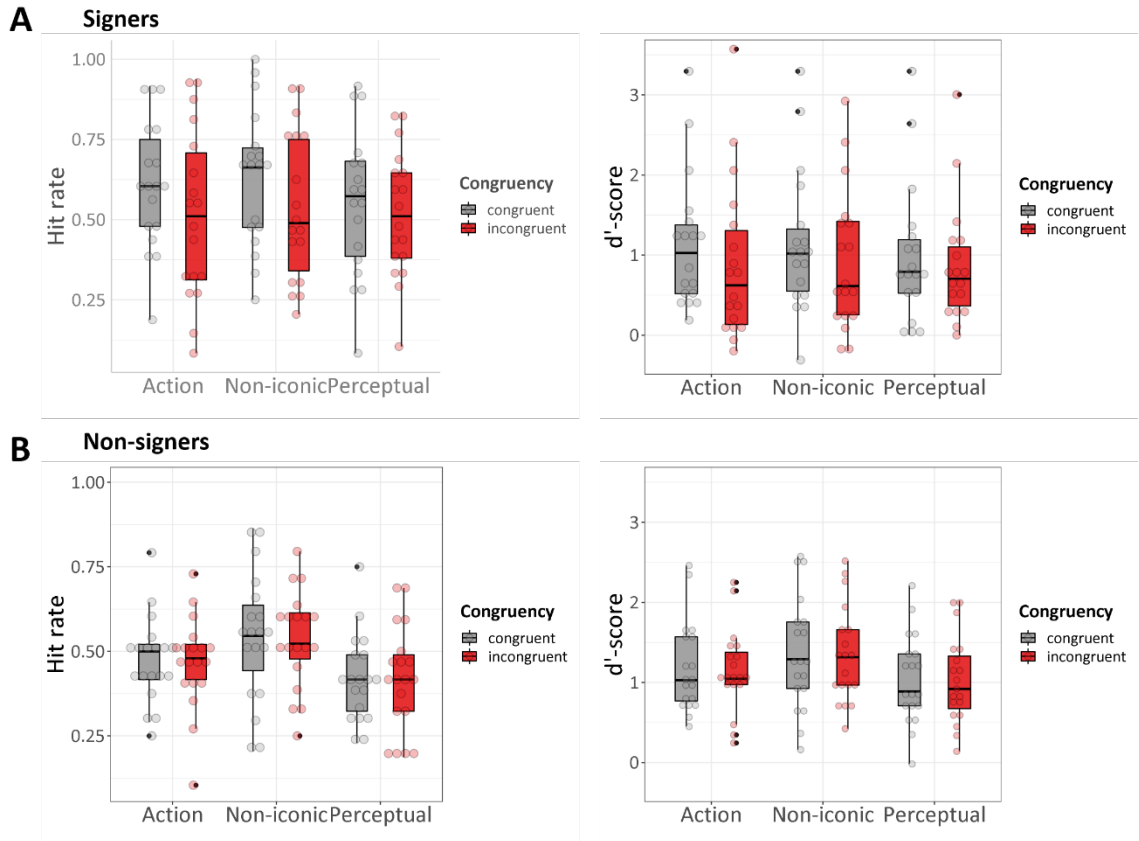


Figure 5. (A) Hit rates (left) and d-prime scores (right) of signers for congruent vs. incongruent pictures whose corresponding signs are action-based iconic, non-iconic, or perceptual-based iconic. (B) Hit rates (left) and d-prime scores (right) for non-signers. The solid horizontal lines indicate the medians, the upper and lower ends of the boxes indicate the 75th and 25th percentiles, the dots indicate individual participants' means.

Discussion

Research has shown that semantic processing engages our perceptual systems (Meteyard et al., 2012) and that words activate perceptual features of objects via perceptual simulation (Lewis & Poeppel, 2014; Ostarek & Huettig, 2017). These claims, however, have been made around spoken languages which are primarily expressed through speech/text and whose words have an arbitrary relationship between the form of a word and the concept they represent (i.e., non-iconic). As such, there is limited evidence whether the processing of sign languages of deaf communities also engage in this form of perceptual simulation and whether iconicity contributes to simulation. While there is no doubt that speech and sign have the same underlying linguistic structure, it is unclear whether both modalities (i.e., speech and sign) operate under similar processes. Here we entertained the possibility that the processing of a sign language may not lead to visual simulation because of excessive demands to the visual system. That is, we could find differences in the effect of spoken vs signed language processing in visual simulation due to the visuo-spatial nature of signs.

To address this question, we probed visual simulation in the processing of signs in a group of deaf users of Sign Language of the Netherlands (NGT). We capitalised on a novel application of Continuous Flash Suppression (CFS) to test whether signs activate visual simulations of the objects they refer to and therefore boost the detection of congruent objects. Our results show that the evocative properties of words can be extended to lexical labels expressed in the manual modality (i.e., signs). Despite important physical differences between words and signs that could result in differences in language processing, we observe that signs have similar knock-on effects on visual processing. Specifically, we found that when viewing signs referring to objects, deaf signers activate visual

representations of these objects which in turn influence detection sensitivity to congruent pictures suppressed with CFS. Signs also activate sensory representations in deaf people at similar rates as spoken words do in hearing people (Lupyan & Ward, 2013; Ostarek & Huettig, 2017). Importantly, the processing of a manual-visual language does not disrupt (visual) simulation processes. As has been shown for motor simulation through motion perception (in MT+ areas) (McCullough et al 2012), and semantic compatibility (Secora & Emmorey, 2015), our results add to the growing body of evidence that deaf signers can juggle the parallel demands of bottom-up visual processing and top-down simulation. Our data provide empirical evidence that signs and words alike hold a unique place in human cognition in that both have the capacity to modulate processes across perceptual systems with striking parallels. A core property of the human capacity for language, regardless of whether it is expressed through speech or sign, seems to be the penetration of perceptual systems.

A result worth highlighting is that sign iconicity did *not* increase visual sensitivity in deaf signers. There is a growing body of evidence showing a processing advantage of iconic signs over non-iconic signs in deaf signers because of the close links between the form of a sign and the concept it represents (Ormel, Hermans, Knoors, & Verhoeven, 2011; Thompson, Vinson, & Vigliocco, 2009; Vinson et al., 2015). A possible mechanistic account of how the iconicity advantage arises is that the physical form of a sign pre-activates processes that are subsequently recruited top-down during simulation. Based on this we predicted that perceptual-iconic signs (those representing the shape of the object) could show stronger effects on visual sensitivity because the form of these signs could

facilitate visual simulation processes. This was not the case suggesting that the locus of the iconicity advantage lies at higher levels of processing.

Embodied theories are typically not very specific about the exact representations or processes predicted to be activated during simulation but generally claim that at some level simulations activate the same processes that are active during perception. If so, one would expect bottom-up processing of the visual form of perceptual-iconic signs and the ensuing visual simulation to have additive effects (as they would both be expected to pre-activate shape-specific representations in the visual system). We did not obtain evidence for such additive effects. The present results indicate a dissociation between bottom-up sensory and top-down (simulation) signals. We propose that this dissociation arises because simulations specifically reflect diagnostic features of an object's category (e.g., the general features of a bottle), whereas bottom-up processing of an object involves a unique exemplar with idiosyncratic low-level features (e.g., the specific features of a given bottle). A similar mechanism has been suggested for featural (David et al., 2008) and categorical attention (Cukur, Nishimoto, Huth, & Gallant, 2013), whereby attention to particular features or semantic categories during natural vision leads to increased tuning to relevant features, with stronger effects in high-level compared to low-level retinotopic visual cortex (Cukur et al., 2013). Further elucidating the neuro-cognitive mechanisms that underlie the dissociation between simulation and perception is an important challenge for future research.

Iconicity had no effect in hearing non-signers either. Sign-naïve participants can guess the meaning of some iconic signs thanks to their resemblance with their referent (Klima & Bellugi, 1979; Pizzuto & Volterra, 2000). They also have exposure to manual

iconicity thanks to the iconic gestures that are commonly produced accompanying speech in face-to-face interactions (Kita, 2000; Özyürek, Willems, Kita, & Hagoort, 2007). However, whereas hearing non-signers can identify the meaning of some iconic signs when they resemble gestures (Ortega et al., 2019), they typically can only identify accurately a small proportion of iconic signs (Klima & Bellugi, 1979; Pizzuto & Volterra, 2000; Sehyr and Emmorey, 2020). Our data suggest that iconicity in signs did not lead to the efficient retrieval of information from the visual system. Thus, whereas hearing non-signers are perceptive to some extent of sign iconicity (Klima & Bellugi, 1979; Ortega et al., 2019), it is not sufficient to activate specific visual representations. This suggests that the ability to retrieve diagnostic conceptual features from the sensorimotor systems in signs requires a fully-fledged linguistic system in the manual modality.

The human capacity of language has the potential to be expressed not only in speech but also with the body, as in the case of the sign languages of deaf communities. The contribution of our study is that both words and signs display unique referential properties through links with our sensorimotor experiences. While objects can be referred to through a wide array of cues (e.g., pictures, sounds), words and signs stand out for their capacity to efficiently activate diagnostic conceptual features that capture the essence of a concept. Our data suggest that the retrieval of sensorimotor features involves mechanisms that modulate surprisingly basic visual processes, namely those involved in the earliest stages of conscious vision. Linguistic labels have the power to transgress other cognitive domains and stimulate representations across systems. This new evidence can help us move towards a more complete picture of the cognitive architecture of language that is not confined to the communicative channels of spoken language.

Context Paragraph

There is increasing evidence that comprehending words that refer to physical objects involves the activation of sensory processes that would be activated if the object was actually perceived. So far, this process called *perceptual simulation* has only been studied in the domain of spoken languages and it is currently not known whether it is used in sign languages, the manual languages used by deaf communities. Here, we adapted a recently developed strong test of perceptual simulation that measures the effect of language on participants' ability to detect pictures that are suppressed from awareness. Our results suggest that signs, similar to spoken words, trigger visual simulations that can make otherwise invisible objects visible. This suggests that perceptual simulation is a general property of language regardless of mode of communication (i.e., speech vs. sign).

References

- Barsalou, L. W. (2008). Grounding Symbolic Operations in the Brain's Modal Systems. In G. Semin & E. Smith (Eds.), *Embodied Grounding. Social, Cognitive, Affective and Neuroscientific Approaches* (pp. 9–42). Cambridge: Cambridge University Press.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390-412.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *ArXiv Preprint ArXiv:1406.5823*. Retrieved from <https://arxiv.org/abs/1406.5823>
- Bavelier, D., Tomann, A., Hutton, C., Mitchell, T., Corina, D., Liu, G., & Neville, H. (2000). Visual attention to the periphery is enhanced in congenitally deaf individuals. *The Journal of Neuroscience*, *20*(17), RC93. <https://doi.org/psycinfo/2001-03295-001>
- Bonvillian, J., Orlansky, M., & Novack, L. (1983). Developmental milestones: Sign language acquisition and motor development. *Child Development*, *54*(6), 1435–1445.
- Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: a label advantage in object recognition. *Journal of Neuroscience*, *35*(25), 9329–9335.
- Correia, J., Formisano, E., Valente, G., Hausfeld, L., Jansma, B., & Bonte, M. (2014). Brain-based translation: fMRI decoding of spoken words in bilinguals reveals

- language-independent semantic representations in anterior temporal lobe. *Journal of Neuroscience*, *34*(1), 332–338.
- De Groot, F., Koelewijn, T., Huettig, F., & Olivers, C. N. L. (2016). A stimulus set of words and pictures matched for visual and semantic similarity. *Journal of Cognitive Psychology*, *28*(1), 1–15.
<https://doi.org/10.1080/20445911.2015.1101119>
- 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>
- Dye, M. W. G., Hauser, P. C., & Bavelier, D. (2009). Is visual selective attention in deaf individuals enhanced or deficient? The case of the useful field of view. *PloS One*, *4*(5), e5640. <https://doi.org/10.1371/journal.pone.0005640>
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307.
- Hesselmann, G., Hebart, M., & Malach, R. (2011). Differential BOLD Activity Associated with Subjective and Objective Reports during “Blindsight” in Normal Observers. *Journal of Neuroscience*, *31*(36), 12936–12944.
<https://doi.org/10.1523/jneurosci.1556-11.2011>
- Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and Gesture* (pp. 162–185). Cambridge: Cambridge University Press.
- Klima, E., & Bellugi, U. (1979). *The Signs of Language*. Harvard: Harvard University Press.

- Lewis, G., & Poeppel, D. (2014). The role of visual representations during the lexical access of spoken words. *Brain and Language*, *134*, 1–10.
- Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General*, *141*(1), 170–186. <https://doi.org/10.1037/a0024904>
- Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(35), 14196–14201.
<https://doi.org/10.1073/pnas.1303312110>
- MacSweeney, M., Waters, D., Brammer, M. J., Woll, B., & Goswami, U. (2008). Phonological processing in deaf signers and the impact of age of first language acquisition. *NeuroImage*, *40*(3), 1369–1379.
<https://doi.org/10.1016/j.neuroimage.2007.12.047>
- Mahon, B. Z. (2015). What is embodied about cognition? *Language, Cognition and Neuroscience*, *30*(4), 420–429.
- Mayberry, R. I. (2007). When timing is everything: Age of first-language acquisition effects on second-language learning. *Applied Psycholinguistics*, *28*(03), 537–549.
<https://doi.org/10.1017/S0142716407070294>
- McCullough, S., Saygin, A. P., Korpics, F., & Emmorey, K. (2012). Motion-sensitive cortex and motion semantics in American Sign Language. *Neuroimage*, *63*(1), 111-118.
- Meier, R. P. (2002). Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In R. P. Meier, K. Cormier,

- & D. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 1–26). Cambridge: Cambridge University Press.
- Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*(7), 788–804. <https://doi.org/10.1016/j.cortex.2010.11.002>
- 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>
- Ormel, E., Hermans, D., Knoors, H., & Verhoeven, L. (2009). The role of sign phonology and iconicity during sign processing: the case of deaf children. *Journal of Deaf Studies and Deaf Education*, *14*(4), 436–448.
- Ormel, E., Hermans, D., Knoors, H., & Verhoeven, L. (2011). Cross-language effects in written word recognition: The case of bilingual deaf children. *Bilingualism: Language and Cognition*, *15*(02), 288–303.
- Ortega, G., & Ostarek, M. (2020, September 24). Evidence for visual simulation during sign language processing. Retrieved from osf.io/7s5uy
- Ortega, G., Ozyurek, A., & Peeters, D. (2019). 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*. Advance Online Publication.
- Ostarek, M., & Huettig, F. (2017). Spoken words can make the invisible visible—Testing the involvement of low-level visual representations in spoken word processing.

- Journal of Experimental Psychology: Human Perception and Performance*, 43(3), 499–508. <https://doi.org/10.1037/xhp0000313>
- Özyürek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: insights from event-related brain potentials. *Journal of Cognitive Neuroscience*, 19(4), 605–616. <https://doi.org/10.1162/jocn.2007.19.4.605>
- 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), 1664–1678. <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 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(20130300). <https://doi.org/10.1098/rstb.2013.0300>
- Pizzuto, E., & Volterra, V. (2000). Iconicity and transparency in Sign Languages: A cross-linguistic cross-cultural view. In K. Emmorey & H. L. Lane (Eds.), *The signs of language revisited: An anthology to Honor Ursula Bellugi and Edward Klima* (pp. 229–250). Mahwah, N. J.: Lawrence Erlbaum Associates.
- Ralph, M., Jeffries, E., Patterson, K., & Rogers, T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42–55. <https://doi.org/10.1038/nrn.2016.150>
- Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge University Press.

- Secora, K., & Emmorey, K. (2015). The action-sentence compatibility effect in ASL: The role of semantics vs. perception. *Language and cognition, 7*(2), 305-318.
- Sehyr, Z. S., & Emmorey, K. (2019). The perceived mapping between form and meaning in American Sign Language depends on linguistic knowledge and task: evidence from iconicity and transparency judgments. *Language and Cognition, 11*(2), 208-234.
- Shtyrov, Y., Butorina, A., Nikolaeva, A., & Stroganova, T. (2014). Automatic ultrarapid activation and inhibition of cortical motor systems in spoken word comprehension. *Proceedings of the National Academy of Sciences, 111*(18), E1918–E1923.
- Stein, T., Thoma, V., & Sterzer, P. (2015). Priming of object detection under continuous flash suppression depends on attention but not on part-whole configuration. *Journal of Vision, 15*(3), 15–15.
- Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2009). The link between form and meaning in American Sign Language: lexical processing effects. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 35*(2), 550–557.
<https://doi.org/dx.doi.org/10.1037/a0014547>
- Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8*(8), 1096.
- Vinson, D., Thompson, R. L., Skinner, R., & Vigliocco, G. (2015). A faster path between meaning and form? Iconicity facilitates sign recognition and production in British Sign Language. *Journal of Memory and Language, 82*, 56–85.
<https://doi.org/10.1016/j.jml.2015.03.002>

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168–171.