

## Cross-situational learning of sign-like gestures in children and adults: a behavioural and event-related potential study

Arianna Colombani, Varghese Peter, Qian Yin Mai, Amanda Saksida, Natalie Boll-Avetisyan, Outi Tuomainen & Mridula Sharma

To cite this article: Arianna Colombani, Varghese Peter, Qian Yin Mai, Amanda Saksida, Natalie Boll-Avetisyan, Outi Tuomainen & Mridula Sharma (04 Aug 2025): Cross-situational learning of sign-like gestures in children and adults: a behavioural and event-related potential study, Language, Cognition and Neuroscience, DOI: [10.1080/23273798.2025.2539136](https://doi.org/10.1080/23273798.2025.2539136)

To link to this article: <https://doi.org/10.1080/23273798.2025.2539136>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 04 Aug 2025.



[Submit your article to this journal](#)



Article views: 418



[View related articles](#)



[View Crossmark data](#)

## Cross-situational learning of sign-like gestures in children and adults: a behavioural and event-related potential study

Arianna Colombani <sup>a,b,c,d,e,f</sup>, Varghese Peter <sup>g</sup>, Qian Yin Mai<sup>h</sup>, Amanda Saksida <sup>ij</sup>,  
Natalie Boll-Avetisyan <sup>f</sup>, Outi Tuomainen <sup>f</sup> and Mridula Sharma <sup>h,k</sup>

<sup>a</sup>International Doctorate for Experimental Approaches to Language and Brain (IDEALAB), University of Potsdam, Potsdam, Germany; <sup>b</sup>Macquarie University, Sydney, Australia; <sup>c</sup>University of Groningen, Groningen, Netherlands; <sup>d</sup>Newcastle University, Newcastle upon Tyne, UK; <sup>e</sup>School of Psychological Sciences, Macquarie University, Sydney, Australia; <sup>f</sup>Department of Linguistics, University of Potsdam, Potsdam, Germany; <sup>g</sup>Discipline of Psychology, School of Health, University of the Sunshine Coast, Queensland, Australia; <sup>h</sup>Department of Linguistics, Macquarie University, Sydney, Australia; <sup>i</sup>Institute for Maternal and Child Health-IRCCS “Burlo Garofolo”-Trieste, Italy; <sup>j</sup>Educational Research Institute, Ljubljana, Slovenia; <sup>k</sup>Caring Futures Institute, College of Nursing and Health Sciences, Flinders University, Adelaide, Australia

### ABSTRACT

Drawing on the innate human ability to detect regularities in the language input (statistical learning), this study applies a cross-situational learning paradigm to test the learning of unfamiliar sign-like gestures (in the form of pseudosigns) for familiar spoken words in children and adults. Twenty-five children (8–11 years) and 19 adults (18–35 years) were familiarised with 8 word-pseudosign pairs and tested on a recognition and a semantic categorisation task, with simultaneous EEG recording. Both groups demonstrated above-chance accuracy, indicating successful learning of word-pseudosign pairs and their meanings, with an advantage of adults over children. Both groups also showed an N400 followed by an LPC response during the recognition task. During categorisation, adults demonstrated an N400 response, whereas, in children, an N400 emerged only when the correctly identified trials were considered. These results suggest that pseudosigns are highly salient linguistic inputs, likely to be learned through statistical computations.

### ARTICLE HISTORY

Received 2 July 2024  
Accepted 9 July 2025

### KEYWORDS

Pseudosigns; statistical learning; cross-situational learning; N400; LPC; semantic categorisation; associative learning

## Introduction

Human communication is, in its essence, multimodal, with both auditory and visual input playing essential roles (Holler & Levinson, 2019). During early childhood, infants initially communicate with adults through a combination of vocal and non-vocal acts. These acts include cries, vocalisations, protophones, eye contact (Long et al., 2020; Oller et al., 2021), and gestures, such as hand, head, eye, and body movements (Goldin-Meadow, 1999; Hübscher & Prieto, 2019). Over time, there is a gradual shift towards verbal communication, typically occurring by the end of the second year (Capirci et al., 1996; Saksida et al., 2023). In this initial phase, gestures serve to represent objects, actions, or events (Acredolo & Goodwyn, 1988; Capirci et al., 1996; Goldin-Meadow & Butcher, 2003) and preverbal children predominantly use them to direct caregivers' attention, make requests, and reference events (Acredolo & Goodwyn, 1988; Bates, 1979; Capirci et al., 1996;

Goldin-Meadow, 1999; Özçalışkan & Dimitrova, 2013). As toddlers progress towards childhood, speech offers a more energy-efficient means of communication; in child and adult communication, gestures persist as integral language components to accompany, emphasise, and complement speech (Colletta et al., 2015; McNeill, 1985, 1992, 2005; Willems et al., 2007).

In both child and adult communication, gestures can be classified according to their dependence on context. Context-dependent gestures such as *pointing*, *reaching* and *giving* serve to establish joint attention and derive meaning from the immediate environment in which they are used (Bates et al., 1975; Capirci et al., 1996). Context-independent gestures, on the other hand, always denote a specific referent, and their meaning is stable and agreed upon across users. Examples of these gestures are putting a thumb up to signify “ok”, flapping the hand to indicate “bird”, or bringing a hand to the mouth to mean “to eat”. In the literature,

**CONTACT** Arianna Colombani  arianna.colombani@uni-potsdam.de  Universität Potsdam Linguistics Department Karl-Liebknecht-Str. 24/25 House 14, 14476 Potsdam

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23273798.2025.2539136>.

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

context-independent gestures are referred to with different definitions, with the most commonly used being *symbolic* (Acredolo & Goodwyn, 1988; Bates et al., 1979; Fasolo & D'Odorico, 2012; Goodwyn et al., 2000; Vallotton, 2012) and *representational* (Batista et al., 2019; Rowe et al., 2008), among others (for a comprehensive list of definitions see Murillo & Casla, 2021 and Suttora et al., 2022). Another type of context-independent gestures, often referred to as *sign language-like gestures* (or *sign-like*; see e.g. Krentz & Corina, 2008; Newman et al., 2015), has been explored in studies on early language acquisition, particularly to investigate how infants process and learn visual versus auditory stimuli and their preferences for gestural or spoken input (e.g. Baker et al., 2006; Krentz & Corina, 2008; Newman et al., 2015; Rabagliati et al., 2012; Stone et al., 2018). These gestures are typically modelled after sign languages but do not belong to an established linguistic system.

Context-independent gestures are particularly relevant during communication as they convey semantic information, allowing to represent and refer to objects and concepts, even in the absence of specific referents (Roberts & Hampton, 2018; Goldin-Meadow & Butcher, 2003; Capirci et al., 1996; Acredolo & Goodwyn, 1996). These gestures can stand alone, functioning similarly to words in conveying meaning (Capone & McGregor, 2004; Goodwyn et al., 2000; Pizer et al., 2007). For these reasons, they can be used to promote communication in the pre-verbal stage of language development (e.g. Acredolo & Goodwyn, 1996; Paul et al., 2019; Vallotton, 2012) and with clinical populations (e.g. Harrigan & Nikolopoulos, 2002). In these cases, context-independent gestures are usually produced along with speech. To guarantee linguistic salience and coherence, they are often drawn from the vocabulary of sign languages (for a review, see Colombani et al., 2022).

These shared characteristics with spoken words suggest that context-independent gestures could likewise be referential and interpreted as labels for referents. However, existing literature on the topic has primarily focused on infants and toddlers (e.g. Capone & McGregor, 2005; McGregor et al., 2009; Namy et al., 2004; Namy & Waxman, 1998; Puccini & Liszkowski, 2012; Suanda et al., 2013), showing that the ability to learn symbols initially applies to both gestures and words but gradually becomes more specific to words, likely due to predominant exposure to verbal communication (Krentz & Corina, 2008). Later in toddlerhood, gesture learning appears to re-emerge, potentially driven by broader symbolic experiences and increased sensitivity to communicative intent, reflecting a “U-shaped”

developmental trajectory (Namy et al., 2004). To date, no direct studies have investigated the use of gestures as object labels in older children or adults. Research in this area has predominantly focused on sign language acquisition (e.g. Lieberman et al., 2022; Lillo-Martin & Henner, 2021), leaving a gap in understanding how gestural labels are acquired and associated with meaning beyond early childhood.

### **Language acquisition and cross-situational learning**

During language development, one of the possible mechanisms in which learning occurs is through the gradual, incidental recognition of patterns of regularities within the language input. This process has been referred to as *statistical learning* (Conway, 2020; for a recent discussion of the term, see Christiansen, 2019). So far, statistical learning abilities in humans have been extensively studied across different domains and populations (for a review, see Isbilen & Christiansen, 2022). However, in the field of language learning in the visual modality, specifically in hearing non-signer (or deaf signers) populations, to date, only Hofweber and colleagues (2022) have applied a statistical learning paradigm to sign learning. This study examined the capacity of sign language-naïve individuals to acquire signs presented within a naturalistic context (i.e. sentences with no auditory input). Results showed that sign acquisition upon initial exposure is feasible, providing preliminary evidence that adult-like language-learning mechanisms are operative across modalities.

In statistical learning studies, a commonly used paradigm is the *cross-situational learning* (CSL) paradigm (Yu & Smith, 2007), which accounts for 11% of all statistical learning studies (Frost et al., 2019). In this paradigm, participants are presented with a series of ambiguous learning trials consisting of multiple referents and words, with no explicit indication of word-referent correspondences. Learning occurs by statistically computing the probability of word-referent co-occurrences across the trials. This identification of patterns across multiple contexts can, therefore, classify cross-situational learning as a form of statistical learning. This form of learning can be considered implicit, as acquired without explicit intent or feedback. It relies on basic mechanisms that function outside conscious awareness (Christiansen, 2019), such as associative learning processes (Perruchet & Pacton, 2006) operating under attention and memory constraints (Kachergis et al., 2013; Vlach & Johnson, 2013; Yurovsky et al., 2014). With this ambiguous learning context, this paradigm mimics real-world learning scenarios in the laboratory setting, providing the

opportunity to test learning mechanisms in a naturalistic-like learning environment.

Cross-situational learning paradigms have been applied to test word learning across ages, stimuli and conditions (for a review, see Isbilen & Christiansen, 2022; Roembke et al., 2023; see Appendix A for a summary of relevant studies). Research includes infants (Escudero et al., 2016a; Smith & Yu, 2008, 2013; Vlach & Johnson, 2013; Yu & Smith, 2011), toddlers (Bunce & Scott, 2017; Mangardich & Sabbagh, 2022; Vlach & DeBrock, 2017, 2019), pre-schoolers and children (Fitneva & Christiansen, 2017; Roembke et al., 2018; Suanda et al., 2014; Vlach & DeBrock, 2017, 2019), and adults (Chen, Gershkoff-Stowe, Wu, Cheung, & Yu, 2017; Escudero et al., 2016b; Monaghan et al., 2019; Mulak et al., 2019; Roembke et al., 2018; Yip, 2022). The standard behavioural cross-situational learning paradigm has been combined with various methods, including eye-tracking (in infants: Yu & Smith, 2011; in adults: Fitneva & Christiansen, 2011; Roembke & McMurray, 2016; Trueswell et al., 2013; Yu et al., 2012), event-related potentials (in children: Mangardich & Sabbagh, 2022; in adults: Angwin et al., 2022), fMRI (in adults: Berens et al., 2018), and trial-by-trial behavioural analysis (in adults: Dautriche & Chemla, 2014; Roembke & McMurray, 2021; Suanda & Namy, 2012; Trueswell et al., 2013). The majority of studies have used pseudo-words (meaningless word-like sound sequences following a language's phonotactic rules) and novel objects as stimuli to simulate the learning of novel nouns. Building on this foundation, other studies extended the cross-situational paradigm to explore the learning of word categories (in children and adults: Wang & Trueswell, 2019; in adults: Chen, Gershkoff-Stowe, Wu, Cheung, & Yu, 2017), verbs (in toddlers: Scott & Fisher, 2012; in adults: Monaghan et al., 2015), grammar rules (Monaghan et al., 2019; Rebuschat et al., 2021; Walker et al., 2020), and written words (in adults: Escudero et al., 2022). Moreover, evidence suggests that cross-situational learning performance is closely linked to cognitive factors, such as working memory and language abilities. These correlations have been observed across age groups, including infants (Smith & Yu, 2013; Vlach & Johnson, 2013), toddlers and children (Aravind et al., 2018; Scott & Fisher, 2012; Vlach & DeBrock, 2017, 2019), and adults (Vlach & Sandhofer, 2014; Walker et al., 2020; Wang, 2020). Cross-situational learning studies have investigated how specific variables shape learning outcomes across different populations. These include bilingualism (in adults: Benitez et al., 2016; Escudero et al., 2016c; Poepsel & Weiss, 2016; in children: Crespo et al., 2023), speaker variability (in children: Crespo & Kaushanskaya, 2021), phonotactics (Dal Ben

et al., 2022), prosody (Filippi et al., 2014; Yurovsky et al., 2013), social cues (MacDonald et al., 2017), and temporal order of stimuli presentation (in adults: Apfelbaum & McMurray, 2017; Benitez et al., 2020; Kachergis et al., 2013; in children: Benitez et al., 2020). Other areas of focus include the influence of partial knowledge (Yurovsky et al., 2014), temporal order and repetition of stimuli (Apfelbaum & McMurray, 2017; Benitez et al., 2020; Blythe et al., 2010; Hendrickson & Perfors, 2019; Kachergis et al., 2013; Vogt, 2012), selective attention (Yu et al., 2012), and second language acquisition (in children: Hu, 2017; in adults: Tuninetti et al., 2020), and symbolic flexibility (in children and adults: Roembke et al., 2018). However, despite the key role of gestures in language and communication, to the best of our knowledge, no studies have been conducted on cross-situational learning of linguistic input in the visual modality (e.g. sign language signs, sign-like or other types of context-independent gestures as stimuli).

### ***Semantic processing and event-related potentials (ERPs)***

While cross-situational learning has advanced our understanding of the mechanisms involved in word-referent association, it remains unclear whether this association also encodes semantic representation or is based solely on associative and perceptual mechanisms (Gliga & Csibra, 2009; Waxman & Gelman, 2009). One approach to address this question is the analysis of event-related potentials (ERPs), averaged neural responses to specific sensory, cognitive, and motor events extracted from the EEG signal (Luck, 2004, p. 4).

A widely regarded index of semantic integration (e.g. Kutas & Federmeier, 2011) is the N400 ERP component, characterised by a negative voltage peaking around 400 ms after stimulus onset. The classical N400 paradigm comprises sentences with semantically congruent or incongruent target words (e.g. “He spread his warm bread with butter” vs. “He spread his warm bread with socks”; Kutas & Hillyard, 1980). The effect is sensitive to the degree of congruency and varies according to task and stimulus type (Kutas & Federmeier, 2009; 2011). One way to measure semantic encoding is to see whether a newly learned word can serve as an effective prime in a *semantic priming task*, where incongruent word pairs (e.g. “dog” and “book”) are shown to evoke stronger N400 responses than congruent ones (e.g. “dog” and “bone”). A robust N400 effect has been demonstrated in word learning studies across age groups, including infants and children (e.g. Hirotani et al., 2009; Mangardich & Sabbagh, 2022; see Junge et al., 2021 for a review) and adults (e.g. Angwin et al.,

2014; Armstrong et al., 2024; Bermúdez-Margaretto et al., 2018; Borovsky et al., 2012; Kaczer et al., 2018; Mestres-Missé et al., 2007; Perfetti et al., 2005; Ramos-Escobar et al., 2021), as well as in gesture learning (e.g. in adults: Bernardis et al., 2008; Drijvers & Özyürek, 2018; Habets et al., 2011; Kelly et al., 2004; 2010; Özyürek et al., 2007; Wu & Coulson, 2005; 2007a; 2007b; in children: Sekine et al., 2020; for review see Özyürek, 2014).

Another neural marker of explicit memory retrieval and integration in word learning is the presence of a late positive component (LPC), usually observed following the N400 response within the 450–850 ms post-stimulus window. The presence of an LPC has been associated with controlled, explicit semantic retrieval and integration of new words into the mental lexicon, with its amplitude modulated by retrieval and integration demands (e.g. Bakker et al., 2015; Hoshino & Thierry, 2012; Kaczer et al., 2018; Martin et al., 2009; Rugg & Curran, 2007; for sign-language in hearing population: Zachau et al., 2014).

To date, only two studies have investigated semantic encoding of word-referent associations by combining the standard cross-situational learning paradigm with an ERP analysis (in 4-year-old children: Mangardich & Sabbagh, 2022; in adults: Angwin et al., 2022). For both children (4 years old) and adults, novel words learned via cross-situational learning encoded semantic representations for the new words and showed a similar neurophysiological profile to that of familiar words. Furthermore, the N400 effect has been observed with pseudowords and across modalities, including American Sign Language signs (see Kutas & Petten, 1994). However, to date, no study has examined the N400 effect elicited by other forms of visually presented language, such as context-independent sign-like gestures.

### Current study

Building on these findings, the current study examines whether context-independent gestures can be effectively learned and assigned meaning, focusing on sign-like gestures. Their similarity to words – arbitrariness, symbolic properties, and referential function – suggests that these gestures have the potential to act as labels for referents and function as an ecological linguistic input. In light of the lack of studies on cross-situational learning of language in the visual modality and the shortage of data regarding the neurophysiological underpinnings of cross-situational learning mechanisms, our study aimed to explore the acquisition of gestures corresponding to familiar spoken words. We investigated whether cross-situational learning would allow the association of gestures with familiar words and their meaning,

employing a combination of behavioural and event-related potential methodologies. Given that cross-situational learning of visual linguistic stimuli has not been previously investigated, including both children and adults allowed us to conduct a preliminary exploration of broad developmental differences in this learning process.

To ensure linguistic salience and significance, our visual stimuli were conceived as language-like pseudosigns. We defined a pseudosign as an arbitrary hand movement, phonotactically legal in sign languages, with no previous cultural meaning, and no iconic relation to the referent it represented (however, despite our attempts, our stimuli still displayed a varying degree of iconicity; see Appendix B for a full report of stimuli validation). The phonological legality of the pseudosigns was verified using the phonological framework of Italian Sign Language (LIS) as described in Branchini and Mantovan (2020), drawing on the first author's (AC) proficiency in this language (see “Visual stimuli”). Following these criteria, we created 8 novel pseudosigns and paired them with 8 familiar spoken words,

To test our hypothesis, first, we familiarised our participants with the 8 word-pseudosign items in a *familiarisation phase* using a standard cross-situational learning paradigm (Yu & Smith, 2007). Then, we used two semantic priming tasks to assess whether the pseudosigns could be correctly associated with their intended spoken words (*recognition task*) and if this association could be extended to the semantic area (*categorisation task*). Similar to previous studies (Angwin et al., 2022; Mangardich & Sabbagh, 2022), we combined the cross-situational learning paradigm and the behavioural measures with the analysis of event-related potentials (ERPs) and investigated the amplitude and latency of the N400 component as a marker of lexical and semantic learning. Based on previous research, we formulated two main hypotheses: (1) We hypothesised that participants in both groups would successfully associate pseudosigns with the intended spoken words. This would be evidenced by significantly higher scores than chance in the *recognition task*, along with a greater N400 negativity when a pseudosign was followed by a different word than the intended one; (2) We also hypothesised that participants would extend the meaning of spoken words to pseudosigns, effectively assigning them a meaning. This would be reflected in above-chance performance in the *categorisation task* and a greater N400 response when a pseudosign was followed by a spoken word from a different semantic category than the intended one. Moreover, we formulated two additional hypotheses: (1) We expected higher accuracy in the *recognition task* compared to the *categorisation*

task, as the *recognition task* required direct retrieval of learned word-pseudosigns pairs, while the *categorisation task* involved a more indirect form of semantic processing. At the neural level, we predicted a reduced N400 amplitude in the *categorisation task*, as indirect semantic associations generally elicit weaker N400 effects than direct pairings (e.g. Kreher et al., 2006); (2) Lastly, we hypothesised a difference in performance between the two groups. General language and memory skills have been linked to cross-situational learning performance in children (Vlach & DeBrock, 2017, 2019), although similar results do not hold within the adult population (e.g. Hamrick, 2015; Walker et al., 2020). Since cognitive abilities, such as working memory capacity and attention, have been shown to differ between age groups (e.g. Booth et al., 2003; Hoyer et al., 2021; for a review, see Cowan, 2016), we hypothesised that adults' cognitive and linguistic maturity facilitated learning and improved task performance, with individual language abilities also contributing to the variability in results. Although we did not measure individual cognitive and linguistic abilities in our study, we expected adults to outperform children in the behavioural tasks, likely due to their greater cognitive and linguistic maturity. At the neural level, however, given evidence that children as young as 6 years old can exhibit adult-like N400 responses (e.g. Coch, 2014; Sekine et al., 2020), we predicted that no significant group differences would emerge in N400 amplitude. In addition, considering the narrow age range of the children's group (8–12 years old), no predictions were made regarding any differences within this group.

## Materials and methods

### Participants

A total of 26 children (14 female, 12 male,  $M_{age} = 9.11$  years, range = 8.1–11.11 years,  $SD = 14.8$ ) and 19 adults (11 female, 8 male,  $M_{age} = 27$  years, range = 18–35 years,  $SD = 4.9$ ) participated in the study. All participants were healthy and normally hearing, native English speakers, having attended school or higher education in Australia. Before the start of the experiment, a hearing screening test comprising pure tone audiometry and otoacoustic emissions was performed to check for normal hearing sensitivity on the day of the experiment. To prevent biasing the participant towards the nature of the experiment, previous exposure to sign language was not mentioned as an exclusion criterion. In fact, if participants had been screened before the experiment, they might have become aware of the linguistic nature of the stimuli, potentially introducing bias toward the

learning aspect of the task. This information was only collected once the experiment was done. Of the 26 children initially recruited, one participant reported substantial experience with sign language and was therefore excluded from the final analysis, resulting in a total sample of 25 children. Participants gave informed consent prior to the start of the experiment. For child participants, consent was given both by the parent/legal guardian and the child, who gave verbal assent to the experiment. After completing the study, participants were provided with a token of appreciation in the form of a \$50 eGift Card. The study was approved by the Macquarie University ethical committee (approval number 520231271648518).

### Stimuli

Stimuli for the familiarisation phase consisted of 8 target words (audio stimuli) and 8 novel pseudosigns (videos). Each spoken target word was matched with a pseudosign to create an item that remained consistent throughout the entire duration of the experiment and for which the learning was tested.

### Verbal stimuli

Target words have been selected from Posnansky (1978) and refer to familiar objects (*bed, dog, car, cold, cup, pink, shirt, toe*) from 8 different semantic categories (*furniture, animals, vehicles, weather, kitchenware, colours, clothes, body parts*). The choice of these words and categories was guided by the study design, specifically by the constraints of the semantic task. First, to assess learning of both gesture form and meaning, the words had to be easily categorisable. Second, the words and categories had to be familiar to the child group. Lastly, monosyllabic words were selected to facilitate ERP analysis, as they reduce variability in acoustic properties and ensure a more precise temporal alignment for ERP components such as the N400. Posnansky (1978) provides validated category norms for verbal items in children aged 7–12, with the same categories also supported for adult populations (Van Overschelde et al., 2004). Each of our target words represented a distinct semantic category and was paired with seven additional monosyllabic words from the same category. This resulted in a total of eight target words and 54 additional words for use in the categorisation task. The audio stimuli were recorded by an adult female native speaker of Australian English using monotone intonation.

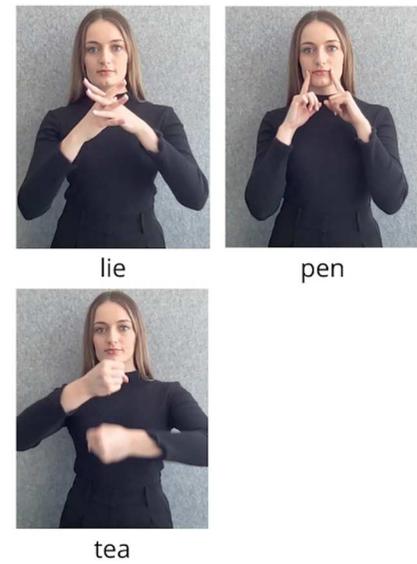
### Visual stimuli

For the visual stimuli (see Figure 1), 8 pseudosigns were created through the following process. To establish a

(A) Target word-pseudosign items



(B) Cover task



**Figure 1.** Static depiction of the 8 items made of a pseudosign linked to the matched target word (A) and the word-pseudosign items used in the cover task (B).

lexical base, signs corresponding to the target words were selected from various sign language corpora – Brazilian Sign Language (Libras), Czech Sign Language (CSE), Finnish Sign Language (FINSL), Lithuanian Sign Language (LGK), Polish Sign Language (PJM), and Portuguese Sign Language (LGP) – available in the online dictionary *Spread the Sign* ([www.spreadthesign.com](http://www.spreadthesign.com)). These signs were modified in one or more phonological parameters (handshape, orientation, location, and movement) to meet the experimental requirements and create one-handed pseudosigns while ensuring they remained phonologically plausible. The final set consisted of eight one-handed pseudosigns, each designed to be maximally distinct through systematic variation of parameter combinations. A detailed phonological description of the pseudosigns is provided in the supplementary material (Appendix C), following the framework proposed by Branchini and Mantovan (2020). This process was followed to avoid any possible risk of previous exposure of the participants to the experimental stimuli while keeping the pseudosigns visually salient and linguistically meaningful. The stimuli were then recorded as short individual video clips by an actress with sign language proficiency (Australian Sign Language), ensuring that the pseudosigns were produced in a natural, sign language-like manner. The actress performed each pseudosign with the right hand, facing the camera with a neutral expression and without producing any other movement, facial expression, or vocalisation. To keep their attention during the *familiarisation phase*, participants were instructed to press a button every time they saw a

pseudosign made with two hands instead of one (cover task). For this cover task, three additional pseudosigns were created (*lie*, *pen*, *tea*), in which the actress performed the pseudosign with two hands. Original video clips were edited using Adobe Premiere to align the onset, duration, and offset of each gesture. Each clip featured a 3-second gesture, preceded and followed by 1-second no-movement intervals where the hands remained still. The gesture stroke began 1 s after the start of the video clip. As a result, all gestures were synchronised, with no noticeable differences.

### Procedure

The experiment comprised three phases, starting with (1) a *familiarisation phase* for cross-situational learning followed by (2) a behavioural *recognition task* with a yes/no choice and (3) a behavioural *categorisation task*. Stimulus presentation and response collection were controlled using the software Presentation® version 24.0 (Neurobehavioral Systems). The appointment lasted ~120 minutes to complete and included hearing screening, electrode application and removal, and short breaks between the tasks. Participants sat in an armchair at ~1.5-meter distance from the screen for the visual stimuli and the speaker for auditory stimuli. To manage participants' fatigue, short breaks of about 5 minutes were taken between tasks. During these breaks, the experimenter checked the EEG electrode impedance, and the participants were offered a small, non-caffeinated drink. The time was also used to give verbal instructions regarding the next task. The

instructions were also repeated in written form on the screen. EEG signal was recorded during the duration of the experiment, and the ERPs extracted from the *recognition* and *categorisation* tasks were analysed offline to extract N400 amplitude during the presentation of congruent vs. incongruent pairs.

### **Familiarisation phase**

For the *familiarisation phase*, we applied a covert cross-situational learning task of the word-pseudosign items. The task comprised 48 trials. In each trial, two videos were presented simultaneously on a screen and played side by side in conjunction with two auditory stimuli (spoken words) played one after the other. The trial started with a fixation cross. One second later, the videos were presented in 400 × 300 pixels in size and were centred 250 pixels to the right or left of the screen's centre. The first audio was presented 750 ms after the start of the video, and the second audio was presented 1500 ms after the start of the first audio. The inter-stimulus interval was 3 s. Throughout the trials, the words and pseudosigns were always matched consistently; that is, the same word was always associated with the same pseudosign, creating an item. The word and associated pseudosign were always presented together. For instance, if the spoken word was “dog,” one of the videos had the pseudosign for “dog”. In 50% of the trials, the first spoken word referred to the video on the left and the second to the video on the right. In the other half, this order was reversed (see Figure 2). The same word-pseudosign item was never presented twice within the same trial. Importantly, participants were not informed about the presence of word-pseudosign associations, nor the learning nature of the task. Participants were instructed to pay attention to the screen and press a button whenever they saw a two-handed gesture (cover task). As two word-pseudosign items were presented on each trial, participants were exposed to each pair a total of 12 times across the 48 trials of the *familiarisation phase*. In addition, 4 cover task trials were presented randomly between trial numbers 5-11, 15-22, 25-32, and 35-42. The cover task was included to maximise participant attention to the screen and ensure all participants received the same amount of exposure to experimental items.

### **Recognition task**

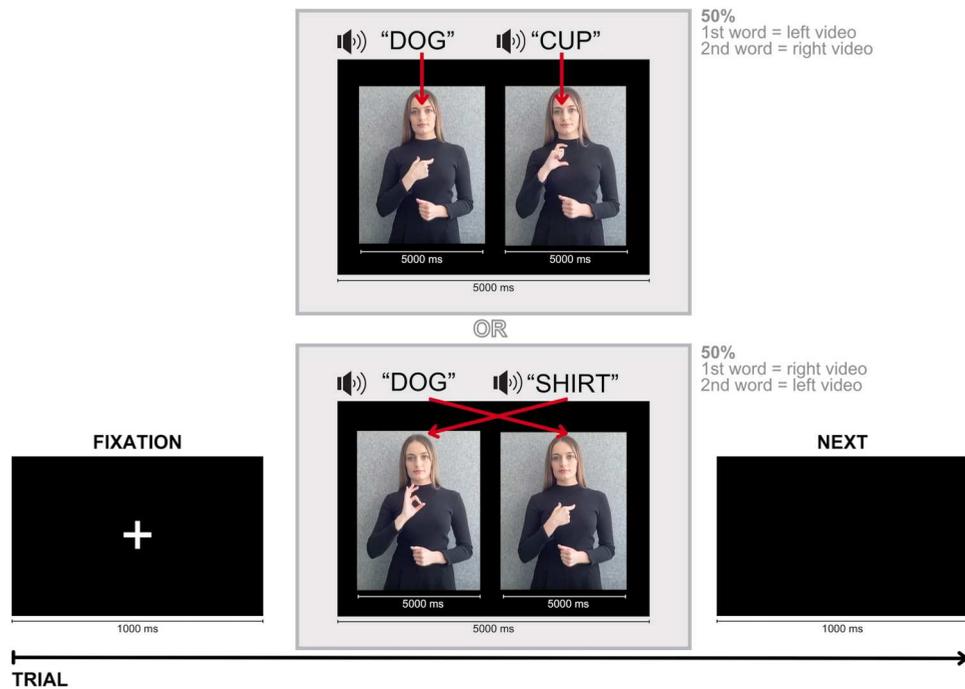
During the *recognition task*, the participants' learning of the word-pseudosign items was tested. This task comprised a total of 96 trials, 48 congruent and 48 incongruent. Each trial consisted of a fixation cross “+” displayed on screen for 1000 ms, followed by the presentation of one pseudosign, followed by one spoken word. A 750-

ms interval of silence was inserted between the video and the audio (see Figure 3A). Participants were asked to indicate via button press on a response pad if the spoken word was or was not the correct match for the pseudosign. No feedback was given to participants about the correctness of their answers. The question remained on the screen until either the participants responded or if 3 s had elapsed, whichever occurred first. A 2-second blank screen preceded the next trial.

### **Categorisation task**

This task was designed to explore how well the participants had learnt the pseudosigns. Once the task began, participants were presented with one pseudosign (from the same set of the *familiarisation phase*) followed by one spoken word, different from the target words but pertaining to one of the eight semantic categories. Participants were asked to indicate whether the spoken word belonged to the same or a different category of the pseudosign (see Figure 3B). No feedback was given about the correctness of their answers. For half of the trials, the semantic categories of the two words were congruent (e.g. pseudosign “dog”, spoken word “cat”). On the other half of trials, the spoken word pertained to a different category, hence considered incongruent (e.g. pseudosign “dog”, spoken word “shirt”). Consequently, to respond to the trials accurately, participants had to draw upon their knowledge of the pseudosigns' meaning learned in the *familiarisation phase*. The trial structure and timing were similar to the recognition task.

Before this task, a practice trial was performed to ensure all participants understood the task, especially in the child group. Using flashcards instead of videos and live voice instead of audio recordings, the experimenter recreated the task by showing the participant a drawing of an object pertaining to the same experimental categories – but not used in any other experimental phases – e.g. a drawing of a *kangaroo*/ANIMALS and then pronouncing a category-congruent or incongruent word such as “mouse”/ANIMALS – congruent or “taxi”/VEHICLES – incongruent. To clarify the nature of the *categorisation task*, some intentionally misleading pairs were introduced to induce responses based on association rather than categorisation. For example: a drawing of a *yellow stain* (category COLOURS) paired with the spoken word “lemon” (misleading association: “the lemon is yellow”) or “purple” (correct association: COLOURS), and a drawing of a *mouth* (category BODY PARTS) paired with the spoken word “cake” (misleading association: “the cake is eaten with the mouth”) or “leg” (correct association: BODY PARTS). In case the participant gave the incorrect answer, the experimenter stressed



**Figure 2.** Static depiction of word-pseudosign learning phase (*familiarisation phase*).

that the words should be considered congruent or incongruent according to their categories and not by other associations such as “the lemon is yellow” or “the cake is eaten with the mouth”.

### EEG recording

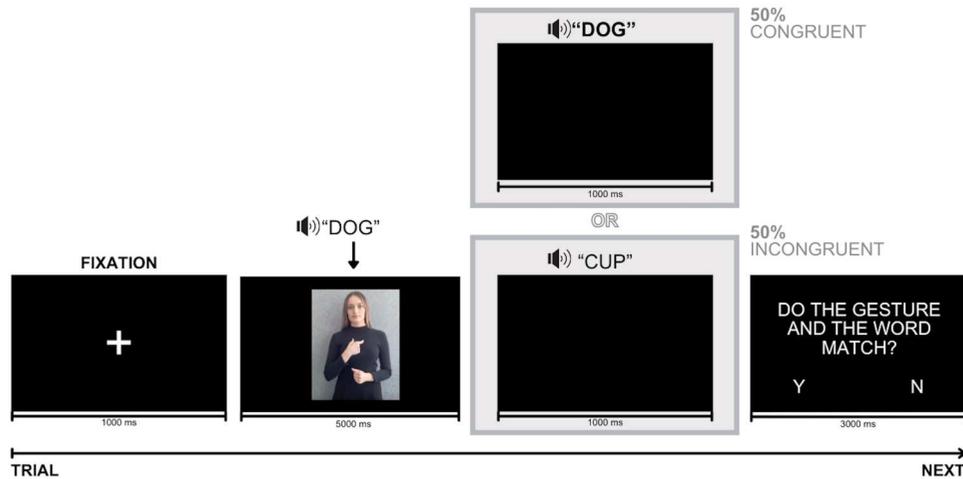
EEG was recorded for the entire duration of the experiment, using a 64-electrode cap (Easycap) on the scalp according to the 10–20 system. EEG was continuously recorded at a sample rate of 1000 Hz using a SynAmps 2 Neuroscan<sup>®</sup> amplifier. The reference electrode (M1) was set on the left mastoid, and additional data were recorded from the right mastoid (M2). Vertical and horizontal oculograms (VEOG, HEOG) were recorded from electrodes placed above and on the side of the right eye and the left and right temple, respectively. The experiment was conducted in an electric – and acoustic-shielded room. Electrode impedances were set at <math><10\text{ k}\Omega</math> at the start of the recording and maintained throughout the experiment.

### EEG analysis

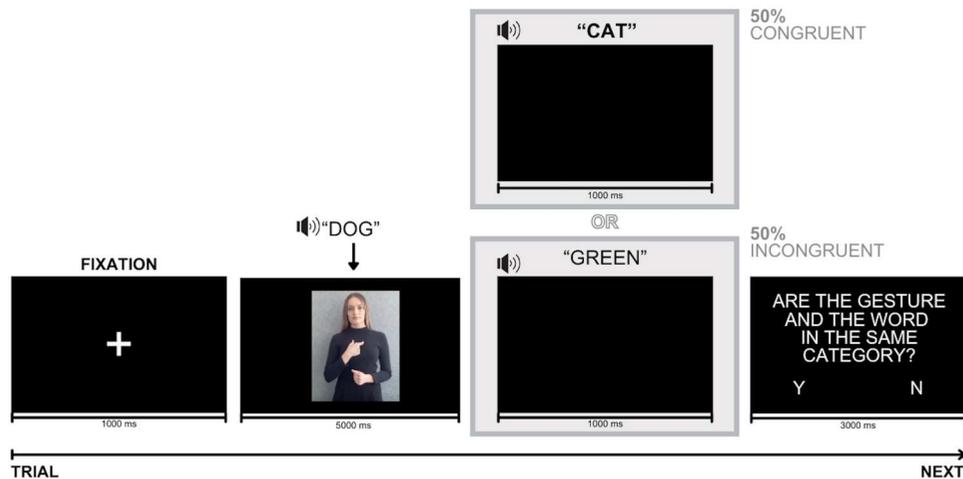
Continuous EEG was recorded for the entire duration of the experiment, but data were only analysed for the *recognition* and *categorisation tasks*. EEG data analysis was conducted in Matlab v.R2023b (The MathWorks Inc., 2023) using a custom script and FieldTrip toolbox (Oostenveld et al., 2011). EEG continuous recording was

downsampled to 250 Hz and bandpass filtered from 0.3–30 Hz using a windowed sinc finite impulse filter and then divided into epochs between  $-100$ – $1000$  ms from stimulus onset. As the pre-stimulus activity is expected to be similarly low across trials, baseline correction was applied between  $-100$ – $0$  ms before stimuli, and the activity in this time window averaged and subtracted throughout the waveform. Epochs and channels contaminated by artefacts were identified and removed in three steps. First, the EEG epochs were manually inspected, and trials with amplitude exceeding  $\pm 500\text{ }\mu\text{V}$  and channels that are consistently noisy were removed from the data (adults:  $M = 0.4$  channels, range [0–1]; children:  $M = 1.19$  channels, range [0–5]). The independent component analysis (ICA) was then performed, and components with the stereotypical features of blinks and eye movements were removed from the data (adults:  $M = 3.8$ , range [2–7]; children:  $M = 2.88$ , range: [2–5]). After this, epochs with amplitudes exceeding  $\pm 100\text{ }\mu\text{V}$  were removed. Any removed channels were then interpolated using a spherical spline interpolation. The epochs were then re-referenced to the average of the mastoids. ERPs were then calculated by averaging both correctly and incorrectly identified trials. Including both types of trials ensured a higher signal-to-noise ratio by increasing the number of trials available for analysis. This approach also helps capture neural responses that might not align with behavioural performance. In fact, studies have shown that the N400 can occur even without conscious identification

(A) Recognition task



(B) Categorization task

**Figure 3.** Static depiction of the *recognition task* (A) and *categorisation task* (B).

(Stenberg et al., 2000). Individual ERPs were obtained for each participant and averaged to obtain a grand average for both groups (adults and children) and both conditions (*recognition task* and *categorisation task*). Data from individual participants were included if more than 35 good epochs out of 48 for each condition

**Table 1.** Signal Detection Analysis.

Participant response	Congruent trial (Signal Present)	Incongruent trial (Signal Absent)
YES (the pair is congruent)	HIT (e.g. <i>rec.</i> : dog – dog; <i>cat.</i> : dog – bear)	FALSE ALARM (e.g. <i>rec.</i> : dog – car; <i>cat.</i> : dog – storm)
NO (the pair is incongruent)	MISS (e.g. <i>rec.</i> : dog – dog; <i>cat.</i> : dog – bear)	CORRECT REJECTION (e.g. <i>rec.</i> : dog – car; <i>cat.</i> : dog – storm)

Note. Participants were presented with the question “Did the word and the gesture match?” appearing on the screen and had to give their response by pressing YES or NO on a response pad.

were available (75% good trials). All participants were included in this step.

### Statistical analysis

Analyses were conducted using the R Statistical language (version 4.2.2; R Core Team, 2022) on macOS Big Sur 11.7.4, using the packages lme4 (version 1.1.34; Bates et al., 2015).

### Behavioural analysis

For behavioural data, the accuracy score was calculated as the percentage of correct trials out of the total number of trials ( $n = 96$ ). To account for individual differences and response strategies, we used signal detection analysis (Macmillan & Creelman, 2004) and calculated detection sensitivity (i.e.  $d$ -prime and response bias measures;

Table 1). These two scores allowed us to quantify participants' ability to correctly identify the congruent and incongruent pairs and model their response strategy (more or less prone to respond "congruent" or "incongruent"). *D*-prime scores were calculated as the standardised difference between the "hit rate" (i.e. the proportion of correctly identified trials relative to the total congruent trials) and the "false alarm rate" (i.e. the proportion of incorrectly identified incongruent trials relative to the total incongruent trials). Larger *d*-prime values implied higher sensitivity, and values near zero indicated at-chance performance. Response bias was calculated as the difference between the *z*-scores of "hit rate" and "false alarm rate". To address potential issues with hit rates or false alarm rates of 0% or 100%, which can affect *z*-score calculations, we applied a smoothing procedure. Specifically, 0.5 was added to the number of hits and false alarms, and the totals were divided by the sum of signal-present or signal-absent trials plus one. This approach ensures that the calculated rates are bounded and avoids infinite *z*-scores. These smoothed hit rates and false alarm rates were then transformed into *z*-scores, and the *d*-prime was calculated by subtracting the *z*-score of the false alarm rate from the *z*-score of the hit rate. Positive *d*-prime scores suggest a more conservative bias; that is, participants are more likely to respond "no, the pair is incongruent" even when uncertain, leading to fewer "false alarms" but more misses. Negative scores depict a liberal bias, that is, participants are more prone to say that the pair is congruent even when uncertain, leading to more "false alarms" but fewer "misses".

A linear model, estimated using Ordinary Least Squares, was used to examine the effect of group (i.e. adult vs. children) and task type (i.e. recognition vs. categorisation) on the *d*-prime scores and the interaction between the two variables ( $d\text{-prime} \sim (\text{Task} * \text{Group})$ ). Standardised parameters were obtained by fitting the model on a standardised version of the dataset, and 95% Confidence Intervals (CIs) and *p*-values were computed using a Wald *t*-distribution approximation. Given that only one child participant was excluded from the analysis and none of the remaining participants had prior exposure to sign language, sign language proficiency was not included as a covariate in the analysis. Similarly, exploratory analyses with age as a covariate revealed that, within our child group, age did not play a significant role in predicting performance ( $p = .621$ ); hence, this factor was not considered further.

### ERP analysis

The significance of the ERP responses was determined using the cluster-based permutation test (CBPT) implemented in the Fieldtrip toolbox in Matlab. This

test is commonly used to analyse EEG data without a priori assumptions about specific temporal effects (Maris & Oostenveld, 2007). The CBPT computed multiple *t*-tests at all electrodes and all time points in the 0.1–1 s time interval. Alpha-level was set at 0.05. Clusters were formed over time and space by grouping at least two nearby electrodes with significant initial *t*-test results at the same time point and then by grouping at least two adjacent time points with significant *t*-values. A cluster *t*-score (mass *t*-score) was then calculated as the sum of all *t*-values within each cluster. A permutation approach was used to control for Type I errors. In order to build a data-driven null hypothesis distribution, standard and deviant waveforms were randomly assigned to conditions, and multiple *t*-tests were repeated in 2000 iterations.

If a significant effect was found in the CBPT, Cohen's *d* was calculated as the effect size. The CBPT showed 2-time windows where the congruent and incongruent ERPs differed, with one of the effects being negative and the other one positive. The topography and time range of the first effect was that of N400, and the second effect was that of LPC. The amplitude of these components was calculated as the mean amplitude in a 200 ms window around the peak of the effect in the grand averaged waveform from 15 electrodes in the centro-parietal region (C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2, CP4, P3, P1, Pz, P2, P4). The time range and electrodes were selected based on previous studies on N400 (Duncan et al., 2009; Šoškić et al., 2022).

For statistical analysis of ERP data, we conducted a series of *t*-tests to check for significant differences in mean N400 amplitude between groups and tasks. Pearson's product-moment correlations were used to check for correlations between ERP and behavioural results. For age, research has shown that by grade 3 (approximately 8 years), children exhibit adult-like N400 responses for semantic processing of words (Coch, 2014) and gestures (Sekine et al., 2020). Consequently, significant age-related differences in ERP components were not expected within this range. Indeed, exploratory analyses with age as a covariate revealed that, within our child group, age did not play a significant role in predicting neural responses ( $p = .715$ ).

## Results

### Behavioural results

To assess the normality of accuracy and *d*-prime scores, we conducted a Shapiro–Wilk test for each group and task, applying Bonferroni corrections for multiple comparisons. For accuracy, the test indicated significant

deviations from normality in the adult group (recognition:  $p = .015$ ; categorisation:  $p = .007$ ) but not in the child group (recognition:  $p = 1.00$ ; categorisation:  $p = .296$ ). For the  $d$ -prime scores, the test indicated no significant deviation from normality in both groups and tasks (all Bonferroni-adjusted  $p > .05$ ). Given these results, we used the non-parametric Wilcoxon signed-rank test to analyse accuracy scores, as normality assumptions were violated in the adult group. For  $d$ -prime scores, we used parametric one-sample  $t$ -tests to check whether the scores significantly differed from zero (chance level). Both tests showed that accuracy and  $d$ -prime scores were significantly above chance in both groups and tasks (all  $p < .001$ ). For response bias, the Shapiro–Wilk test showed no deviation from normality in adults ( $W = 0.958$ ,  $p = .538$ ) but a significant deviation in children ( $W = 0.905$ ,  $p = .024$ ). As a result, we used the non-parametric Mann–Whitney  $U$  test to compare response bias between the two groups. The calculation showed that both groups were more conservative (i.e. more prone to respond that the word-pseudesign items were incongruent) in their decision-making, with no significant difference in response bias between the two groups ( $U = 0.286$ ,  $p = .255$ ). Behavioural results (accuracy,  $d$ -prime scores, and response bias) are reported in Table 2.

As expected, the linear model ( $d$ -prime  $\sim$  (Task \* Group)) showed a significant main effect of group,  $F(3, 84) = 6.55$ ,  $p < .001$ ,  $R^2 = .16$ , with adults performing better than children ( $\beta = -1.24$ , 95% CI  $[-1.81, -0.68]$ ,  $t(84) = -4.36$ ,  $p < .001$ ). Contrary to expectations, no significant effect of task was found ( $\beta = -0.28$ , 95% CI  $[-1.14, 0.57]$ ,  $t(84) = -0.66$ ,  $p = .512$ ), neither a significant interaction between task and group ( $\beta = 0.10$ , 95% CI  $[-1.04, 1.23]$ ,  $t(84) = 0.17$ ,  $p = .864$ ) (Figure 4).

### ERP results

The CBPT results are reported in Table 3. In the adult group, the main analysis identified clusters of electrodes and time points with significant differences between congruent and incongruent trials. A negative effect, consistent with the N400 component, peaked at approximately 400 ms in the *recognition task* and at 646 ms in the *categorisation task*, both with a centro-parietal distribution (see Figure 5). For the child group, a cluster of significant differences between trial types was observed in the *recognition task*, with a negative effect peaking at about 368 ms and a centro-parietal distribution (Figure 6). No clusters of significant differences were detected in the *categorisation task* for the child group, indicating that no N400 effect was detected. These findings contrasted with the results of the behavioural analysis and

motivated additional investigation to further explore this discrepancy between behavioural and ERP results (see *Exploratory analysis: EEG analysis of correct trials only*).

To check for any significant difference in the mean N400 between groups during the *recognition task*, we used a linear model (mean N400  $\sim$  Group). The results indicated no significant difference between groups,  $F(1, 42) = 0.07$ ,  $p = .799$ ,  $R^2 = -.02$ , with adults ( $M = -1.30$ ) and children ( $M = -1.40$ ) exhibiting comparable N400 amplitude ( $\beta = -0.11$ , 95% CI  $[-0.94, 0.73]$ ,  $t(42) = -0.26$ ,  $p = .799$ ). This suggests that both groups exhibit similar neural responses during the *recognition task*. Comparison of mean N400 amplitude during the *categorisation task* was not possible in the child group as no significant differences in time points or electrodes were identified, indicating the absence of an N400 effect.

To test for any significant differences in the N400 amplitude between tasks within the adult group, we fitted a linear model (Amplitude  $\sim$  Task). The results indicated no significant effect of task,  $F(1, 36) = 3.75$ ,  $p = .061$ ,  $R^2 = .094$ , with no statistical difference in N400 amplitude between *recognition* and *categorisation* tasks ( $\beta = 0.51$ , 95% CI  $[-0.02, 1.04]$ ,  $t(36) = 1.94$ ,  $p = .061$ ) (Figure 7A). The same comparison was not possible in the child group, as analysis of ERP data from the *categorisation task* did not identify any clusters showing significant differences between conditions.

Finally, we explored the relationship between behavioural performance and mean N400 amplitude using Pearson's correlations. In the adult group, the correlation between accuracy and N400 amplitude was not significant (recognition:  $r(17) = -0.42$ ,  $t(17) = -1.92$ ,  $p = .072$ ; categorisation:  $r(17) = -0.40$ ,  $t(17) = -1.83$ ,  $p = .086$ ). While these results do not provide conclusive evidence, they suggest a potential indication of a negative relationship that could be explored further in future studies. In the child group, the correlation between recognition accuracy and N400 amplitude was also negative and not statistically significant ( $r(23) = -0.34$ ,  $t(23) = -1.73$ ,  $p = .098$ ). Interestingly, categorisation accuracy and N400 in this group showed a significant positive correlation ( $r(12) = 0.54$ ,  $t(12) = 2.20$ ,  $p = .048$ ), indicating higher categorisation accuracy associated with more negative N400 amplitudes.

### Exploratory analyses: Late Positive Component (LPC) analysis

Although not predicted, a Late Positive Component (LPC) was observed in the *recognition task* for both adults and children, but not in the *categorisation task*. Adults showed a positive cluster peaking at approximately 825 ms, with a centro-parietal distribution,

**Table 2.** Descriptive statistics of behavioural performance of the two groups.

Group	Task	Accuracy (%)			<i>d</i> -prime			Response Bias <i>M</i>
		<i>M</i>	<i>SD</i>	Wilcoxon <i>W</i>	<i>M</i>	<i>SD</i>	<i>t</i> -statistics	
Adults ( <i>N</i> = 19)	Recognition	87.5	13.6	190, <i>p</i> < .001	2.76	1.27	9.47, <i>p</i> < .001	0.398
	categorisation	83.8	16.6	170, <i>p</i> < .001	2.48	1.34	8.04, <i>p</i> < .001	
Children ( <i>N</i> = 25)	Recognition	72.0	17.5	314, <i>p</i> < .001	1.47	1.33	5.52, <i>p</i> < .001	0.292
	categorisation	69.0	18.2	295, <i>p</i> < .001	1.29	1.34	4.79, <i>p</i> < .001	

Note. *M* = mean, *SD* = standard deviation, and correlation coefficient, respectively. Accuracy scores were analysed with the Wilcoxon signed-rank test and *d*-prime scores were analysed using one-sample *t*-tests to determine if they were significantly above chance levels. Bonferroni corrections were applied to *p*-values to control for multiple comparisons.

while children showed a significant positive cluster peaking at around 756 ms, with a prevalently parietal distribution (Table 3). As the LPC wave was found for the *recognition task* only but in both adults and children, the effect was compared between the groups by fitting a linear model ( $Amplitude \sim Group$ ). The results indicated no significant effect of group,  $F(1, 42) = 0.016$ ,  $p = .901$ ,  $R^2 = -.02$ , with no statistically significant difference in LPC amplitude between adults and children ( $\beta = 0.04$ , 95% *CI* [-0.66, 0.75],  $t(42) = 0.13$ ,  $p = .901$ ) (Figure 7B). In addition, no correlation was found between behavioural outcomes and mean LPC amplitude in either group (adults:  $r(17) = 0.31$ ,  $t(17) = 1.35$ ,  $p = .193$ ; children:  $r(23) = -0.05$ ,  $t(23) = -0.22$ ,  $p = .824$ ).

### Exploratory analyses: EEG analysis of correct trials only

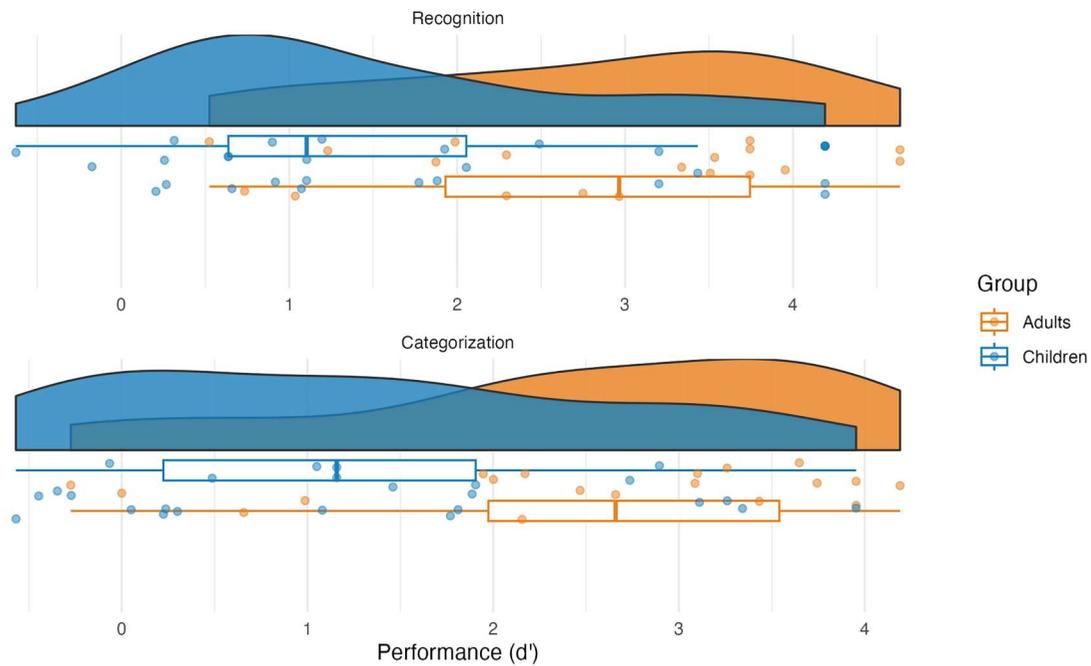
After the main analysis was conducted, no clusters showed significant differences between congruent and incongruent trials for the child group in the *categorisation task*. This absence of an N400 effect was unexpected, given multiple pieces of evidence suggesting learning in both tasks. First, the behavioural data showed above-chance performance in both *recognition* and *categorisation tasks*, indicating that children had successfully learned in both conditions. Second, the linear model applied to the behavioural results did not reveal any significant difference between the two tasks, suggesting comparable performance levels. Finally, the ERP analysis for the recognition task revealed an N400 effect, providing electrophysiological evidence of learning in support of behavioural data. Taken together, these findings made the lack of an N400 effect in the *categorisation task* surprising. To address this discrepancy, we proceeded with further exploratory analysis focused exclusively on correct trials. This approach aimed to minimise the influence of task performance errors (e.g. lapses of attention or other factors unrelated to the process under investigation) and to target the cognitive processes of interest more directly. Importantly, this analysis was not planned a priori and was conducted to gain deeper insights and offer a more comprehensive understanding of the data.

The continuous EEG was analysed following the same steps as described in the previous paragraph, with the only difference being that epochs were selected only for trials in which participants provided a correct response. Data from individual participants were included if more than 24 good epochs out of 48 for each condition (50% good trials). 14 out of 25 participants were included in this analysis. The additional analysis comprising ERPs from correct trials only revealed a significant negative cluster ( $p = .04$ ) in the time window 760–876 ms (Table 4; Figure 8).

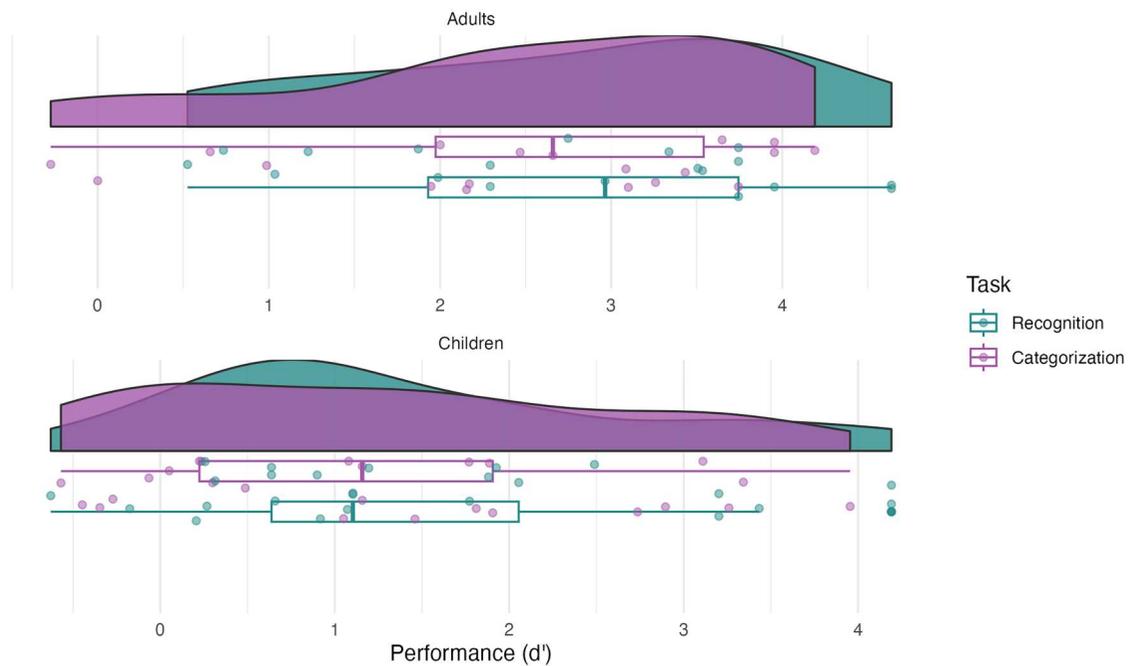
## Discussion

In this study, we investigated both behavioural and electrophysiological markers of the learning of unfamiliar sign language-like gestures (in the form of pseudosigns) corresponding to familiar spoken words in adults and children, using a statistical learning paradigm known as cross-situational learning. Following a brief exposure to eight word-pseudosign pairs during a *familiarisation phase*, we assessed the association of pseudosigns with spoken words through two yes/no tasks: (1) a *recognition task* to test the learning of pseudosign forms and (2) a *categorisation task* to evaluate the learning of pseudosign meanings. To account for individual differences, participants' performance was measured using sensitivity scores (*d*-prime) to distinguish correctly identified congruent from incongruent word-pseudosign items. Additionally, an EEG signal was recorded, and event-related potentials (ERPs) were analysed to identify markers of semantic processing differences between congruent and incongruent word-pseudosign items, focusing on the N400 response. Our hypothesis was that both groups would successfully associate pseudosigns with spoken words and their meanings, as reflected in above-chance performance and an N400 effect. We expected higher accuracy in the *recognition task* than in the *categorisation task*, with a reduced N400 amplitude in the latter. Additionally, we predicted that adults would outperform children in the behavioural tasks, though no group differences in N400 amplitude were expected.

## (A) Group comparison



## (B) Task comparison



**Figure 4.** Comparisons of performance scores expressed as sensitivity scores ( $d'$ -prime). The graph shows probability density, means, medians, and interquartile range for the  $d'$ -prime scores in the two tasks and groups. A small amount of random jitter was added to the raw data points to facilitate figure interpretation. (A) By-group comparison shows a better performance of adults over children. (B) By-task comparison reveals no difference in the performance in the two tasks for both groups.

### Behavioural results

In both tasks, adult and children's performance was significantly above chance, confirming successful learning

of the pseudosigns' forms and meanings. In the *recognition task*, pseudosigns were reliably associated with their corresponding spoken words, demonstrating that participants learned the specific word-pseudosign

**Table 3.** Cluster statistics results with effect sizes.

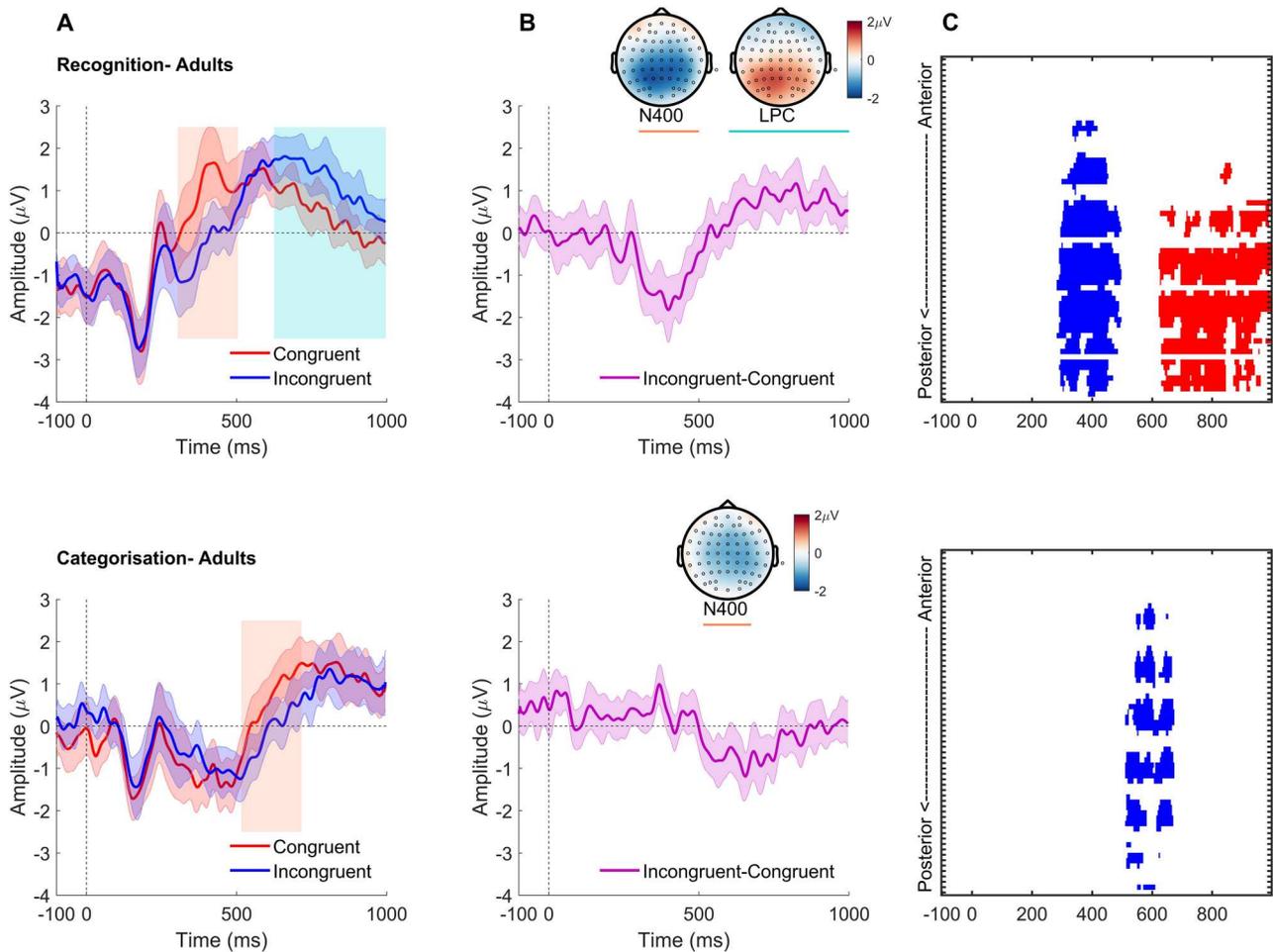
Group	Task	Cluster type	Latency	$p$	Cohen's $d$
Adults ( $N = 19$ )	Recognition	Negative	285–497	.001	-1.77
		Positive	625–997	.001	1.62
Children ( $N = 25$ )	Categorisation	Negative	513–673	.005	-1.44
	Recognition	Negative	156–464	.001	-1.39
		Positive	572–896	.002	1.23

pairings presented during familiarisation. In the *categorisation task*, both adults and children were able to draw upon the pseudosign semantic connection with their referent to correctly distinguish between congruent and incongruent categories. This result indicated that they successfully extended the meaning of the spoken word to the pseudosign, thus learning the pseudosign's meaning.

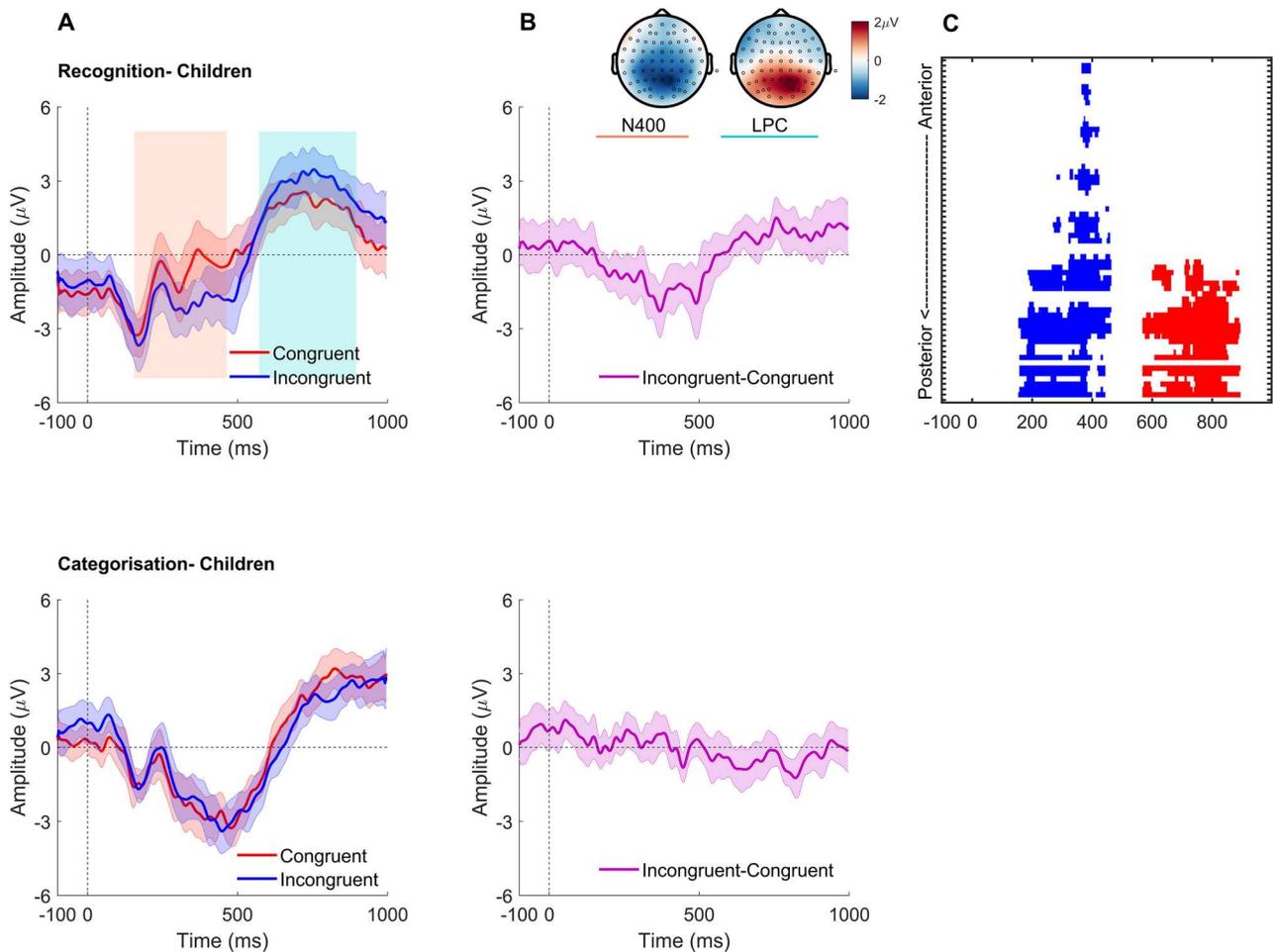
Following expectations, adults performed better than children in both tasks. Such a difference was anticipated

due to adults' cognitive and linguistic advantage, which we assumed could facilitate learning and task performance. This advantage is consistent with evidence showing that working memory and attention mature with age (e.g. Booth et al., 2003; Hoyer et al., 2021; see Cowan, 2016 for a review) and that cross-situational learning performance in children is predicted by individual cognitive and language skills (Vlach & DeBrock, 2017, 2019). In the present study, no individual data were collected in this regard, and future research should consider testing participants to clarify the role of individual cognitive and linguistic abilities in cross-situational learning.

Contrary to expectations, no significant difference in performance was found between the two tasks. Given the complex nature of the *categorisation task*, we expected the scores to be lower than the *recognition task*. Unlike the *recognition task*, in fact, the word-pseudosign correspondence was never direct, as the pseudosigns and words presented were never congruent with



**Figure 5.** Grand averaged ERPs of the adult group divided by task (*recognition task*, upper row; *categorisation task*, bottom row). (A) The N400 and LPC time windows are highlighted in red and blue, respectively (B) The shading of the difference waveform encompasses 95% confidence intervals. (C) Representation of time points and electrodes where significant differences between conditions were detected: channels (y-axis) vs. time (x-axis). Blue indicates negative effects, red indicates positive effects.



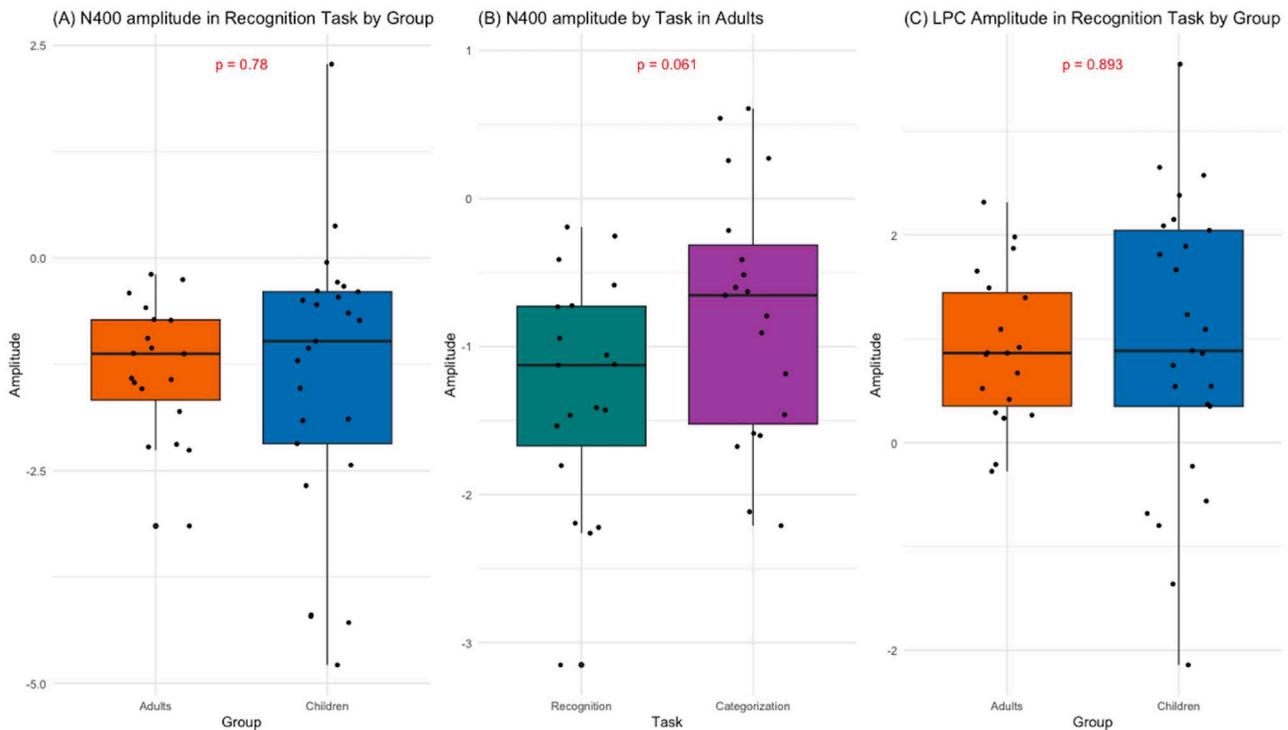
**Figure 6.** Grand averaged ERPs of the child group divided by task (*recognition task*, upper row; *categorisation task*, bottom row). No significant effect was found in the *categorisation task*. (A) The N400 and LPC time windows are highlighted in red and blue, respectively (B) The shading of the difference waveform encompasses 95% confidence intervals. (C) Representation of time points and electrodes where significant differences between conditions were detected: channels (y-axis) vs. time (x-axis). Blue indicates negative effects, red indicates positive effects.

the form with which the participants had been familiarised. In addition, there were multiple steps in the judgement between congruent and incongruent trials. Participants had first to access the correct semantic content of the pseudosigns based on the meaning of the familiar word paired with the pseudosign during the *familiarisation phase*. Subsequently, used that knowledge to decide whether the following word pertained to the same category. The fact that our hypothesis was disproved and participants performed without significant differences in the two tasks demonstrates the robustness of cross-situational learning mechanisms across different tasks, age groups, and language modalities.

### ***N400 as a marker of semantic integration***

Electrophysiological data generally supported the behavioural results. Following expectations, we observed that when pseudosigns were not followed by

the corresponding spoken word or category, the processing effort for incongruent pairs increased. This “congruency effect” was reflected in a centro-parietal N400 observed in the grand-averaged ERP waveform of congruent minus incongruent trials. Its presence suggests that participants successfully learned the pseudosigns, forming correct associations with both their corresponding spoken words and meaning. These findings are in line with previous research using semantic priming paradigms to investigate word (e.g. Angwin et al., 2014; Armstrong et al., 2024; Bermúdez-Margaretto et al., 2018; Borovsky et al., 2012; Hirotani et al., 2009; Kaczer et al., 2018; Mangardich & Sabbagh, 2022; Mestres-Missé et al., 2007; Perfetti et al., 2005) and gesture learning (e.g. Bernardis et al., 2008; Drijvers & Özyürek, 2018; Habets et al., 2011; Kelly et al., 2004; 2010; Özyürek et al., 2007; Sekine et al., 2020; Wu & Coulson, 2005; 2007a; 2007b; for review see Özyürek, 2014). However, differences in the effect emerged across tasks and



**Figure 7.** (A) N400 amplitude in the *recognition task* by group (adults vs. children), showing no significant difference between groups. (B) N400 amplitude by task in adults (recognition vs. categorisation), indicating no significant difference between tasks. (C) LPC amplitude in the *recognition task* by group (adults vs. children) showed no significant difference between groups.

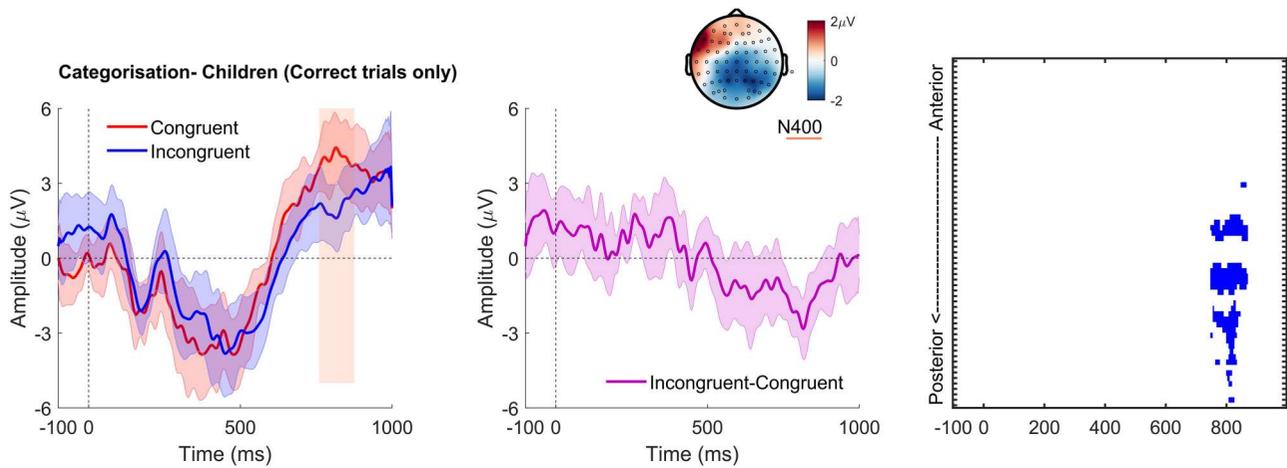
groups. In the *recognition task*, the N400 effect was reported in the usual 200–600 ms time window (Kutas & Federmeier, 2011, p. 623) for both groups, with a centro-parietal distribution across the two hemispheres. This greater retrieval effort was observed when a mismatched spoken word followed the pseudosigns prime, confirming that children and adults learned the pseudosigns. In the *categorisation task*, however, ERP findings did not follow a clear pattern. In adults, the N400 effect was also observed in the centro-parietal region, indicating that this group also successfully associated pseudosigns with the intended meaning. Conversely, in the child group, no significant effect was found during the main analysis. In addition, no difference in mean N400 amplitude between tasks was observed in the adult group. While this result contrasts with our initial hypothesis, it aligns with the behavioural findings, which showed no differences in learning between tasks. This suggests that both tasks were equally demanding for adults, resulting in similar behavioural and electrophysiological responses. A similar

comparison was not possible in the child group due to the absence of N400 in the main analysis.

This absence of an N400 effect in the child group was unexpected for three main reasons. First, despite previous literature indicating that semantic processing may differ between adults and children (e.g. Benau et al., 2011), N400 amplitude in the *recognition task* did not significantly differ between the two groups (as was our hypothesis). This suggests that, within our sample, children’s neurophysiological responses were in line with those of adults (see Coch, 2014; Sekine et al., 2020). Second, in adults, the mean N400 amplitude remained consistent across both tasks. Therefore, assuming that children in our study processed semantic information similarly to adults, we would expect a comparable N400 response across tasks. Lastly, children’s behavioural performance indicated learning in both *recognition* and *categorisation*, with *d*-prime scores showing comparable accuracy. Importantly, response bias analysis showed no significant difference between groups, indicating that both adults and children adopted a similar conservative response strategy. These considerations, underpinned by the presence of the N400 effect in the *recognition task*, lead us to expect a similar N400 in the *categorisation task*. To explore this discrepancy, we conducted an exploratory EEG analysis on correctly identified trials. Although not part of the

**Table 4.** Cluster stats results with effect sizes for correctly identified trials only.

Group	Task	Cluster type	Latency	<i>p</i>	Cohen’s <i>d</i>
Children ( <i>N</i> = 14)	categorisation	Negative	760–876	.043	–1.44



**Figure 8.** Grand averaged ERPs of children ( $n = 25$ ) in the *categorisation task*, with ERPs from correct trials only (additional analysis). The N400 time window is highlighted in red (A). The shading of the difference waveform (B) encompasses 95% confidence intervals. (C) Representation of clusters showing significant differences: channels (y-axis) vs. time (x-axis). Blue indicates negative effects.

initial plan, this approach aimed to minimise the influence of task-related difficulties and unrelated noise, providing a more focused measure of categorisation processing. Results revealed an N400 peak with centro-parietal distribution, showing that while giving the correct answers, children also exhibited evidence of semantic integration.

Several factors could have contributed to these results. First, the influence of the small sample size (adults  $N = 19$ , children  $N = 25$ ). Although statistical analysis of the behavioural data showed a significant effect of the group, such an effect did not emerge in the ERP analysis. Considering the small sample size and the greater individual variability in children compared to adults, this might have attenuated the magnitude of the N400 effect and compromised its detectability. Future research with a larger sample size is needed to determine whether the absence of a significant effect in the child ERP analysis reflects a true null result or is due to individual variability and limited statistical power. Additionally, including a broader age span would better inform on developmental differences that did not emerge in our study. Second, previous studies have shown a mismatch between behavioural and electrophysiological responses. At the individual level, ERPs are susceptible to artefact noise (e.g. eye blink, eye movement, fidgeting), which can vary among participants and consequently impacts the number of trials included in the analysis (Dalebout & Fox, 2000; Henderson et al., 2011; McLaughlin et al., 2004; Uwer & Von Suchodoletz, 2000). In contrast, behavioural data tend to be more consistent across participants. In our dataset, no correlation was observed between  $d$ -prime scores and ERP amplitudes when considering all trials. Thus, in line

with previous research, this could account for the mismatch between behavioural and ERP results in children. This hypothesis is further supported by the correlation between  $d$ -prime and N400 found in our exploratory analysis of correctly identified trials, indicating that when noise is reduced, the relationship between behavioural and ERP data becomes clearer. Another potential explanation is given by the complexity of the *categorisation task*. Participants had to process two within-task levels: first, they needed to extend the direct association learned between a pseudosign and a spoken word (e.g. pseudosign “dog” = spoken word “dog”) to a more abstract, indirect association at the category level (e.g. pseudosign “dog”/ANIMALS = spoken word “cat”/ANIMALS  $\rightarrow$  congruent); second, they had to do this task while switching between modalities, integrating information from both the visual (pseudosign) and auditory (spoken word) domains. Although, as discussed, no significant effect of task was found in the N400 amplitude between the *recognition* and *categorization task* in adults, we could assume that, for children, the multiple layers of processing in the *categorisation task* may have attenuated the N400 effect. In fact, in the additional analysis in which only correctly identified trials were considered (and thus, presumably, where the processing was more linear), a significant N400 was found. This result suggests that children who performed successfully on the task were able to construct a semantic representation of pseudosigns. Still, the lack of an effect in the entire dataset raises questions about the generalisability of the effect. It is possible that children’s performance was affected by uncertainty about the correctness of their answers, which may have added additional cognitive load during the task. To address

this concern, future studies should consider implementing the paradigm with an additional phase to provide participants with feedback on their learning (e.g. indicating which word-pseudosign items they correctly identified). This would help disentangle effects driven by semantic processing from those influenced by task complexity, ensuring that ERP responses reflect word-pseudosigns integration more directly. Furthermore, it is important to note that the restriction of the analysis further reduced the sample size ( $n = 14$ ). A larger sample will, therefore, be needed to determine the robustness of this effect.

Overall, the N400 effect observed in our study seems to be broadly distributed in the central-parietal region, with no hemispheric dominance. This pattern aligns with findings from ERP studies on gesture-speech integration in post-learning semantic tasks, where centro-parietal regions have been associated with the processing of online (e.g. Drijvers & Özyürek, 2018; Habets et al., 2011; Kelly et al., 2010) and sequential (e.g. Wu & Coulson, 2005; 2007b) gesture-speech integration. In children, the ERP study of Sekine et al. (2020) also reported a centro-parietal N400 effect for 6-7-year-olders, showing online integration of semantic information similar to adults. In the context of cross-situational learning of novel words, previous studies have reported a left-central N400 with a more frontal distribution in adults (Angwin et al., 2022) and a bilaterally distributed occipito-parietal effect in children (Mangardich & Sabbagh, 2022). Altogether, our neurophysiological findings are broadly consistent with previous research, yet further studies are needed to better understand how gestures learned through cross-situational learning are processed. This could offer insights into the influence of task structure, stimulus properties, and developmental differences on N400 distribution, as well as the potential role of individual differences in linguistic and cognitive abilities in modulating this processing.

### **LPC as a marker of explicit memory retrieval and integration**

Interestingly, in the *recognition task*, a late positive component (LPC) with a parietal (children) and centro-parietal (adults) distribution was also observed. Although the exploration of the LPC was not initially included in our hypothesis, its presence in the *recognition task* and absence in the *categorisation task* were noteworthy.

While some studies have indicated the LPC as an index of semantic processing in word recognition (Hoshino & Thierry, 2012; Martin et al., 2009), evidence

suggests that semantic activation alone is not sufficient to elicit an LPC effect. Contrary to the presence of the N400 – partially underpinned by automatic processes (Kutas & Federmeier, 2011) – the presence of the LPC seems modulated by explicit retrieval and integration demands (Bakker et al., 2015; Kaczer et al., 2018; Kandhadai & Federmeier, 2010; Rugg & Curran, 2007; Zachau et al., 2014). These findings suggest that the LPC tends to emerge when tasks require an explicit, controlled retrieval of a learned association. For example, Bakker and colleagues (2015) reported that recently learned words elicit an LPC response when participants are required to actively retrieve learned meanings, but this effect decreases as words become more integrated into the lexicon after 24-h consolidation.

In our study, an LPC effect was observed in the *recognition task* but not in the *categorisation task*. In light of previous literature, we attribute this discrepancy to differences in retrieval demands of the two tasks. In the *recognition task*, participants were presented with the same learned word-pseudosign items and were required to recall the direct association, which likely engaged explicit memory retrieval mechanisms. This aligns with previous findings indicating that retrieval effort and integration demands modulate LPC amplitude (e.g. Bakker et al., 2015). In contrast, in the *categorisation task*, participants were required to process a more abstract, indirect association, as the spoken word corresponded to a broader category rather than a directly learned association. As category-level associations were never explicitly trained, participants may not have engaged in explicit retrieval strategies and/or may have failed to recognise the pseudosigns as a prime for the learned association. This likely reduced the likelihood of controlled retrieval processes being triggered, leading to the absence of the LPC. This interpretation is consistent with the idea that LPC emergence depends on whether retrieval demands are strong enough to engage explicit processing mechanisms. Future studies should further investigate this aspect by either providing explicit training on category-level associations or introducing feedback to enhance retrieval confidence, allowing for an examination of potential differences in LPC modulation.

### **Iconic features in our stimuli**

At the suggestion of two anonymous reviewers, a survey was conducted after the completion of the experiment to assess whether the pseudosigns carried any prior cultural meaning and to evaluate their degree of iconicity. A full report of the validation study (methods, results, and discussion) can be found in Appendix B.

The survey, administered online via Qualtrics (Qualtrics, 2025), consisted of two tasks assessing *transparency* and *iconicity*, respectively. Following Sevcikova Sehyr and Emmorey (2019), *transparency* is defined as the ease of inferring a sign's meaning without prior knowledge, and *iconicity* as the extent to which a sign visually resembles its referent (e.g. through shape, movement, or other features). To assess transparency, participants were asked to assign a noun freely to each pseudosign. To assess iconicity, they were presented with the list of the eight experimental target words and asked to match each pseudosign to a word.

The results of the survey showed that transparency responses were highly variable across pseudosigns, with no consistent interpretations emerging. This confirmed that the pseudosigns lacked inherent transparency and, thus, were culturally meaningless to the participants. The iconicity assessment revealed that the majority of the pseudosigns (*bed, car, cold, dog, shirt*) did not show significant above-chance accuracy, but three pseudosigns (*cup, pink, and toe*) were matched with the spoken words we had coincidentally associated them within our experiment with above-chance accuracy. This result suggests that despite our attempt to create non-iconic stimuli, these pseudosigns contained some visual features that facilitated their association with the intended spoken words.

Since these results were unexpected, we analysed the data further to investigate whether the different levels of iconicity influenced learning performance in the main experiment. In line with prior studies, the learning accuracy for the three iconic pseudosigns was significantly higher than for the five non-iconic ones. This result was not surprising, as studies suggest that iconicity can influence gesture processing and acquisition (e.g. Caselli & Pyers, 2020; Hofweber et al., 2022, 2023; Lücke & Ritterfeld, 2014; Namy et al., 2004; Orlansky & Bonvillian, 1984; Ortega & Morgan, 2015). However, all pseudosigns were learned significantly above chance, suggesting that learning was not dependent on a pseudosign's iconicity.

We would like to note that the imbalance of our stimulus characteristics does not weaken our overall conclusion that children and adults can associate pseudosigns with word meanings in a cross-situational learning context. However, we recommend for future research to pre-validate the experimental stimuli as a common practice to control for unexpected iconic features. Importantly, this pre-validation could be done not only in gesture studies but in any type of word-learning experiments, for example, to control for coincidental relations between pseudowords and novel objects. In fact, iconicity has been suggested to be substantially present in spoken language word forms (for a discussion, see Vigliocco

et al., 2014). Therefore, assessing the iconicity of experimental stimuli could help disentangle the effects of learning from those driven by iconicity itself.

## Conclusions

In conclusion, our findings suggest that sign-like gestures can be effectively acquired and associated with intended meanings by adults and children, even under the referential ambiguity inherent in the cross-situational learning context and after limited exposure. The lexical and semantic association was evidenced by high scores in the behavioural tasks and the presence of an N400 effect in the ERPs. This suggests that sign-like gestures are likely to be learned through statistical learning mechanisms. These findings support our hypothesis that these gestures constitute a highly salient linguistic input and are perceived as meaningful forms of communication.

Given that research on the processing of sign-like gestures, particularly in the context of cross-situational learning, remains a relatively unexplored field, we believe our findings provide a valuable foundation for further research. Despite its limitations, our study offers insights into the behavioural and neurophysiological aspects of multimodal language acquisition and the potential language-specific characteristics of gestural communication.

## Acknowledgements

The authors would like to thank Madison Brown, Xinyu Guo, Ponsuang Luengtaweekul, and Kriengkrai Puripanyo at Macquarie University for their help in data collection and Marisa Varrica for her help in recording the stimuli, as well as two anonymous reviewers for their constructive comments on a previous version of the manuscript.

## Author contributions (CRediT author statement)

**Arianna Colombani:** Conceptualisation, Methodology, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualisation, Project administration, Funding acquisition. **Varghese Peter:** Conceptualisation, Methodology, Formal analysis, Writing – Original Draft, Writing – Review & Editing, Visualisation, Supervision. **Qian Yin Mai:** Investigation. **Amanda Saksida:** Writing – Review & Editing, Supervision. **Natalie Boll-Avetisyan:** Writing – Review & Editing, Supervision. **Outi Tuomainen:** Writing – Review & Editing, Supervision, Funding acquisition. **Mridula Sharma:** Conceptualisation, Methodology, Formal analysis, Writing – Review & Editing, Supervision, Funding acquisition.

## Data availability statement

Stimuli, data frames, and analysis scripts are publicly available on this paper's project page on the OSF platform at <https://doi.org/10.17605/OSF.IO/YU94Q>

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by an IDEALAB scholarship from the University of Potsdam and Macquarie University (CTIMQRES IDEALAB, [grant number 20223934]). Costs for participant reimbursement were covered by Macquarie University under the Beijing Higher Talent Foundation [grant number 5282500] to MS. Funding sources had no involvement in the study design, collection, analysis and interpretation of data, in the writing of the report, or in the decision to submit the article for publication.

## ORCID

Arianna Colombani  <https://orcid.org/0000-0003-1199-8286>  
 Varghese Peter  <http://orcid.org/0000-0002-4007-507X>  
 Amanda Saksida  <http://orcid.org/0000-0001-7434-9625>  
 Natalie Boll-Avetisyan  <https://orcid.org/0000-0001-5446-946X>  
 Outi Tuomainen  <https://orcid.org/0000-0002-8654-2446>  
 Mridula Sharma  <https://orcid.org/0000-0002-0448-6429>

## References

- Acredolo, L., & Goodwyn, S. (1988). Symbolic gesturing in normal infants. *Child Development*, 59(2), 450. <https://doi.org/10.2307/1130324>
- Acredolo, L., & Goodwyn, S. (1996). *Baby signs: how to talk with your baby before your baby can talk* (Chicago III, Ed.). Contemporary Books.
- Angwin, A. J., Armstrong, S. R., Fisher, C., & Escudero, P. (2022). Acquisition of novel word meaning via cross situational word learning: An event-related potential study. *Brain and Language*, 229(March), 105111. <https://doi.org/10.1016/j.bandl.2022.105111>
- Angwin, A. J., Phua, B., & Copland, D. A. (2014). Using semantics to enhance new word learning: An ERP investigation. *Neuropsychologia*, 59(1), 169–178. <https://doi.org/10.1016/j.neuropsychologia.2014.05.002>
- Apfelbaum, K. S., & McMurray, B. (2017). Learning during processing: Word learning doesn't wait for word recognition to finish. *Cognitive Science*, 41(S4), 706–747. <https://doi.org/10.1111/cogs.12401>
- Aravind, A., De Villiers, J., Pace, A., Valentine, H., Golinkoff, R., Hirsh-Pasek, K., Iglesias, A., & Sweig Wilson, M. (2018). Fast mapping word meanings across trials: Young children forget all but their first guess. *Cognition*, 177, 177–188. <https://doi.org/10.1016/j.cognition.2018.04.008>
- Armstrong, S. R., Copland, D. A., Escudero, P., & Angwin, A. J. (2024). Tracking the acquisition and retention of novel word representations: An ERP study. *Language, Cognition and Neuroscience*, 39(4), 401–422. <https://doi.org/10.1080/23273798.2024.2310549>
- Baker, S. A., Golinkoff, R. M., & Petitto, L.-A. (2006). New insights into Old puzzles from infants' categorical discrimination of soundless phonetic units. *Language Learning and Development*, 2(3), 147–162. [https://doi.org/10.1207/s15473341l1d0203\\_1](https://doi.org/10.1207/s15473341l1d0203_1)
- Bakker, I., Takashima, A., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Tracking lexical consolidation with ERPs: Lexical and semantic-priming effects on N400 and LPC responses to newly-learned words. *Neuropsychologia*, 79, 33–41. <https://doi.org/10.1016/j.neuropsychologia.2015.10.020>
- Bates, E. (1979). The emergence of symbols. In E. A. Hammel (Ed.), *The emergence of symbols*. Elsevier. <https://doi.org/10.1016/c2013-0-10341-8>
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. Academic Press.
- Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly*.
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), <https://doi.org/10.18637/jss.v067.i01>
- Batista, A., Le Normand, M.-T., & Colletta, J.-M. (2019). Chapitre 3. Rôle et évolution des combinaisons bimodales au cours de l'acquisition du langage. Données chez l'enfant francophone âgé de 18 à 42 mois. In A. Mazur-Palandre & I. Colón de Carvajal (Eds.), *Multimodalité du langage dans les interactions et l'acquisition*. UGA Éditions. <https://doi.org/10.4000/books.ugaeditions.10947>
- Benau, E. M., Morris, J., & Couperus, J. W. (2011). Semantic processing in children and adults: Incongruity and the N400. *Journal of Psycholinguistic Research*, 40(3), 225–239. <https://doi.org/10.1007/s10936-011-9167-1>
- Benitez, V. L., Yurovsky, D., & Smith, L. B. (2016). Competition between multiple words for a referent in cross-situational word learning. *Journal of Memory and Language*, 90, 31–48. <https://doi.org/10.1016/j.jml.2016.03.004>
- Benitez, V. L., Zettersten, M., & Wojcik, E. (2020). The temporal structure of naming events differentially affects children's and adults' cross-situational word learning. *Journal of Experimental Child Psychology*, 200, 104961. <https://doi.org/10.1016/j.jecp.2020.104961>
- Berens, S. C., Horst, J. S., & Bird, C. M. (2018). Cross-Situational learning Is supported by propose-but-verify hypothesis testing. *Current Biology*, 28(7), 1132–1136.e5. <https://doi.org/10.1016/j.cub.2018.02.042>
- Bermúdez-Margaretto, B., Beltrán, D., Cuetos, F., & Domínguez, A. (2018). Brain signatures of new (pseudo-) words: Visual repetition in associative and non-associative contexts. *Frontiers in Human Neuroscience*, 12(September), 1–12. <https://doi.org/10.3389/fnhum.2018.00354>
- Bernardis, P., Salillas, E., & Caramelli, N. (2008). Behavioural and neuropsychological evidence of semantic interaction between iconic gestures and words. *Cognitive Neuropsychology*, 25, 1114–1128.
- Blythe, R. A., Smith, K., & Smith, A. D. M. (2010). Learning times for large lexicons through cross-situational learning. *Cognitive Science*, 34(4), 620–642. <https://doi.org/10.1111/j.1551-6709.2009.01089.x>

- Booth, J. R., Burman, D. D., Meyer, J. R., Lei, Z., Trommer, B. L., Davenport, N. D., Li, W., Parrish, T. B., Gitelman, D. R., & Mesulam, M. M. (2003). Neural development of selective attention and response inhibition. *NeuroImage*, 20(2), 737–751. [https://doi.org/10.1016/S1053-8119\(03\)00404-X](https://doi.org/10.1016/S1053-8119(03)00404-X)
- Borovsky, A., Elman, J. L., & Kutas, M. (2012). Once is enough: N400 indexes semantic integration of novel word meanings from a single exposure in context. *Language Learning and Development*, 8(3), 278–302. <https://doi.org/10.1080/15475441.2011.614893>
- Branchini, C., & Mantovan, L. (2020). A grammar of Italian sign language (LIS). In *A grammar of Italian sign language (LIS)*. Fondazione Università Ca' Foscari. <https://doi.org/10.30687/978-88-6969-474-5>.
- Bunce, J. P., & Scott, R. M. (2017). Finding meaning in a noisy world: Exploring the effects of referential ambiguity and competition on 2-5-year-olds' cross-situational word learning. *Journal of Child Language*, 44(3), 650–676. <https://doi.org/10.1017/S0305000916000180>
- Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language*, 23(3), 645–673. <https://doi.org/10.1017/S0305000900008989>
- Capone, N. C., & McGregor, K. K. (2004). Gesture development: A review for clinical and research practices. *Journal of Speech, Language, and Hearing Research*, 47, 173–186.
- Capone, N. C., & McGregor, K. K. (2005). The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, and Hearing Research*, 48(6), 1468–1480. [https://doi.org/10.1044/1092-4388\(2005\)102](https://doi.org/10.1044/1092-4388(2005)102)
- Caselli, N. K., & Pyers, J. E. (2020). Degree and not type of iconicity affects sign language vocabulary acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(1), 127–139.
- Chen, C., Gershkoff-Stowe, L., Wu, C., Cheung, H., & Yu, C. (2017). Tracking multiple statistics: Simultaneous learning of object names and categories in English and mandarin speakers. *Cognitive Science*, 41(6), 1485–1509. <https://doi.org/10.1111/cogs.12417>
- Christiansen, M. H. (2019). Implicit statistical learning: A tale of Two literatures. *Topics in Cognitive Science*, 11(3), 468–481. <https://doi.org/10.1111/tops.12332>
- Coch, D. (2014). The N400 and the fourth grade shift. *Developmental Science*, 18(2), 254–269. <https://doi.org/10.1111/desc.12212>
- Colletta, J.-M., Guidetti, M., Capirci, O., Cristilli, C., Demir, O., Kunene-Nicholas, R., & Levine, S. (2015). Effects of age and language on co-speech gesture production: An investigation of French, American, and Italian children's narratives. *Journal of Child Language*, 42(1), 122–145. <https://doi.org/10.1017/S0305000913000585>
- Colombani, A., Saksida, A., Pavani, F., & Orzan, E. (2022). Symbolic and deictic gestures as a tool to promote parent-child communication in the context of hearing loss: a systematic review. *International Journal of Pediatric Otorhinolaryngology*, 165(December 2022), 111421. <https://doi.org/10.1016/j.ijporl.2022.111421>
- Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neuro-cognitive mechanisms of statistical learning. *Neuroscience & Biobehavioral Reviews*, 112, 279–299. <https://doi.org/10.1016/j.neubiorev.2020.01.032>
- Cowan, N. (2016). Working memory maturation: Can We Get at the essence of cognitive growth? *Perspectives on Psychological Science*, 11(2), 239–264. <https://doi.org/10.1177/1745691615621279>
- Crespo, K., & Kaushanskaya, M. (2021). Is 10 better than 1? The effect of speaker variability on children's cross-situational word learning. *Language Learning and Development*, 17(4), 397–410. <https://doi.org/10.1080/15475441.2021.1906680>
- Crespo, K., Vlach, H., & Kaushanskaya, M. (2023). The effects of bilingualism on children's cross-situational word learning under different variability conditions. *Journal of Experimental Child Psychology*, 229, 105621. <https://doi.org/10.1016/j.jecp.2022.105621>
- Dal Ben, R., Souza, D. D. H., & Hay, J. F. (2022). Combining statistics: The role of phonotactics on cross-situational word learning. *Psicologia: Reflexão e Crítica*, 35(1), 30. <https://doi.org/10.1186/s41155-022-00234-y>
- Dalebout, S., & Fox, L. (2000). Identification of the mismatch negativity in the responses of individual listeners. *Journal of the American Academy of Audiology*, 11, 12–22. <https://doi.org/10.1055/s-0042-1748004>
- Dautriche, I., & Chemla, E. (2014). Cross-situational word learning in the right situations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(3), 892–903. <https://doi.org/10.1037/a0035657>
- Drijvers, L., & Özyürek, A. (2018). Native language status of the listener modulates the neural integration of speech and iconic gestures in clear and adverse listening conditions. *Brain and Language*, 177–178, 7–17. <https://doi.org/10.1016/j.bandl.2018.01.003>
- Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., Polich, J., Reinvang, I., & Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, 120(11), 1883–1908. <https://doi.org/10.1016/j.clinph.2009.07.045>
- Escudero, P., Mulak, K. E., Fu, C. S. L., & Singh, L. (2016c). More limitations to monolingualism: Bilinguals outperform monolinguals in implicit word learning. *Frontiers in Psychology*, 7(August), 1–13. <https://doi.org/10.3389/fpsyg.2016.01218>
- Escudero, P., Mulak, K. E., & Vlach, H. A. (2016a). Infants encode phonetic detail during cross-situational word learning. *Frontiers in Psychology*, 7, <https://doi.org/10.3389/fpsyg.2016.01419>
- Escudero, P., Mulak, K. E., & Vlach, H. A. (2016b). Cross-Situational learning of minimal word pairs. *Cognitive Science*, 40(2), 455–465. <https://doi.org/10.1111/cogs.12243>
- Escudero, P., Smit, E. A., & Angwin, A. J. (2022). Investigating orthographic versus auditory cross-situational word learning With online and laboratory-based testing. *Language Learning*, 73(2), 543–577. <https://doi.org/10.1111/lang.12550>
- Fasolo, M., & D'Odorico, L. (2012). Gesture-plus-word combinations, transitional forms, and language development. *Gesture*, 12(1), 1–15. <https://doi.org/10.1075/gest.12.1.01fas>
- Filippi, P., Gingras, B., & Fitch, W. T. (2014). Pitch enhancement facilitates word learning across visual contexts. *Frontiers in Psychology*, 5, <https://doi.org/10.3389/fpsyg.2014.01468>
- Fitneva, S. A., & Christiansen, M. H. (2011). Looking in the wrong direction correlates With more accurate word learning. *Cognitive Science*, 35(2), 367–380. <https://doi.org/10.1111/j.1551-6709.2010.01156.x>

- Fitneva, S. A., & Christiansen, M. H. (2017). Developmental changes in cross-situational word learning: The inverse effect of initial accuracy. *Cognitive Science*, 41(51), 141–161. <https://doi.org/10.1111/cogs.12322>
- Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, 145(12), 1128–1153. <https://doi.org/10.1037/bul0000210>
- Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, 20(3), 347–353. <https://doi.org/10.1111/j.1467-9280.2009.02295.x>
- Goldin-Meadow, S. (1999). What children contribute to language-learning. *Science Progress*, 82(1), 89–102. <https://doi.org/10.1177/003685049908200105>
- Goldin-Meadow, S., & Butcher, C. (2003). Pointing toward two-word speech in young children. In N. E. A. Mahwah, & S. Kita (Eds.), *Pointing: Where language, culture, and cognition meet* (pp. 85–107). Lawrence Erlbaum Associates Publishers.
- Goodwyn, S. W., Acredolo, L., & Brown, C. A. (2000). Impact of symbolic gesturing on early language development. *Journal of Nonverbal Behavior*, 24(2), 81–103. <https://doi.org/10.1023/A:1006653828895>
- Habets, B., Kita, S., Shao, Z., Özyurek, A., & Hagoort, P. (2011). The role of synchrony and ambiguity in speech – gesture integration during comprehension. *Journal of Cognitive Neuroscience*, 23(8), 1845–1854. <https://doi.org/10.1162/jocn.2010.21462>
- Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, 44, 9–15. <https://doi.org/10.1016/j.lindif.2015.10.003>
- Harrigan, S., & Nikolopoulos, T. P. (2002). Parent interaction course in order to enhance communication skills between parents and children following pediatric cochlear implantation. *International Journal of Pediatric Otorhinolaryngology*, 66(2), 161–166. [https://doi.org/10.1016/S0165-5876\(02\)00243-4](https://doi.org/10.1016/S0165-5876(02)00243-4)
- Henderson, L. M., Baseler, H. A., Clarke, P. J., Watson, S., & Snowling, M. J. (2011). The N400 effect in children: Relationships with comprehension, vocabulary and decoding. *Brain and Language*, 117(2), 88–99. <https://doi.org/10.1016/j.bandl.2010.12.003>
- Hendrickson, A. T., & Perfors, A. (2019). Cross-situational learning in a zipfian environment. *Cognition*, 189, 11–22. <https://doi.org/10.1016/j.cognition.2019.03.005>
- Hirotsu, M., Stets, M., Striano, T., & Friederici, A. D. (2009). Joint attention helps infants learn new words: Event-related potential evidence. *NeuroReport*, 20(6), 600–605. <https://doi.org/10.1097/WNR.0b013e32832a0a7c>
- Hofweber, J., Aumonier, L., Janke, V., Gullberg, M., & Marshall, C. (2023). Which aspects of visual motivation aid the implicit learning of signs at first exposure?. *Language Learning*, 1–31. <https://doi.org/10.1111/lang.12587>
- Hofweber, J. E., Aumonier, L., Janke, V., Gullberg, M., & Marshall, C. (2022). Breaking into language in a New modality: The role of input and individual differences in recognising signs. *Frontiers in Psychology*, 13(May), <https://doi.org/10.3389/fpsyg.2022.895880>
- Holler, J., & Levinson, S. C. (2019). Multimodal language processing in human communication. *Trends in Cognitive Sciences*, 23(8), 639–652. <https://doi.org/10.1016/j.tics.2019.05.006>
- Hoshino, N., & Thierry, G. (2012). Do Spanish-English bilinguals have their fingers in two pies – or is it their toes? An electrophysiological investigation of semantic access in bilinguals. *Frontiers in Psychology*, 3(FEB), 1–6. <https://doi.org/10.3389/fpsyg.2012.00009>
- Hoyer, R. S., Elshafei, H., Hemmerlin, J., Bouet, R., & Bidet-Caulet, A. (2021). Why Are children So distractible? Development of attention and motor control from childhood to adulthood. *Child Development*, 92(4), <https://doi.org/10.1111/cdev.13561>
- Hu, C.-F. (2017). Resolving referential ambiguity across ambiguous situations in young foreign language learners. *Applied Psycholinguistics*, 38(3), 633–656. <https://doi.org/10.1017/S0142716416000357>
- Hübscher, I., & Prieto, P. (2019). Gestural and prosodic development Act as sister systems and jointly pave the Way for children’s sociopragmatic development. *Frontiers in Psychology*, 10, 1259. <https://doi.org/10.3389/fpsyg.2019.01259>
- Isbilen, E. S., & Christiansen, M. H. (2022). Statistical learning of language: A meta-analysis into 25 years of research. *Cognitive Science*, 46(9), <https://doi.org/10.1111/cogs.13198>
- Junge, C., Boumeester, M., Mills, D. L., Paul, M., & Cosper, S. H. (2021). Development of the N400 for word learning in the first 2 years of life: A systematic review. *Frontiers in Psychology*, 12(June), <https://doi.org/10.3389/fpsyg.2021.689534>
- Kachergis, G., Yu, C., & Shiffrin, R. M. (2013). Actively learning object names across ambiguous situations. *Topics in Cognitive Science*, 5(1), 200–213. <https://doi.org/10.1111/tops.12008>
- Kaczer, L., Bavassi, L., Petroni, A., Fernández, R. S., Laurino, J., Degiorgi, S., Hochman, E., Forcato, C., & Pedreira, M. E. (2018). Contrasting dynamics of memory consolidation for novel word forms and meanings revealed by behavioral and neurophysiological markers. *Neuropsychologia*, 117, 472–482. <https://doi.org/10.1016/j.neuropsychologia.2018.07.001>
- Kandhadai, P., & Federmeier, K. D. (2010). Automatic and controlled aspects of lexical associative processing in the two cerebral hemispheres. *Psychophysiology*, <https://doi.org/10.1111/j.1469-8986.2009.00969.x>
- Kelly, S. D., Creigh, P., & Bartolotti, J. (2010). Integrating speech and iconic gestures in a stroop-like task: Evidence for automatic processing. *Journal of Cognitive Neuroscience*, 22(12), 2939. <https://doi.org/10.1162/jocn.2009.21254>
- Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and Language*, 89(1), 253–260. [https://doi.org/10.1016/S0093-934X\(03\)00335-3](https://doi.org/10.1016/S0093-934X(03)00335-3)
- Kreher, D. A., Holcomb, P. J., & Kuperberg, G. R. (2006). An electrophysiological investigation of indirect semantic priming. *Psychophysiology*, 43(6), 550–563. <https://doi.org/10.1111/j.1469-8986.2006.00460.x>
- Krentz, U. C., & Corina, D. P. (2008). Preference for language in early infancy: The human language bias is not speech specific. *Developmental Science*, 11(1), 1–9. <https://doi.org/10.1111/j.1467-7687.2007.00652.x>
- Kutas, M., & Federmeier, K. D. (2009). N400. *Scholarpedia*, 4(10), 7790.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*,

- 62, 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. In *Science* (Vol. 207, Issue 4427, pp. 203–205). American Assn for the Advancement of Science. <https://doi.org/10.1126/science.7350657>.
- Kutas, M., & Petten, C. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). Academic Press.
- Lieberman, A. M., Fitch, A., & Borovsky, A. (2022). Flexible fast-mapping: Deaf children dynamically allocate visual attention to learn novel words in American sign language. *Developmental Science*, 25(3), e13166. <https://doi.org/10.1111/desc.13166>
- Lillo-Martin, D., & Henner, J. (2021). Acquisition of sign languages. *Annual Review of Linguistics*, 7(1), 395–419. <https://doi.org/10.1146/annurev-linguistics-043020-092357>
- Long, H. L., Bowman, D. D., Yoo, H., Burkhardt-Reed, M. M., Bene, E. R., & Oller, D. K. (2020). Social and endogenous infant vocalizations. *PLoS One*, 15(8), e0224956. <https://doi.org/10.1371/journal.pone.0224956>
- Luck, S. J. (2004). *An introduction to the event – related potential technique* (2nd ed.). MIT Press.
- Lüke, C., & Ritterfeld, U. (2014). The influence of iconic and arbitrary gestures on novel word learning in children with and without SLI. *Gesture*, 14(2), 204–225. <https://doi.org/10.1075/gest.14.2.04luk>
- MacDonald, K., Yurovsky, D., & Frank, M. C. (2017). Social cues modulate the representations underlying cross-situational learning. *Cognitive Psychology*, 94, 67–84. <https://doi.org/10.1016/j.cogpsych.2017.02.003>
- Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory*. Psychology Press. <https://doi.org/10.4324/9781410611147>.
- Mangardich, H., & Sabbagh, M. A. (2022). Event-related potential studies of cross-situational word learning in four-year-old children. *Journal of Experimental Child Psychology*, 222, 105468. <https://doi.org/10.1016/j.jecp.2022.105468>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG – and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Martin, C. D., Dering, B., Thomas, E. M., & Thierry, G. (2009). Brain potentials reveal semantic priming in both the ‘active’ and the ‘non-attended’ language of early bilinguals. *NeuroImage*, 47(1), 326–333. <https://doi.org/10.1016/j.neuroimage.2009.04.025>
- The MathWorks Inc. (2023). *MATLAB version: 23.2.0 (R2023b)*. The MathWorks Inc. <https://www.mathworks.com>
- McGregor, K. K., Rohlfing, K. J., Bean, A., & Marschner, E. (2009). Gesture as a support for word learning: The case of under. *Journal of Child Language*, 36(4), 807–828. <https://doi.org/10.1017/S0305000908009173>
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nat. Neurosci*, 7(7), 703–704.
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, 92(3), 350–371. <https://doi.org/10.1037/0033-295X.92.3.350>
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press, Ed.
- McNeill, D. (2005). Gesture and thought. In *Gesture and thought*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226514642.001.0001>.
- Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, 17(8), 1858–1866. <https://doi.org/10.1093/cercor/bhl094>
- Monaghan, P., Mattock, K., Davies, R. A. I., & Smith, A. C. (2015). Gavagai Is as gavagai does: Learning nouns and verbs from cross-situational statistics. *Cognitive Science*, 39(5), 1099–1112. <https://doi.org/10.1111/cogs.12186>
- Monaghan, P., Schoetensack, C., & Rebuschat, P. (2019). A single paradigm for implicit and statistical learning. *Topics in Cognitive Science*, 11(3), 536–554. <https://doi.org/10.1111/tops.12439>
- Mulak, K. E., Vlach, H. A., & Escudero, P. (2019). Cross-Situational learning of phonologically overlapping words across degrees of ambiguity. *Cognitive Science*, 43(5), 1–19. <https://doi.org/10.1111/cogs.12731>
- Murillo, E., & Casla, M. (2021). Multimodal representational gestures in the transition to multi-word productions. *Infancy*, 26(1), 104–122. <https://doi.org/10.1111/infa.12375>
- Namy, L. L., Campbell, A. L., & Tomasello, M. (2004). The changing role of iconicity in Non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development*, 5(1), 37–57. [https://doi.org/10.1207/s15327647jcd0501\\_3](https://doi.org/10.1207/s15327647jcd0501_3)
- Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants’ interpretations of different forms of symbolic reference. *Child Development*, 69(2), 295–308. <https://doi.org/10.1111/j.1467-8624.1998.tb06189.x>
- Newman, A. J., Supalla, T., Fernandez, N., Newport, E. L., & Bavelier, D. (2015). Neural systems supporting linguistic structure, linguistic experience, and symbolic communication in sign language and gesture. *Proceedings of the National Academy of Sciences of the United States of America*, 112(37), 11684–11689. <https://doi.org/10.1073/pnas.1510527112>
- Oller, D. K., Ramsay, G., Bene, E., Long, H. L., & Griebel, U. (2021). Protophones, the precursors to speech, dominate the human infant vocal landscape. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1836), <https://doi.org/10.1098/rstb.2020.0255>
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). Fieldtrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, <https://doi.org/10.1155/2011/156869>
- Orlansky, M. D., & Bonvillian, J. D. (1984). The role of iconicity in early sign language acquisition. *Journal of Speech and Hearing Disorders*, 49(3), 287–292. <https://doi.org/10.1044/jshd.4903.287>
- Ortega, G., & Morgan, G. (2015). The effect of iconicity in the mental lexicon of hearing non-signers and proficient signers: Evidence of cross-modal priming. *Language, Cognition and Neuroscience*, 30(5), 574–585. <https://doi.org/10.1080/23273798.2014.959533>
- Özçalışkan, Ş., & Dimitrova, N. (2013). How gesture input provides a helping hand to language development. *Seminars in Speech and Language*, 34(4), 227–236. <https://doi.org/10.1055/s-0033-1353447>
- Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical*

- Transactions of the Royal Society B: Biological Sciences*, 369(1651), <https://doi.org/10.1098/rstb.2013.0296>
- 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>
- Paul, I. M., Hohman, E. E., Birch, L. L., Shelly, A., Vallotton, C. D., & Savage, J. S. (2019). Exploring infant signing to enhance responsive parenting: Findings from the INSIGHT study. *Maternal and Child Nutrition*, 15(3), <https://doi.org/10.1111/mcn.12800>
- Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(6), 1281–1292. <https://doi.org/10.1037/0278-7393.31.6.1281>
- Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, 10(5), 233–238. <https://doi.org/10.1016/j.tics.2006.03.006>
- Pizer, G., Walters, K., & Meier, P. (2007). Bringing up infant with hearing signs: Language ideologies and socialisations in hearing families. *Sign Language Studies*, 7(4), 143–159.
- Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, 152, 9–19. <https://doi.org/10.1016/j.cognition.2016.03.001>
- Posnansky, C. J. (1978). Category norms for verbal items in 25 categories for children in Grades 2–6. *Behavior Research Methods and Instrumentation*, 10(6), 819–832. <https://doi.org/10.3758/BF03205407>
- Puccini, D., & Liszkowski, U. (2012). 15-month-old infants fast map words but not representational gestures of multimodal labels. *Frontiers in Psychology*, 3(APR), <https://doi.org/10.3389/fpsyg.2012.00101>
- Qualtrics. (2025, January). Qualtrics survey software. <https://www.qualtrics.com>.
- Rabagliati, H., Senghas, A., Johnson, S., & Marcus, G. F. (2012). Infant rule learning: Advantage language, or advantage speech? *PLoS One*, 7(7), e40517. <https://doi.org/10.1371/journal.pone.0040517>
- Ramos-Escobar, N., Laine, M., Sanseverino-Dillenburg, M., Cucurell, D., François, C., & Rodriguez-Fornells, A. (2021). The interplay between domain-general and domain-specific mechanisms during the time-course of verbal associative learning: An event-related potential study. *NeuroImage*, 242, 118443. <https://doi.org/10.1016/j.neuroimage.2021.118443>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-Project.Org>.
- Rebuschat, P., Monaghan, P., & Schoetensack, C. (2021). Learning vocabulary and grammar from cross-situational statistics. *Cognition*, 206, 104475. <https://doi.org/10.1016/j.cognition.2020.104475>
- Roberts, M. Y., & Hampton, L. H. (2018). Exploring cascading effects of multimodal communication skills in infants with hearing loss. *Journal of Deaf Studies and Deaf Education*, 23(1), 95–105. <https://doi.org/10.1093/deafed/enx041>
- Roembke, T. C., & McMurray, B. (2016). Observational word learning: Beyond propose-but-verify and associative bean counting. *Journal of Memory and Language*, 87, 105–127. <https://doi.org/10.1016/j.jml.2015.09.005>
- Roembke, T. C., & McMurray, B. (2021). Multiple components of statistical word learning are resource dependent: Evidence from a dual-task learning paradigm. *Memory & Cognition*, 49(5), 984–997. <https://doi.org/10.3758/s13421-021-01141-w>
- Roembke, T. C., Simonetti, M. E., Koch, I., & Philipp, A. M. (2023). What have we learned from 15 years of research on cross-situational word learning? A focused review. *Frontiers in Psychology*, 14(July), 1–9. <https://doi.org/10.3389/fpsyg.2023.1175272>
- Roembke, T. C., Wiggs, K. K., & McMurray, B. (2018). Symbolic flexibility during unsupervised word learning in children and adults. *Journal of Experimental Child Psychology*, 175, 17–36. <https://doi.org/10.1016/j.jecp.2018.05.016>
- Rowe, M. L., Özçalışkan, Ş., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture’s role in predicting vocabulary development. *First Language*, 28(2), 182–199. <https://doi.org/10.1177/0142723707088310>
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11(6), 251–257. <https://doi.org/10.1016/j.tics.2007.04.004>
- Saksida, A., Rebesco, R., Colombani, A., Pintonello, S., Tonon, E., Santoro, A. M., & Orzan, E. (2023). The timeline of non-vocal and vocal communicative skills in infants with hearing loss. *Frontiers in Pediatrics*, 11(January), <https://doi.org/10.3389/fped.2023.1209754>
- Scott, R. M., & Fisher, C. (2012). 2.5-Year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, 122(2), 163–180. <https://doi.org/10.1016/j.cognition.2011.10.010>
- Sekine, K., Schoechl, C., Mulder, K., Holler, J., Kelly, S., Furman, R., & Özyürek, A. (2020). Evidence for children’s online integration of simultaneous information from speech and iconic gestures: An ERP study. *Language, Cognition and Neuroscience*, 35(10), 1283–1294. <https://doi.org/10.1080/23273798.2020.1737719>
- Sevcikova Sehyr, Z., & 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. <https://doi.org/10.1017/langcog.2019.18>
- Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568. <https://doi.org/10.1016/j.cognition.2007.06.010>
- Smith, L. B., & Yu, C. (2013). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Language Learning and Development*, 9(1), 25–49. <https://doi.org/10.1080/15475441.2012.707104>
- Šoškić, A., Jovanović, V., Styles, S. J., Kappenman, E. S., & Ković, V. (2022). How to do better N400 studies: Reproducibility, consistency and adherence to research standards in the existing literature. *Neuropsychology Review*, 32(3), 577–600. <https://doi.org/10.1007/s11065-021-09513-4>
- Stenberg, G., Lindgren, M., Johansson, M., Olsson, A., & Rosén, I. (2000). Semantic processing without conscious identification: Evidence from event-related potentials. *Journal of*

- Experimental Psychology: Learning, Memory, and Cognition*, 26(4), 973–1004. <https://doi.org/10.1037/0278-7393.26.4.973>
- Stone, A., Petitto, L.-A., & Bosworth, R. (2018). Visual sonority modulates infants' attraction to sign language. *Language Learning and Development*, 14(2), 130–148. <https://doi.org/10.1080/15475441.2017.1404468>
- Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology*, 126, 395–411. <https://doi.org/10.1016/j.jecp.2014.06.003>
- Suanda, S. H., & Namy, L. L. (2012). Detailed behavioral analysis as a window into cross-situational word learning. *Cognitive Science*, 36(3), 545–559. <https://doi.org/10.1111/j.1551-6709.2011.01218.x>
- Suanda, S. H., Walton, K. M., Broesch, T., Kolkin, L., & Namy, L. L. (2013). Why Two-year-olds fail to learn gestures as object labels: Evidence from looking time and forced-choice measures. *Language Learning and Development*, 9(1), 50–65. <https://doi.org/10.1080/15475441.2012.723189>
- Suttora, C., Guarini, A., Zuccarini, M., Aceti, A., Corvaglia, L., & Sansavini, A. (2022). Integrating gestures and words to communicate in full-term and Low-risk preterm late talkers. *International Journal of Environmental Research and Public Health*, 19(7), <https://doi.org/10.3390/ijerph19073918>
- Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, 66(1), 126–156. <https://doi.org/10.1016/j.cogpsych.2012.10.001>
- Tuninetti, A., Mulak, K. E., & Escudero, P. (2020). Cross-Situational word learning in Two foreign languages: Effects of native language and perceptual difficulty. *Frontiers in Communication*, 5, 602471. <https://doi.org/10.3389/fcomm.2020.602471>
- Uwer, R., & Von Suchodoletz, W. (2000). Stability of mismatch negativities in children. *Clinical Neurophysiology*, 111(1), 45–52. [https://doi.org/10.1016/S1388-2457\(99\)00204-7](https://doi.org/10.1016/S1388-2457(99)00204-7)
- Vallotton, C. D. (2012). Infant signs as intervention? Promoting symbolic gestures for preverbal children in low-income families supports responsive parent-child relationships. *Early Childhood Research Quarterly*, 27(3), 401–415. <https://doi.org/10.1016/j.ecresq.2012.01.003>
- Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the battig and montague (1969) norms. *Journal of Memory and Language*, 50(3), 289–335. <https://doi.org/10.1016/j.jml.2003.10.003>
- Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130292. <https://doi.org/10.1098/rstb.2013.0292>
- Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, 93, 217–230. <https://doi.org/10.1016/j.jml.2016.10.001>
- Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, 45(4), 700–711. <https://doi.org/10.1037/xlm0000611>
- Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, 127(3), 375–382. <https://doi.org/10.1016/j.cognition.2013.02.015>
- Vlach, H. A., & Sandhofer, C. M. (2014). Retrieval dynamics and retention in cross-situational statistical word learning. *Cognitive Science*, 38(4), 757–774. <https://doi.org/10.1111/cogs.12092>
- Vogt, P. (2012). Exploring the robustness of cross-situational learning under zipfian distributions. *Cognitive Science*, 36(4), 726–739. <https://doi.org/10.1111/j.1551-6709.2011.1226.x>
- Walker, N., Monaghan, P., Schoetensack, C., & Rebuschat, P. (2020). Distinctions in the acquisition of vocabulary and grammar: An individual differences approach. *Language Learning*, 70(52), 221–254. <https://doi.org/10.1111/lang.12395>
- Wang, F. H. (2020). Explicit and implicit memory representations in cross-situational word learning. *Cognition*, 205, 104444. <https://doi.org/10.1016/j.cognition.2020.104444>
- Wang, F. H., & Trueswell, J. C. (2019). Spotting dalmatians: Children's ability to discover subordinate-level word meanings cross-situationally. *Cognitive Psychology*, 114, 101226. <https://doi.org/10.1016/j.cogpsych.2019.101226>
- Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*, 13(6), 258–263. <https://doi.org/10.1016/j.tics.2009.03.006>
- Willems, R. M., Özyürek, A., & Hagoort, P. (2007). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex*, 17(10), 2322–2333. <https://doi.org/10.1093/cercor/bhl141>
- Wu, Y. C., & Coulson, S. (2005). Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology*, 42(6), 654–667. <https://doi.org/10.1111/j.1469-8986.2005.00356.x>
- Wu, Y. C., & Coulson, S. (2007a). How iconic gestures enhance communication: An ERP study. *Brain and Language*, 101(3), 234–245. <https://doi.org/10.1016/j.bandl.2006.12.003>
- Wu, Y. C., & Coulson, S. (2007b). Iconic gestures prime related concepts: An ERP study. *Psychonomic Bulletin & Review*, 14(1), 57–63. <https://doi.org/10.3758/BF03194028>
- Yip, M. C. W. (2022). Cross-situational word learning of cantonese Chinese. *Psychonomic Bulletin & Review*, 30(3), 1074–1080. <https://doi.org/10.3758/s13423-022-02217-7>
- Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics: Research article. *Psychological Science*, 18(5), 414–420. <https://doi.org/10.1111/j.1467-9280.2007.01915.x>
- Yu, C., & Smith, L. B. (2011). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science*, 14(2), 165–180. <https://doi.org/10.1111/j.1467-7687.2010.00958.x>
- Yu, C., Zhong, Y., & Fricker, D. (2012). Selective attention in cross-situational statistical learning: Evidence from Eye tracking. *Frontiers in Psychology*, 3, <https://doi.org/10.3389/fpsyg.2012.00148>
- Yurovsky, D., Fricker, D. C., Yu, C., & Smith, L. B. (2014). The role of partial knowledge in statistical word learning. *Psychonomic Bulletin & Review*, 21(1), 1–22. <https://doi.org/10.3758/s13423-013-0443-y>

Yurovsky, D., Yu, C., & Smith, L. B. (2013). Competitive processes in cross-situational word learning. *Cognitive Science*, 37(5), 891–921. <https://doi.org/10.1111/cogs.12035>

Zachau, S., Korpilahti, P., Hämäläinen, J. A., Ervast, L., Heinänen, K., Suominen, K., Lehtihalmes, M., &

Leppänen, P. H. T. (2014). Electrophysiological correlates of cross-linguistic semantic integration in hearing signers: N400 and LPC. *Neuropsychologia*, 59, 57–73. <https://doi.org/10.1016/j.neuropsychologia.2014.04.011>