Brain & Language 277 (2026) 105736

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Metaphors we learn to sign by: Positive transfer effects of existing sign language expertise and gestural repertoire on metaphor inferencing in a novel sign language☆

Julia Hofweber[a,*], Lizzie Aumônier[a], Theresa Latham[b], Katherine Trice[c], Zhenghan Qi[b,c]

[a] Department of Psychology, Northeastern University London, London, UK
[b] Department of Communication Sciences and Disorders, Northeastern University, Boston, MA, United States
[c] Department of Psychology, Northeastern University, Boston, MA, United States

ARTICLE INFO

Keywords:
Language learning
Transfer
Cognition
Sign languages
Bilingualism
Metaphors
Iconicity
Gesture
Implicit learning

ABSTRACT

Our study investigated metaphor inferencing in second languages (L2). According to universalist accounts, L2 metaphor comprehension should be intuitive. However, cross-linguistic differences suggest that successful metaphor inferencing is subject to transfer effects. Little is known about the interaction of transfer effects involving modality differences in the context of sign languages. We investigated whether existing sign language expertise enhances SPACE-TIME metaphor inferencing in signed L2s. Crucially, we focused on SPACE-TIME metaphor inferencing because the need to abstract from salient concrete SPACE interpretations may pose challenges to sign novices.

Methodologically, we compared native signers (N = 27, L1 = American or British Sign Language) to hearing sign novices (N = 37, L1 = English). Participants were exposed to a novel sign language through a short weather forecast, mirroring real-life implicit immersion. Subsequently, they completed tasks tapping into meaning assignment. Our analyses focused on metaphor inferencing, considering transfer effects from cognateness regarding L1-L2 sign similarity and gesture-sign similarity.

Our findings revealed a marked dissociation in transfer effects by group. Amongst native signers metaphoricity facilitated L2 meaning assignment. However, metaphoricity posed a challenge to sign novices who defaulted to concrete SPACE interpretations, misled by their saliency and iconicity. At the item-level, cognate effects occurred both from existing linguistic knowledge and gestural repertoires. Our study suggests that whilst all participants recognised universal iconic patterns facilitating concrete meaning assignment, abstraction and metaphor inferencing relied on positive transfer from specific cognates or generic knowledge of modality-specific metaphor formation strategies. Hence, teaching materials should build on existing knowledge in a bespoke manner.

1. Introduction

Whilst cross-linguistic transfer effects in spoken languages have been widely studied, few studies have explored their interaction with modality in the context of sign languages (Pichler & Koulidobrova, 2023). Likewise, research on metaphor inferencing in second languages is scarce (Athanasopoulos et al., 2017; Athansopoulos & Bylund, 2020). This paper investigates transfer effects of modality-specific knowledge on meaning assignment in second sign languages, notably on metaphor inferencing. Specifically, we compared native signers to sign novices regarding SPACE-TIME metaphor comprehension in a novel sign language.

Our title pays homage to Lakoff & Johnson's seminal 1980 book "metaphors we live by", which outlines the ubiquity of metaphors in conceptual and linguistic processing in languages across the world. Languages systematically exploit conceptual metaphors by associating abstract notions with iconically grounded concrete physical referents. A common example are so-called SPACE-TIME metaphors, in which abstract notions of TIME are expressed by referring to the concrete bodily experience of SPACE. Spoken-language examples are expressions such as "short days" or "extended deadlines".

The prevalence of metaphors across languages is in line with

☆ This article is part of a special issue entitled: 'Gest Cogn Lang' published in Brain and Language.

* Corresponding author.
E-mail address: j.hofweber@northeastern.edu (J. Hofweber).

https://doi.org/10.1016/j.bandl.2026.105736
Received 30 July 2025; Received in revised form 25 February 2026; Accepted 28 February 2026
Available online 24 April 2026

0093-934X/© 2026 Northeastern University. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Embodied Cognition theories, which suggest that abstract thought and language are rooted in and evolved from physical reality. According to this position, conceptual representations are mediated by sensorimotor experiences. However, it is unclear to which extent the final step of abstraction is arbitrary and language-specific. Hence, hybrid explanations involving both embodied and embodiment-independent cognitive processes may be more suitable explanatory frameworks (Dove, 2016). SPACE-TIME Metaphors are hybrid in that they involve both embodiment (SPACE-TIME conceptual overlap), and abstraction (abstract TIME inference).

The debated relationship between abstract thought and real-world referents is overtly observable in sign languages. Compared to spoken languages, sign languages display relatively higher levels of iconicity, i.e. similarity between linguistic forms and real-world referent (Ortega, 2017). Moreover, embodied and iconic aspects of sign languages are partly due the overt visibility of the articulatory organs, i.e. the hands and body. Sign languages exploit this visibility when expressing abstract temporal notions, resulting in rich SPACE-TIME metaphor repertoires using spatial cues to mark temporal notions (Marschark, 2004; Meir & Cohen, 2018).

British Sign Language BSL, for instance, systematically marks "timelines" using body parts. The passing of weeks, for instance, is marked on the arms. The term "one week ago" is expressed by running one finger of the dominant hand backwards along the non-dominant arm. Likewise, its future-designating equivalent "in one week's time" is denoted by moving one finger forwards along the arm. SPACE-TIME metaphors in sign languages constitute instances of iconicity as they directly depict TIME moving in SPACE, based on the folk notion that timelines are linear (Taub, 2001). In a similar vein, co-speech gestures have been shown to represent temporal notions by making reference to SPACE (Heng, 2016).

The evolution, development and universality of metaphors is a topic of debate. Four different positions prevail (Cassanto, 2013). The first position argues that metaphors are a cross-culturally universal component of embodied cognition, cf. explanation above, arising from humans' shared environment (Lakoff & Johnson, 2008). The second view also takes a universalist stance but suggests that universality is caused by our innate cognitive architecture, rather than shared environments. Certain metaphors have been hardwired into our conceptual repertoires through evolution (Pinker, 1997). The evolutionary explanation suggests that hard-wired conceptual metaphors emerge from an interaction of genes and environment, providing a conceptual link between the two universalist positions.

The third position emphasizes that metaphors are language-specific and arbitrary (Gentner, 2001). Indeed, studies on cross-linguistic differences and second language learning reveal that any universality of metaphoric relationships is mediated by arbitrary language-specific conventions (Gentner, 2001; Gu, & Swerts, 2017). Importantly, metaphors differ in the extent to which they are universal versus language-specific and arbitrary (Kövecses, 2005). SPACE-TIME metaphors are partially shaped by language-specific internal, rather than external factors (Máthé, 2022). To give an example from the domain of sign languages, the direction and orientation of timelines in both sign languages and gestures is modulated by cultural and linguistic conventions, suhc as writing systems. Although exposed to the same objective physics of time, European culture and writing systems express timelines horizontally, whilst Mandarin Chinese culture also uses vertical timelines. This difference is reflected in how SPACE-TIME metaphors are expressed in both sign and gesture (Gu & Swerts, 2019).

The fourth position takes an interactionist approach suggesting that metaphorical thinking and the development of metaphors arises from a complex interplay of all three factors, i.e. embodied cognition, genetic hardwiring, and cross-linguistic differences (Cassanto, 2013; Athanasopoulos et al, 2017). The hybrid approach is supported by neuroimaging findings revealing partial overlap but also distinctiveness of the neural correlates of spatial and temporal processing (e.g., Lai & Desai, 2016). Investigations of L2 metaphor processing could contribute to our understanding of the mutual influence of the factors affecting conceptual metaphorical processing. The degree of cross-linguistic intelligibility of metaphors may provide insights into the relative contribution of universal versus arbitrary factors to metaphor processing. However, L2 metaphor processing remains under-researched (Athansopoulos & Bylund, 2024).

Theories suggesting that metaphoricity is universal, based on either innateness or embodied cognition, would suggest that the processing of metaphors is independent of language-specific conventions. Hence, they would predict that metaphors comprehension in second languages should be effortless. Indeed, the universality of metaphors may facilitate their comprehension in L2s, compared to non-metaphorical items. Theories that emphasize the arbitrary aspects of metaphoricity imply a need to acquire novel conceptual mappings for L2 comprehension. Hence, metaphoricity in L2 may pose a challenge to comprehension. The interactionist perspective would take a nuanced approach investigating the relative degree of contribution of several factors influencing the role of metaphoricity in L2 comprehension (Howe, 2024). Existing research corroborates the latter position. Whilst L2 metaphor comprehension generally challenges learners, positive transfer through cognate effects and cross-linguistic overlaps in metaphor expression may facilitate learning (Chen & Lai, 2014).

A range of factors modulate L2 metaphor comprehension. Crucially, it may be affected by whether the novel language is in a familiar or novel modality. In the context of sign L2, it is therefore crucial to distinguish between two scenarios: (1) M1L2: native signers who are already familiar with visuo-spatial communication, e.g. native signers acquiring a new sign language, (2) M2L2: sign novices acquiring a novel language in an unfamiliar modality (Pichler & Koulidobrova, 2023). Few studies have investigated the comprehension of L2 metaphors, let alone in L2 sign. A study on Swedish sign language comparing M1L2 versus M2L2 interpreters of sign found that the expression of SPACE-TIME metaphors was less idiomatic in non-native than native signer interpreters (Nilson, 2018). These findings suggest that metaphoricity poses a challenge for M2L2 signers, even when they are highly proficient. Moreover, the Nilson (2018) study suggests that sign experts may develop a sensitivity towards detecting and using metaphors in the visual modality, possibly due to existing knowledge of modality-specific metaphor formation strategies.

Another important factor when considering metaphoricity in sign languages is iconicity. Iconicity is associated with transparency and embodied intuitiveness of meaning (Sehyr & Emmorey, 2019). Hence, it should facilitate L2 metaphor comprehension. Indeed, iconicity and transparency have been shown to influence sign learning (Ortega, 2017; Hofweber et al., 2022). Interestingly, signedSPACE-TIME metaphors commonly encapsulate iconicity, creating an iconically mediated link between the concrete real-word referent, the iconic SPACE meaning, and the abstract TIME notion. This involves a two-step process of meaning derivation, as depicted in Fig. 1. In the first step, iconic signs literally depict or embody the visual properties of the real-world referent SPACE. The second step involves abstraction to the metaphorically represented concept of TIME (Taub, 2001).

To give an example, in Swedish sign language, the sign for TODAY consists of a downward palm movement depicting the spatial notion of HERE. To derive the metaphorical temporal meaning TODAY, further abstraction is required. To give another example, the BSL sign for SHORT iconically depicts shortness of SPACE and TIME by showing two pairs of fingers close together. Applying this observation to L2 contexts, it could be argued that the first step of deriving iconic meaning should be universally intuitive, whilst the second step of abstraction would be subject to language- and modality-specific metaphor formation strategies. This mediating role of iconicity is more prevalent in sign than spoken languages. In spoken English, the spatial term "short" is also used metaphorically to describe temporal duration, but the word form "short" itself is fully arbitrary bearing no direct iconic link to the real-word

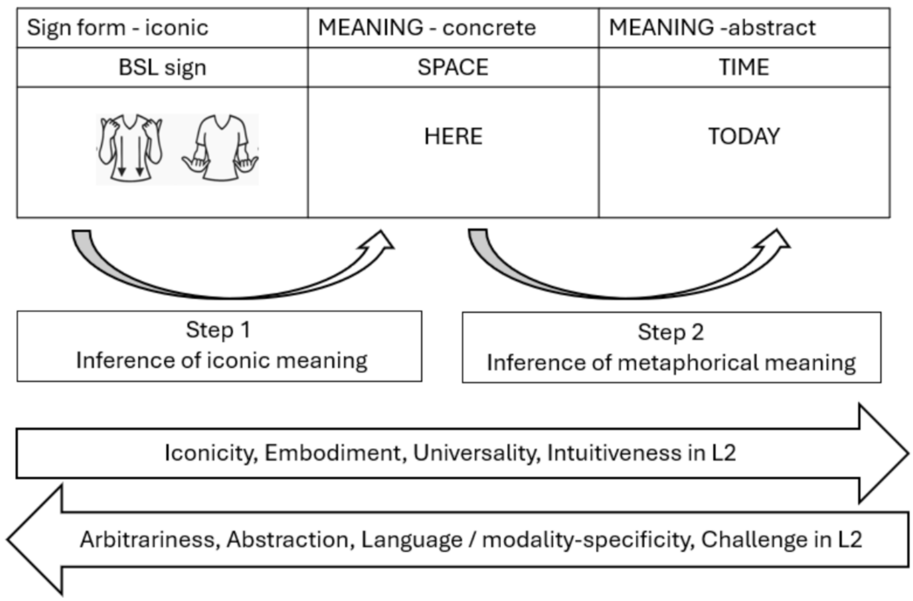

Fig. 1. Metaphor inferencing in a novel sign language.

referent shortness. The embodied aspect of spoken language SPACE-TIME metaphors arises mainly from the conceptual similarity between SPACE and TIME.

In the context of L2 sign, the ubiquity of iconicity in sign languages may facilitate the first step, i.e. assigning concrete spatial meaning iconically represented in the sign form. However, the additional step of inferencing temporal metaphorical meaning may in fact be hindered by the visual saliency of the iconic meaning. A study on implicit learning of Swedish sign language amongst M2L2 sign-novices found that M2L2 learners struggled to detect SPACE-TIME metaphors (Hofweber et al., 2023). Participants identified relevant visuo-spatial properties of signs assigning concrete SPACE meanings but then failed to extrapolate to the metaphorical TIME meaning. For the TODAY sign, for instance, M2L2 participants intuitively associated the downward palm movement with concrete SPACE interpretations, such as HERE or DROP, but failed to abstract to the TIME meaning of TODAY. It appears that the visuo-spatial nature of sign languages facilitates the detection of SPACE iconicity, which may in turn discourage further search processes for abstract TIME interpretations. Thus, for M2L2 leaners, the detection of SPACE-TIME metaphors may be particularly challenging due to the saliency and iconicity of the concrete spatial interpretation. To derive the accurate temporal meaning, learners need to "overwrite" the salient spatial interpretation. Hence, there is a risk of "premature closure", i.e. defaulting to the concrete SPACE interpretation.

The Hofweber et al. (2023) study is limited in that it only assessed M2L2 learners. To investigate the effects of cross-modal and cross-linguistic differences on L2 metaphor detection, we need to consider M1L2 learners with existing knowledge about strategies to express metaphoricity in the visuo-spatial modality. M1L2 learners may transfer their awareness of how SPACE-TIME metaphors are expressed in sign languages onto their interpretation of metaphors in novel sign languages.

Assuming cross-linguistic differences in metaphorical mappings, a crucial factor in L2 comprehension is transfer of existing linguistic knowledge to the second language. Transfer may arise at the macro-level regarding global strategies within a language or modality, or it may be limited to item-specific cognate effects. The Typological Primacy Model (Rothman, 2011) posits that learners subconsciously recognise the proximity of whole language systems, which regulates whether and how existing linguistic knowledge is transferred onto novel linguistic input. Other models suggest that transfer operates at the micro-level for each individual linguistic item (Chen & Myers, 2017; Lee, 2022). An investigation of transfer effects in L2 sign provides interesting insights into whether transfer effects of cross-modal or modality-specific metaphor formation strategies are macro-systemic, or whether they operate at the micro-level of each individual metaphor based on cognate effects. Differences in the modality in which languages are expressed could be considered as indicative of high levels of typological distance. This raises the question whether macro-level transfer effects are subject to existing knowledge of langauges expressed visuo-spatially.

Two types of macro-level transfer could arise. Firstly, the case of sign language is particularly informative because, arguably, postitive or negative transfer arising from cross-modal similarities or differences in metaphor fomation could be considered as a form of high-level or macro-level transfer. Sign and spoken languages share cross-modally that TIME is expressed through SPACE-TIME metaphors. Cross-modal transfer of the awareness that SPACE may designate TIME could thus result in positive transfer, especially given that sign and spoken language processing share overlapping neural correlates (Trettenbein et al., 2021). Secondly, the transfer of modality-specific metaphor formation strategies could be interpreted as another form of macro-level transfer. Due to the ubiquity of iconically-rooted SPACE-TIME metaphors, M1L2 native signers may develop a generic awareness of the fact that in sign languages concrete iconic SPACE meanings commonly need to be overwritten in favour of TIME interpretations. This may result in modality-bespoke macro-level transfer effects in metaphor formation strategies. If existing linguistic expertise is a pre-requisite for this modality-specific positive transfer effects, then M1L2 learners may be more successful at detecting L2 metaphors than M2L2 learners, who have been shown to be led astray by concrete spatial meanings (Hofweber et al., 2023).

Moreover, transfer may operate at the micro-level of individual cognate signs, i.e. the degree of similarity between each L2 sign and its L1 translation equivalent. The facilitative effect of cognates on L2 processing has been widely attested for spoken languages (Midgley et al., 2011; Peeters et al., 2013). Regarding micro-level cognate effects for metaphors, Chen & Lai (2014) describe positive transfer from conceptual metaphors that overlap in the first and second language. Such positive transfer from L1 representations could be predicted for M1L2 learners.

M2L2 sign novices may not benefit from positive transfer effects of existing sign language knowledge. Instead, they may be assisted by their gestural repertoire, which has been shown to lead to analogous cognate effects (Ortega et al., 2017; Karadöller et al., 2024). Given the intuitiveness of SPACE-TIME metaphors, they may share similarities with co-speech gestures depicting time. This is in line with empirical findings revealing neural overlap in gesture and sign processing (Willems et al., 2007). However, other studies emphasize that sign language processing amongst proficient signers draws upon more language-specific brain regions than gesture processing (Newman et al., 2015). Although signs and gestures share the same locus and manner of articulation, they have different neural correlates. Whilst the processing of signs takes place in language-specialised left-lateralized parts of the brain, gestures elicit less language-specific neural responses (Newman et al., 2015). Moreover, gesture-sign overlap did not predict semantic learning amongst M2L2 signers in a study by Hofweber et al. (2023), although metaphors were not specifically investigated. Inconsistencies in findings thus call for further investigations of cognate effects from gestural repertoires.

Given the majority of sign language users are deaf or hard-of-hearing DHH, any potentially observed transfer effects could be due to neural plasticity adaptations resulting from either sign language expertise or Deafness (Que et al., 2018; Simon et al., 2020). Early onset sign expertise has been shown to modulate and enhance left-lateralised neural structures associated with language processing (McCullough & Emmorey, 2021), and white matter integrity (Simon et al., 2020). At the same time, cross-modal usage of auditory cortex resources may result in enhanced visuo-spatial processing amongst the DHH (Ding et al., 2015). This may in turn facilitate the interpretation of input in a novel visual language. It therefore is important to understand whether transfer effects in the context of metaphor inferencing in a novel sign language are due to language-specific transfer of sign language expertise, or enhanced domain-general visuo-spatial cognition resulting from Deafness.

Finally, metaphor inferencing in L2 sign may depend on whether receptive or productive skills are involved. The study on implicit learning of sign amongst M2L2 learners (Hofweber et al., 2023) administered both a transparency task assessing comprehension and meaning assignment, and a gesture elicitation task in which participants were asked to manually depict and produce a range of meanings (Hofweber et al., 2023). In the gesture elicitation task, participants produced SPACE-TIME metaphoricity in a sign-idiomatic manner. However, in the meaning assignment task, sign novices failed to detect SPACE-TIME metaphors, displaying a preference for the visually salient concrete meaning., i.e. SPACE over TIME. For instance, when asked to produce an improvised gesture representing TODAY, participants drew upon SPACE-TIME metaphoricity, e.g. indicating present location, but when guessing the meaning, they would guess HERE rather than TODAY. This suggests that in the novel visuo-spatial modality, L2 metaphor inferencing may be more challenging during comprehension than during production. Hence, our study focused on comprehension.

This study.

The current study explores metaphor inferencing in a novel sign language, and how it interacts with existing knowledge of sign languages. We explored whether metaphoricity poses a challenge or facilitates meaning assignment in the second language. Thus, our study contributes to debates about the cross-linguistic and cross-modal universality of metaphor comprehension. We focused on the comprehension of SPACE-TIME metaphors because they are ubiquitous in sign languages.

Crucially, we, differentiated between M1L2 and M2L2 contexts. For this purpose, we compared metaphor inferencing in M2L2 sign novices to M1L2 signers with existing sign language expertise. Metaphoricity has been shown to pose a challenge to M2L2 novices (Hofweber et al., 2023), which questions the notion of universality. However, within the same modality, metaphor formation strategies may apply universally cross-linguistically and cross-modal transfer. Hence, we investigated whether existing knowledge and awareness of such modality-specific metaphor formation strategies may facilitate L2 metaphor comprehension in M1L2 learners.

Thus, we explored the effects of existing sign language expertise on learners' ability to detect and comprehend metaphorical meaning in a novel sign language. For that purpose, we zoomed in further on transfer effects in the context of metaphor comprehension in novel sign languages. We explored the extent to which transfer applied at the macro–level, as expressed through group differences, or whether it operated at the micro–level, as expressed through cognate effects. At the item-level, we investigate whether transfer effects were limited to linguistic transfer from existing sign language knowledge, or whether cognate effects from non-signers' gestural repertoires also affected metaphor comprehension.

Methodologically, we focused on an implicit first-exposure context without explicit instructions, mirroring real-life immersion and exposure to novel languages (Hofweber et al., 2022). All participants were exposed to a novel sign language, i.e. Swedish Sign Language, through a 4-minute weather forecast containing 22 target signs, based on an existing paradigm (Gullberg et al., 2012; Hofweber et al., 2022). Crucially, we manipulated a common input-related predictor of L2 comprehension, i.e. frequency:11 target items occurred with a low frequency of 3–4 times, whilst the other 11 target items occurred with a high frequency of 8 times. The rationale of the input phase was that the weather forecast allowed participants to be exposed to target signs in context, which would enhance metaphor inferencing. For instance, if the concrete SPATIAL interpretation of metaphorical signs mismatched the weather forecast context, then this would trigger additional semantic search processes, potentially facilitating metaphor inferencing. Following the input exposure phase, participants completed a meaning assignment task assessing comprehension.

To investigate the effects of existing sign language expertise on SPACE-TIME metaphor inferencing, we compared native M1L2 signers ($N = 27$) with L1 Americal Sign Language (ASL) or British Sign Language (BSL) to hearing M2L2 sign novices ($N = 37$) with L1 spoken English. At the item level, we explored cognate effects (Chen & Lai, 2014). In M1L2 native signers, formal similarity was operationalized as formal overlap between L2 target signs with their L1 sign translation equivalents. In hearing non-signers, we assessed the formal overlap between target signs and meaning-equivalent gestures intuitively produced in a gesture elicitation task (cf. Hofweber et al., 2023).

To tease apart whether causality should be attributed to linguistic transfer from sign language expertise or Deafness, the M1L2 group consisted of both hearing and DHH native signers. If the DHH M1L2 signers outperform both hearing groups, this would suggest that any transfer effects arise from Deafness. If, on the other hand, both hearing and DHH M1L2 native signers display similar performance patterns, and outperform the M2L2 non-signers, this would be taken as evidence that any transfer effects are due to sign language expertise. Given that the effects we investigated were linguistic in nature, we predicted the latter pattern, i.e. that causality would be attributable to sign language expertise.

Moreover, the steps involved in metaphor inferencing (Fig. 1) may differ in universal intuitiveness. Whilst the initial recognition of the iconic SPACE meaning may be intuitive and universal, ultimate metaphor inferencing may be subject to modality-specific expertise. In this study, we therefore carefully considered whether iconicity mediated metaphor inferencing. A variety of measures can be used to assess the facilitative effect of visual similarity between a linguistic form and its semantic referent (Motamedi et al. 2019). So-called "iconicity tasks" present participants with the linguistic form and its meaning and ask for a rating of similarity. A more naturalistic measure is the "transparency task", in which participants are presented with a linguistic form, and asked to assign meaning to it. This mirrors real-life L2 comprehension settings, in which learners are not explicitly presented with meanings, but need to engage in semantic interpretation based solely on contextual cues. Existing research reveals both correlations between iconicity ratings and transparency scores, but also dissociations between transparency and iconicity scores in that some signs rated as highly iconic are not necessarily fully transparent (Sehr & Emmorey; Hofweber et el., 2023).

Indeed, a recent study on implicit semantic learning of signs amongst M2L2 learners found both iconicity ratings and transparency scores to predict meaning assignment to novel signs, but the effect sizes suggested that transparency scores were stronger predictors (Hofweber et al., 2023). Importantly for this study, the Hofweber et al. (2023) study observed that iconicity and transparency diverged particularly for meaning assignment in the case of signs involving additional levels of inferencing over and above direct iconicity, notably metaphoricity (Hofweber et al., 2023). Metaphorical signs tended to have low transparency but high iconicity scores. In the meaning assignment task, iconicity facilitated the derivation of the iconic and universally intuitive concrete SPACE interpretation. However, the visual saliency of this iconicity hindered further metaphorical inferencing. Hence, the M2L2 learners defaulted to the concrete spatial interpretation. Transparency scores were thus better predictors of metaphor inferencing than iconicity scores. Given this study focused on metaphoricity, we used transparency scores to measure sign-referent similarity.

To summarize, our study was guided by the following research questions (RQs):

RQ1: How does existing sign language expertise influence SPACE-TIME metaphor inferencing in a novel sign language, and how does this interact with potential predictors, notably exposure frequency and iconicity? If effects of sign language expertise arise, should causality be attributed to sign expertise or neural plasticity effects arising from Deafness?

RQ1 Prediction for effects of sign language expertise on metaphor inferencing.

If metaphor inferencing is entirely universal across modalities and languages, then it should not be affected by existing sign language expertise, so the M1L2 and M2L2 groups should perform on a par. Likewise, if cross-modal transfer of global awareness of SPACE-TIME metapors is sufficient for comprehension, then the M1L2 group should perform the same as the M2L2. If metaphor inferencing involves modality-specific expertise, then existing familiarity with modality-specific strategies of expressing SPACE-TIME metaphors would facilitate metaphor inferencing amongst M1L2 signers. Any effects of sign expertise on meaning assignment will be more salient for metaphorical than non-metaphorical items.

RQ1 Prediction for effects of metaphoricity in M2L2 signers.

If the awareness of SPACE-TIME metaphors is sufficient for metaphor inferencing cross-modally, then metaphoricity should have no effect or a facilitative effect on meaning assignment. If metaphor inferencing relies on modality-specific knowledge, then meaning assignment will be more challenging for metaphorical than non-metaphorical items due to M2L2 signers' unfamiliarity with modality-specific metaphor formation strategies. They will show a trend of "premature closure", i.e. defaulting to attributing the concrete SPACE meaning. Based on existing studies (Hofweber et al., 2023), we predict the latter scenario of "premature closure".

RQ1 Prediction for effects of metaphoricity in M1L2 signers.

Metaphoricity will facilitate meaning assignment. Firstly, iconicity will facilitate the derivation of the iconic and universally intuitive concrete SPACE interpretation. However, when the concrete interpretation does not match the context of the weather forecast, M1L2 learners will search for alternative interpretations. Awareness of modality-specific metaphor formation strategies will facilitate metaphor inferencing. In addition, metaphoricity, once successfully detected, will help anchor novel signs in existing knowledge.

RQ1 Predictions relating to causality attribution

If sign expertise affects metaphor inferencing, this raises the question whether such effects are due to sign language expertise or neural plasticity associated with Deafness. If the DHH M1L2 signers outperform both hearing groups (M1 and M2), this would suggest that any transfer effects arise from Deafness. If, on the other hand, both hearing and DHH M1L2 native signers display similar performance patterns, and outperform the M2L2 non-signers, this would be taken as evidence that any transfer effects are due to sign language expertise. Given that we investigate linguistic transfer, we predict causality effects to be attributable to sign language expertise.

RQ2: To which extent is metaphor inferencing in a novel sign language influenced by micro-level cognate effects?

RQ2 Prediction for M2L2 signers.

Within the non-signer M2L2 group, metaphor inferencing in a novel sign language may be facilitated by cognate effects, i.e. the similarity of each L2 target sign to participants' equivalent intuitive gestural repertoire, as assessed in the gesture elicitation task.

RQ2 Prediction for M1L2 signers.

In native M1L2 signers, metaphor inferencing in a novel sign language may be facilitated by cognate effects, i.e. the formal similarity between L1-L2 translation equivalents.

Aside from its theoretical implications, our study will result in impactful recommendations for educational practices by identifying the factors affecting metaphor detection and comprehension in second sign languages. A teaching intervention study (Beréndi et al., 2008) showed that explicit instructions raising learners' general metalinguistic awareness of the principles underlying metaphor construction enhances the learning of novel metaphors in spoken languages. If we find evidence of similar macro-level transfer effects amongst M1L2 learners, this would, for instance, suggest that M2L2 learners may benefit from explicit training about sign-idiomatic strategies of metaphor formation.

2. Methods

All experiments were conducted online using Gorilla experimental software by a team of researchers based at Northeastern University Boston and London campuses.

2.1. Participants

Our participants consisted of three sub-groups differing in sign language expertise and hearing abilities. Firstly, we tested 37 M2L2 learners, i.e. hearing sign novices with no prior knowledge of any sign languages. The first language of the hearing non-signers was kept constant as English (American English: N = 19, British English: N = 18). In addition, we tested 24 M1L2 native signers of sign languages as first language(s). For the majority of M1L2 signers, the native language was American Sign Language ASL (N = 21), with the exception of some native signers of British Sign Language BSL as their first language (N = 3). Crucially, within the (ASL) native signer group, 14 individuals were DHH, whilst 10 were hearing bimodal bilinguals proficient in both sign and spoken languages from birth. This resulted in three groups: (1) hearing M2L2: non-signers, (2) hearing M1L2: native signer bimodal bilinguals, (3) DHH M1L2: native DHH signers. None of the participants reported to have potentially confounding expertise in the target language Swedish Sign Language, or its spoken equivalent Swedish.

All participants were assessed for demographic background regarding AGE, GENDER and EDUCATION (Table 1). Whilst there was no age difference between hearing and DHH M1L2 native signers ($p >$

Table 1
Descriptive statistics for background variables by group.

Variable	Group	Mean	Standard Deviation	N
Age	M2L2 Non-signers	22.94	4.50	37
Years	M1L2 Signers hearing	28.00	6.00	10
	M1L2 Signers DHH	31.00	6.90	14
IQ KBIT	M2L2 Non-signers	21.92	3.43	37
Scores	M1L2 Signers hearing	21.8	2.94	10
	M1L2 Signers DHH	19.14	5.5	14
English Vocabulary	M2L2 Non-signers	0.76	0.13	37
Scores	M1L2 Signers hearing	0.73	0.07	10
	M1L2 Signers DHH	0.62	0.18	14

0.05), the M2L2 group was significantly younger than both M1L2 groups [$F(1,63) = 23.84$, $\eta 2 = 0.28$, $p < 0.001$]. Importantly, AGE correlated positively with our main dependent variable, i.e. performance in the meaning assignment task [$R(1,64) = 0.39$, $p < 0.001$]. To be prudent, we therefore included AGE as a covariate in the statistical analyses. Gender-wise, the majority of participants was female (M1L2 learners: 76% female; M1L2 learners: 78% female). It should be noted that gender neither had a significant effect on the dependent variable, i.e. semantic similarity between response and target meaning as measured in so-called Wu-Palmer scores [$F(1,63) = 1.13$, $\eta 2 = 0.04$, $p = 0.33$], nor interacted with group effects on Wu-Palmer scores [$F(1,63) = 0.37.84$, $\eta 2 = 0.02$, $p = 0.78$]. In terms of education, all participants had completed secondary school and obtained A-levels. However, the M1L2 group displayed a slightly more highly educated profile with 67% of participants holding Higher Education qualifications, as opposed to 49% in the non-signer group. However, Education did not predict Wu-Palmer scores [$F(1,63) = 0.14$, $\eta 2 = 0.002$, $p = 0.71$], nor did it interact with group effects [$F(1,63) = 0.13$, $\eta 2 = 0.004$, $p = 0.88$]. Hence, neither gender nor education were considered in further analyses of main effects on Wu-Palmer scores.

In addition, we assessed some crucial language and cognitive variables that may affect performance, notably English vocabulary size using the Vocabulary Size Test and non-verbal IQ using a modified version of KBIT-2 matrices task (Drown et al., 2022; Kaufman & Kaufman, 2004). The tasks are described in detail in the following section. Whilst the groups did not differ on IQ ($p > 0.05$), the DHH signers performed less well on the English vocabulary test than the hearing non-signers [$F(1,36) = 1222.97$, $\eta 2 = 0.97$, $p < 0.001$]. This observation is unsurprising, given that English is a second language for the DHH signers, but a first language for the hearing non-signers. However, the discrepancy is not considered to be crucial for the outcome, given that the target language is in the visual modality. Importantly, English Vocabulary scores did not correlate with our dependent variable, the Wu-Palmer scores indicative of the semantic similarity between the target meaning and the response [$r(1,64) = 0.06$, $p = 0.62$].

2.2. Tasks and Procedure

Overview. Most of the tasks and procedures used in the current study were developed by predecessor studies (Hofweber et al., 2022, 2023) and are available on the OSF entry | Which aspects of visual motivation aid the implicit learning of signs at first exposure?. In addition to the tasks shared on the OSF website, we administered the modified version of the KBIT non-verbal IQ test (Kaufman & Kaufman, 2004) and the English Vocabulary Size Test (Drown, et al., 2022). Whilst the predecessor studies used PsychoPy, this study administered all tasks using the Gorilla Experimental Builder (Anwyl-Irvine et al., 2020).

Procedure. Participants first watched a 4-minute weather forecast in the target Swedish Sign Language. Subsequently, they underwent an experimental task assessing meaning assignment. This paper only presents the metaphor inferencing aspect of the meaning assignment task.

Implicit exposure task. The task consisted of a 4-minute weather forecast in the target language Swedish Sign Language, Svenskt teckenspråk STS, deliberately chosen to differ from sign languages using mouthing based on M2L2 participants' first language English, i.e. ASL or BSL. Our paradigm is derived from an implicit learning paradigm developed for spoken languages (Gullberg et al., 2010, 2012). Crucially, participants watched the forecast without any implicit instructions. This was assumed to trigger automatic implicit meaning assignment processes akin to real-life immersion. Importantly, the forecast contained 22 target signs participants would later be tested about. The frequency with which targets occurred in the forecast was manipulated: 11 target signs occurred with a high frequency of 8 times, relative to the 11 low frequency targets shown only three times. For further details about the task, readers can refer to Hofweber et al. (2022, 2023).

Meaning assignment task. This task presented participants with the 22 weather forecast target signs, as well as 22 distractors, in fully randomised order. The 22 distractors consisted of formally similar STS signs that had not occurred in the forecast. For all 44 signs, participants were asked to indicate whether they recalled seeing them in the forecast. For the 22 target signs only, participants were given 20 s to assign meaning to the sign by typing it into the text box. This paper selectively zooms in on items involving TIME-SPACE metaphors.

Transparency task. The transparency scores for the 22 target signs had been collected by Hofweber et al. (2023). To assess the degree of transparency of individual target signs, 30 native speakers of English without any sign language expertise had been asked to assign meaning to each sign. Participants were presented with a PowerPoint presentation of slides individually displaying short videos of the 22 target items in two different pseudorandomized presentation orders. They were instructed to assign meaning to the signs. No contextual clues were provided. Responses were coded for accuracy. Transparency scores for each item were calculated as the proportion of accurate to total number of responses, cf. details below.

Gesture elicitation task. The scores were taken from Hofweber et al. (2023). They had administered a task eliciting spontaneous gestures for the meaning of the 22 target signs from 29 native English non-signers. Participants were presented with a PowerPoint presentation displaying the meanings of the 22 target items in written format on individual slides. Participants were randomly assigned to two different pseudorandomized presentation orders. To generate gesture-sign similarity scores, the elicited gestures were coded for formal similarity to target signs, cf. scoring details below.

KBIT. To assess IQ, participants completed a modified version of the nonverbal matrices of the Kaufman Brief Intelligence Test Second edition. This task was fully digitized, with sample question instructions written above each relevant matrix to introduce the task. Beyond the sample questions, only the abstract matrices were retained, e.g. the first test question was item 18, where subjects must intuit that if a heart matches with a heart with a vertical line, then an oval matches with an oval with a vertical line, rather than item 11, where a rabbit matches with rabbit ears and therefor an elephant matches with elephant ears. Participants were instructed that the task would continue for 7 min or until they got 4 questions in a row incorrect, and that they should answer as quickly and accurately as possible. The resulting non-verbal IQ score is computed as the number of questions correct within 7 min (or before the discontinue rule) for each participant.

English Vocabulary. This task was a modified, web-based version of the Vocabulary Size Test developed by Paul Nation (Drown et al., 2022; Nation 2012). Participants were presented with 42 items from the original assessment. For each, participants were asked to select the best definitions of a given English word from a set of four definitions. All words and definitions were presented in written English. For validation testing of this version of the Vocabulary Size Test, please see https://osf.io/pcsu6/.

2.3. Coding and scoring

Performance in the meaning assignment task. Performance was assessed by computing the semantic similarity between the target and actual response using the Wu-Palmer score, which is indicative of taxonomical relatedness, and based on the online WebVector calculator WebVectors: Computing associates and WordNet corpus (Oram, 2001). Scores range from 0.01 (low relatedness) to 0.9 (high levels of relatedness), to 1.0 (semantically identical). Although we identified only 4 metaphorical items, the task generated 244 data points involving metaphorical items across participants. Importantly, to match metaphorical and non-metaphorical items on transparency and iconicity, we only included signs with transparency scores of 0, 1, or 2 in the current analyses, cf. explanations below.

Transparency scores. The transparency score was computed as the number of times meaning was assigned accurately, e.g. if 10 participants assigned accurate meaning to a given sign, then that sign would have a transparency score of 10.

Gesture similarity scores. The gestures elicited in the gesture similarity task were recorded and subsequently scored for their formal similarity to target signs in terms of three criteria: (1) phonological overlap (Ortega et al., 2017); (2) representational strategy (acting, representing, drawing/moulding, or personifying, as per Ortega & Özyürek 2020); and (3) identity of the source of iconicity (Hofweber et al., 2023). The comparison resulted in gesture similarity scores ranging from 0 to 4.

Sign language similarity scores. The sign language similarity scores were created by adapting the method used to generated the gesture similarity scores described above (Hofweber et al., 2023). The 22 target signs were translated into either BSL or ASL, depending on the signers' first languages. All three involved sign languages (STS, BSL, ASL) were coded for the following criteria: (1) phonological make-up (handshape, location, orientation, and movement); (2) representational strategy (i.e. acting, handling, personifying), (3) identity of the source of iconicity (i.e. physical shape of a rainbow, movement of wind). For the phonological make-up, each aspect of phonology that matched between the target language and the L1 language was given a score of 1. Therefore, the similarity scores for this section ranged from 0 to 4. For the representational strategy, a match equated a score of 4, in order to grant it equal weight to the phonology. For the iconicity, a complete match gave a score of 4, a partial match gave a score of 2, and no match gave a score of 0. This meant that the three main categories (phonology, mode of representation, and source of iconicity) were each given equal weight in the final score. The sign similarity scores ranged from 0, representing no overlap at all, to 12, representing identical signs.

Metaphorical item selection. For the purpose of this study, four items were classified as involving SPACE-TIME metaphors, two in each frequency category: (1) high frequency items were AFTERNOON and NIGHT, (2) low frequency items were TOMORROW and TODAY. Non-metaphorical versus metaphorical items were distributed in equal 7/2 proportions across the high versus low frequency categories. Although the number of metaphorical items was low, repeated exposure in the weather forecast resulted in 366 occurrences for metaphorical low-frequency items, and 976 for metaphorical high-frequency items. For the purpose of this study, four target items were identified as involving TIME-SPACE metaphors.

All items involving metaphorical inferencing received low transparency scores of 0, 1, or 2. In fact, it had been this observation that motivated the predictions of the current paper (Hofweber et al., 2023). Hence, to match metaphorical and non-metaphorical items on transparency, we only included non-metaphorical "benchmark" signs with comparably low transparency scores of 0, 1, or 2 in the current analyses. This has importantly implications for the result interpretation. The study by Hofweber et al. (2023) had shown that learning from the weather forecast was limited to highly transparent items. Given that we only analyse low transparency items in this paper, we therefore describe the observed meaning assignment as comprehension, not semantic learning per se. However, exposure to the weather forecast still provided participants with implicit exposure to semantic context that may influence meaning assignment. For instance, if the concrete SPATIAL interpretation of a novel sign mismatches sense-making in the weather forecast context, this may trigger further semantic search processes, including metaphor inferencing.

3. Results

3.1. RQ1 The effects of sign language expertise on meaning assignment to metaphors in a novel sign language

As can be seen from the descriptive statistics presented in Table 2, overall performance was low, particularly amongst the non-signers, suggesting that metaphor inferencing in a novel sign language is a challenging task involving a certain level of language-specific arbitrariness.

Table 2
Descriptive Statistics Wu-Palmer scores by Frequency, Metaphoricity and Group.

Frequency	Metaphoricity	Group	M	SD	N
Low	Non-metaphorical	M1L2 Signers hearing	0.24	0.18	12.00
		M1L2 Signers DHH	0.24	0.12	15.00
		M2L2 Non-signers	0.22	0.10	37.00
	metaphorical	M1L2 Signers hearing	0.46	0.31	12.00
		M1L2 Signers DHH	0.55	0.27	15.00
		M2L2 Non-signers	0.26	0.13	37.00
High	Non-metaphorical	M1L2 Signers hearing	0.18	0.12	12.00
		M1L2 Signers DHH	0.29	0.33	15.00
		M2L2 Non-signers	0.31	0.25	37.00
	metaphorical	M1L2 Signers hearing	0.58	0.24	12.00
		M1L2 Signers DHH	0.40	0.18	15.00
		M2L2 Non-signers	0.20	0.08	37.00

The distribution of Wu-Palmer semantic similarity scores in each sub-groups was normal (Kolmogorov-Smith test hearing non-signers: $p = 0.18$, hearing signers: $p = 0.19$, DHH signers: $p = 0.20$), legitimizing the use of linear regression models. To explore the predictors of the dependent variable WuPalmer scores, we conducted linear mixed effects models considering random effects by slope and intercept using IBM SPSS 29 software. Firstly, we assessed the fixed effects of Group (M2L2 hearing, M1L2 hearing, M1L2 DHH), Frequency (high, low), Metaphoricity (metaphorical, non-metaphorical), and interactions. Importantly, we controlled for random effects of Items, Subjects, and the covariate Age, resulting in the statistics presented in Table 3. Most importantly, the fixed effects of Group and Metaphoricity were both significant. Overall, the two M1L2 signer groups performed better on meaning assignment than M2L2 non-signers, and metaphoricity overall facilitated meaning assignment. However, the Metaphoricity and Group effects interacted, suggesting differential patterns across groups. Moreover, the fixed effect of Frequency, i.e. evidence of learning, was not significant. This is in line with previous studies (Hofweber et al., 2020) showing that learning was reserved to highly transparent items, whilst this study zoomed in on low transparency items to match non-metaphorical to metaphorical items. The Frequency effect was similarly absent across sub-groups and metaphoricity conditions because the interactions were not significant. There was however a significant 3-way interaction between Group, Frequency and Metaphoricity that requires further exploration

To explore the interactions, we conducted mixed models considering the random effects and covariates described above. To avoid Type I errors, we conservatively interpreted post-hoc analyses applying Bonferroni corrections, although we discuss marginal effects as indicative of trends. The post-hoc tests revealed marked performance dissociations between M1L2 signers and M2L2 non-signers in terms of metaphor inferencing. Our first step was to explore Group effects by metaphoricity, so we modelled metaphorical and non-metaphorical items

Table 3
mixed model results for WuPalmer ~ Group + Metaphoricity + Frequency.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept	1	436	404.88	0<.001	0.48
Group	2	436	8.78	0<.001	0.04
Frequency	1	436	0.03	0.87	0.00
Metaphoricity	1	436	38.95	0<.001	0.08
Group * Frequency	2	436	0.52	0.59	0.00
Group * Metaphoricity	2	436	21.54	0<.001	0.09
Frequency * Metaphoricity	1	436	0.60	0.44	0.00
3-way interaction	2	436	3.73	0.02	0.02

separately (Tables 4 and 5; Fig. 2). This revealed that the effect of Group on meaning assignment performance was limited to metaphorical items. Whilst M1L2 and M2L2 participants performed identically on non-metaphorical items, the M1L2 signers significantly outperformed the M2L2 signers on metaphorical items.

Our second step was to explore differences in the effect of Metaphoricity by group, so we modelled the different sub-groups separately (Tables 6, 7 and 8; Fig. 3). Amongst both groups of M1L2 signers (hearing and DHH), the effect of Metaphoricity was significant and facilitative, and prevailed regardless of item exposure frequency. The facilitation effect was strongest amongst the M1L2 hearing signers and interacted marginally with Frequency in M1L2 DHH signers. Post-hoc analyses revealed that amongst the DHH signers, the Metaphoricity only facilitated meaning assignment in the low frequency condition (Table 9).Fig. 4.Fig. 5..

For non-signers, the reverse effect was observed, i.e. they trended to perform less well on metaphorical than non-metaphorical items. Although the overall effect was marginal, it interacted significantly with frequency. Further post-hoc tests revealed that the negative effect of Metaphoricity was significant in the high frequency condition, and marginally significant in the low frequency condition (Table 10).

To tease apart the effects of Deafness versus sign language expertise, it should be noted that throughout all steps of analysis, identical patterns were observed across the two groups with sign expertise, i.e. the DHH and hearing M1L2 native signers, rather than across the two hearing groups. Moreover, when group differences were observed, they occurred between signers and non-signers, not between DHH and hearing signers. This suggests that the causality for the phenomena observed in this study arises primarily from sign language expertise, not Deafness. To summarise, hearing and DHH signers displayed the identical opposite pattern to hearing non-signers, as illustrated in Figs. 2 and 3.

3.2. RQ2 cognate effects on meaning assignment to metaphors in a novel sign language

To explore item-level predictors of metaphor inferencing, we conducted further linear mixed models with the dependent variable Wu-Palmer scores for metaphorical items only, controlling for items and subjects. The analyses were conducted for the three groups separately. Amongst the group of M1L2 native signers, we investigated cognate transfer effects based on the similarity of linguistic L1 and L2 signs. Amongst the M2L2 non-signers, we explored whether gesture-sign similarity scores predicted metaphor inferencing.

In terms of cognate effects from formal similarity between L1 and L2 signs, these significantly predicted metaphor inferencing amongst the DHH M1L2 signers (Table 12; figure 4) but not amongst the hearing M1L2 signers (Table 11). Moreover, even in the DHH group, the predictive effect of L1-L2 sign similarity was only a marginal trend, when applying Bonferroni corrections.

The formal similarity between target signs and spontaneously elicited gestures representing target meaning predicted M2L2 metaphor inferencing (Table 13). However, figure 5 would suggest that the correlation is driven by only three data points with Wu-Palmer scores above 0.9, challenging the reliability of the finding. Hence, we conducted additional identical analyses removing these outlier Wu-Palmer scores.

Table 4
Mixed model results for non-metaphorical items wupalmer ~ Group + Frequency.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept	1	186	232.29	0<.001	0.56
Group	2	186	1.44	0.24	0.02
Frequency	1	186	0.50	0.48	0.00
Group * Frequency	2	186	1.73	0.18	0.02

Table 5
Mixed model results for metaphorical items for wupalmer ~ Group + Frequency.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept	1	61	459.67	0<.001	0.88
Group	2	61	30.42	0<.001	0.50
Frequency	1	61	1.55	0.22	0.02
Group * Frequency	2	61	9.38	0<.001	0.24

These analyses generated identical results, confirming a robust effect.

4. Discussion

The current study investigated how the detection and inferencing of metaphorical meaning in a novel language, in our case a sign language, is modulated by existing linguistic and gestural knowledge. From a theoretical perspective, we thus explored the cross-modal universality of metaphor comprehension, notably SPACE-TIME metaphors. For M2L2 non-signers, we predicted that they would find metaphor inferencing difficult because it involves the necessity to overwrite salient iconic spatial interpretations in favour of the metaphorical temporal interpretation. Conversely, for M1L2 native signers, metaphoricity should facilitate meaning assignment to novel signs because the iconic spatial cues help meaning assignment, and M1L2 signers possess an awareness of SPACE-TIME metaphor formation strategies in the visual modality, so they consider plausible further extrapolation to the metaphorical meaning level. At the items-level, we predicted that meaning assignment would be facilitated by L1-L2 cognate effects in the M1L2 learners, and by positive transfer of gestural repertoires in the M2L2 non-signers.

Our findings revealed a marked performance dissociation by group, in line with predictions. Firstly, M1L2 signers outperformed M2L2 signers at meaning assignment on metaphorical items. Secondly, in M1L2 signers, metaphoricity facilitated meaning assignment, whilst in M2L2 non-signers, the opposite trend was observed. Overall, M2L2 signers thus performed better at metaphor inferencing in the novel sign language than M2L2 non-signers. Hence, in M1L2 learners, an existing global awareness of strategies of metaphor formation is sign languages facilitated meaning assignment to metaphorical items. Moreover, M1L2 learners performed better on metaphorical than non-metaphorical items, suggesting that metaphoricity helped them integrate novel input with existing knowledge. The M2L2 learners displayed the reverse pattern performing less well on metaphorical than non-metaphorical signs. At the item-level, we detected cognate effects of L1-L2 sign similarity amongst the M1L2 learners, as well as analogous cognate effects of gestural repertoires on metaphor inferencing amongst the M2L2 learners. In this section, we discuss the implications of our findings for theories about the universality of metaphor processing, theories of transfer, the role of iconicity in sign language comprehension, as well as the implications for educational practices. Finally, we discuss limitations, and avenues for future research.

Regarding different theories of metaphor formation and comprehension, our results support an interactionist approach incorporating both universalist and language-specific aspects. Firstly. our results show that, in principle, it is possible to detect and correctly infer metaphors in a novel language because several participants succeeded in inferencing metaphorical meaning. On metaphorical items, the M1L2 native signers achieved semantic proximity to target scores in the range of 40% – 58%, which suggests high levels of mutual intelligibility. It supports the notion that metaphors are partially based on an intuitive universality, stemming either from humans' shared environment, their cognitive architecture, and, most likely, the mutual interrelation of environment and cognition. This intuitiveness was further supported by the finding that metaphor inferencing was predicted by the similarity of target signs to intuitively elicited gestures. However, participants' ability to assign meaning to metaphorical signs strongly interacted with existing

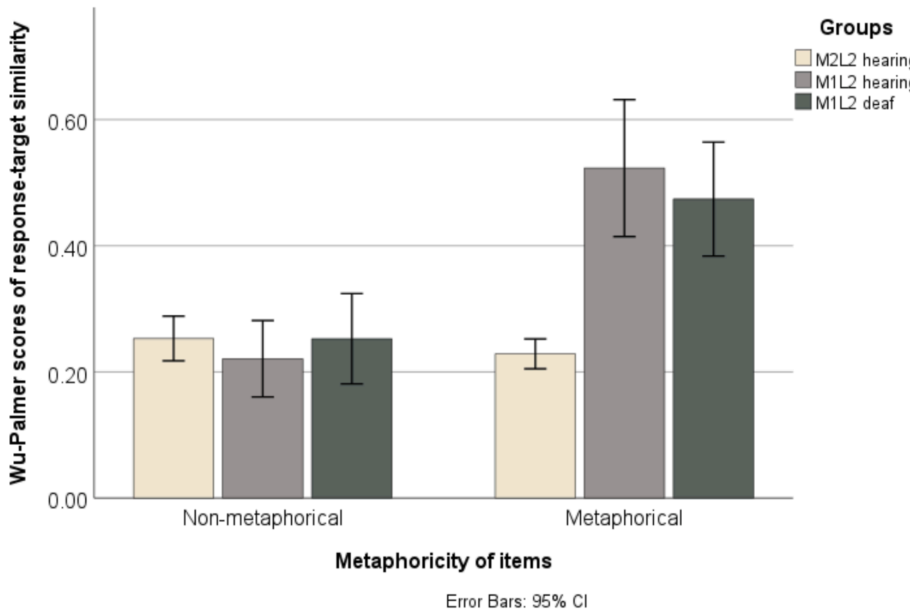

Fig. 2. Semantic similarity Target-Response Wu-Palmer scores by metaphoricity and Group.

Table 6
Mixed model results for m2l2 non-signers for wupalmer ~ Metaphoricity + Frequency.

Factor	*Df numerator*	*Df denominator*	*F-value*	*p-value*	η^2
Intercept	1	255	231.12	0<.001	0.48
Metaphoricity	1	255	3.05	0.08	0.01
Frequency	1	255	0.13	0.72	0.00
Metaphoricity * Frequency	1	255	6.29	0.01	0.02

Table 7
Mixed model results for m1l2 hearing signers for wupalmer ~ Metaphoricity + Frequency.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept	1	80	85.77	0<.001	0.52
Metaphoricity	1	80	14.37	0<.001	0.15
Frequency	1	80	0.14	0.71	0.00
Metaphoricity * Frequency	1	80	0.28	0.60	0.00

Table 8
Mixed model results for m1l2 dhh signers for wupalmer ~ Metaphoricity + Frequency.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept	1	101	111.25	0<.001	0.52
Metaphoricity	1	101	14.42	0<.001	0.12
Frequency	1	101	0.61	0.44	0.01
Metaphoricity * Frequency	1	101	3.41	0.07	0.03

knowledge and experience in modality-specific metaphor formation strategies, which emphasizes the role of cross-linguistic and cross-modal differences.

The importance of modality-specific knowledge was highlighted by the fact that the effect of metaphoricity on meaning assignment was reversed in native signers versus sign novices. Amongst M1L2 native metaphoricity resulted in enhanced performance, whilst amongst M2L2 non-native signers it represented a challenge. This suggests that metaphor inferencing was clearly influenced by whether participants disposed of language- and modality-specific knowledge. Strategies of metaphor formation across language modalities therefore differ considerably, resulting in a challenge for sign novices. Cross-modal transfer of generic awareness of the existence of SPACE-TIME metaphors was not sufficient for successful metaphor inferencing.

A key difference between sign and spoken languages is the greater presence of iconicity in sign languages. Hence, sign novices have limited experience in interpreting iconic signs (Hofweber et al., 2023). In the context of iconic signs with metaphorical meaning, even if the concrete iconic meaning is detected, there remains the need to overwrite the concrete meaning to abstract to the metaphorical meaning (cf. figure 1). Sign novices were often led astray by the iconic concrete SPATIAL meaning of SPACE-TIME metaphors, defaulting to a concrete iconic interpretation. To derive the accurate meaning, they would have needed to ignore the concrete SPATIAL meaning and abstract to the TIME metaphor. This seems to be particularly challenging in sign language, due to the very saliency of iconically expressed SPACE meanings. M1L2 signers, on the other hand, are experts in comprehending iconically mediated SPACE-TIME metaphors. They also initially draw upon the iconic SPACE meaning, but then successfully abstract to the abstract TIME interpretation. The failure of M2L2 sign novices to transition to the metaphor inference stage emphasizes the partially arbitrary nature of metaphors, posing a challenge to sign novices.

The contextualising role of exposure to the weather forecast further contributes to explaining the group difference. Clearly, M1L2 signers were more successful at drawing upon visual contextual cues than M2L2 non-signers. When the concrete SPATIAL interpretation mismatched the weather forecast context, they continued creative search processes for plausible alternative semantic interpretation. Aided by their modality-specific understanding of metaphors in sign languages, they were open to considering metaphorical interpretations. The M2L2 non-signer on the other hand reached a "premature closure", defaulting to the concrete iconic SPACE meaning, regardless of context.

The observed advantages of M1L2 native signers over M2L2 sign-novices may be due to sign language experience or neural plasticity effects of Deafness, i.e. enhanced visual processing drawing upon additional resources available in the auditory cortex (Scott et al., 2014). To tease apart the effects of sign language expertise and Deafness, we

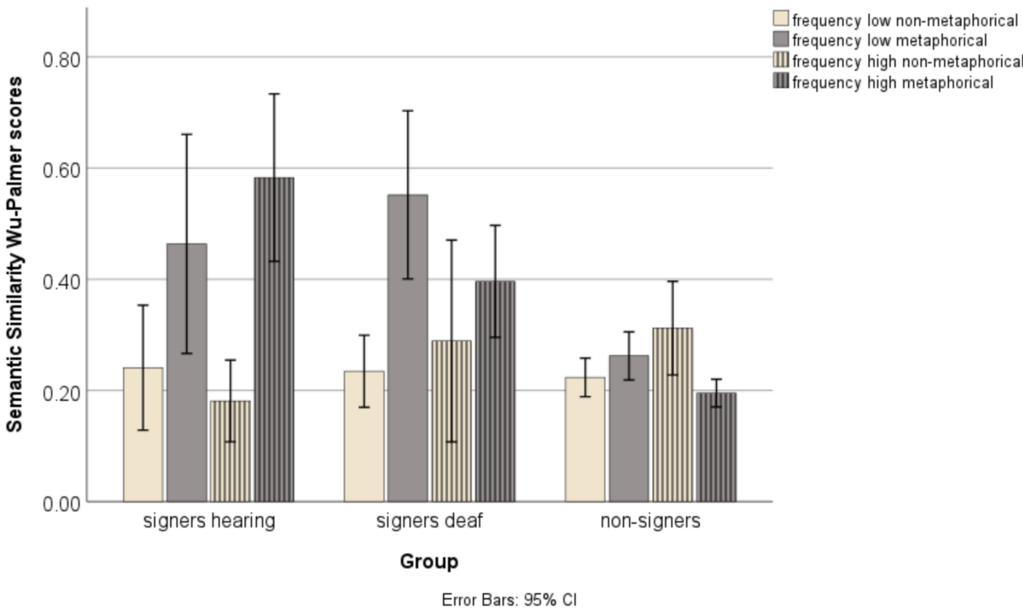

Fig. 3. Semantic similarity Target-Response Wu-Palmer scores by Group, Frequency and Metaphoricity.

Table 9
Mixed model results for m1l2 dhh signers by frequency condition for wupalmer ~ Metaphoricity.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept Low frequency	1	58	59.13	0<.001	0.50
Metaphoricity Low frequency	1	58	12.24	0<.001	0.17
Intercept High frequency	1	43	31.34	0<.001	0.42
Metaphoricity High frequency	1	43	1.91	0.17	0.04

compered the performance patterns of non-DHH hearing M1L2 signers, i.e. bimodal bilinguals, to the performance patterns of DHH M1L2 signers and non-DHH hearing non-signers. The hearing signers patterned with the DHH signers, rather than the hearing non-signers. This suggests that the observed effects were due to sign language expertise, rather than neural plasticity arising from Deafness.

With regards to different models of transfer, we found evidence for both macro-level and micro-level cognate effects. At the macro-level, M1L2 native signers transferred their existing knowledge of metaphor formation in sign languages onto the novel target language input, resulting in successful inferencing of metaphorical meaning. The positive transfer effect cancelled out and reversed the difficulty M2L2 non-native signers experienced in inferencing metaphorical meaning. Hence, our results reveal that global strategies of expressing metaphoricity transfer from first to second (sign) languages. Native signers were highly accurate in recognising SPACE-TIME metaphoricity in the novel sign language, suggesting that modality-specific linguistic expertise is relevant for detecting metaphoricity in second sign languages. However, modality-neutral more global conceptual knowledge of SPACE-TIME metaphors accessibke to the M2L2 group failed to result in positive transfer across sensory modalities, It is possible that the transfer of coceptual SPACE-TIME metaphoricty was blocked by the saliency of the concrete iconic SPACE interpretation.

At the micro-level, we identified cognate effects for linguistic items, as well as analogous transfer of gestural repertoires. Amongst M1L2 native signers, we observed facilitative transfer based on the similarity

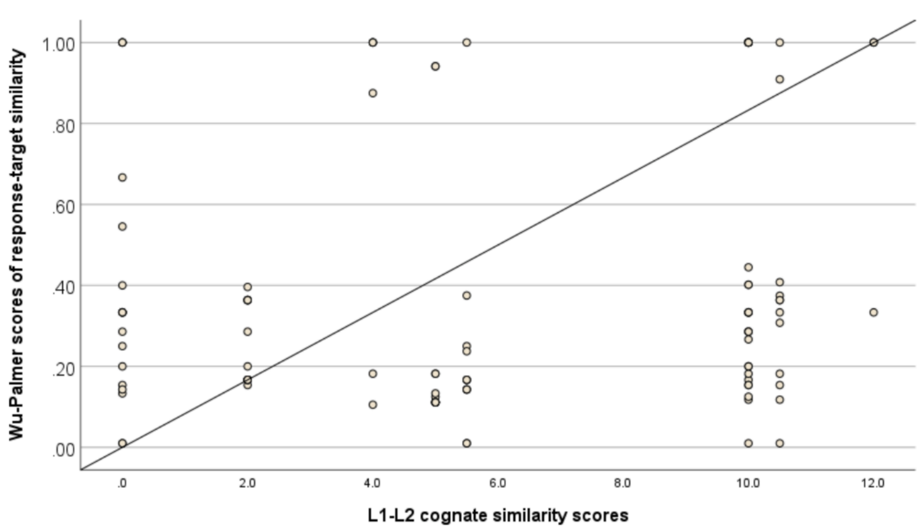

Fig. 4. Correlation between L1-L2 sign similarity and Wu-Palmer scores for metaphorical items.

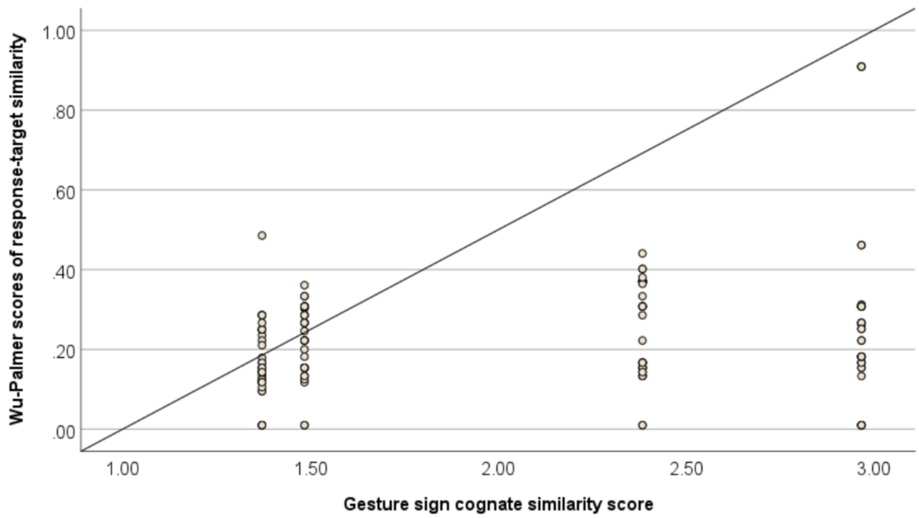

Fig. 5. Correlation of gesture-sign similarity scores and Wu-Palmer scores for metaphorical items.

Table 10
Mixed model results for m2l2 non-signers by frequency condition for wupalmer ~ Metaphoricity.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept Low frequency	1	146	294.41	0<.001	0.67
Metaphoricity Low frequency	1	146	3.89	0.05	0.03
Intercept High frequency	1	109	103.02	0<.001	0.49
Metaphoricity High frequency	1	109	4.05	0.047	0.04

Table 11
Cognate effects of L1-L2 sign similarity in M1L2 hearing signers.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept	1	11	95.53	0.00	0.90
L1-L2 sign similarity	2	11	0.25	0.78	0.04

Table 12
Cognate effects of L1-L2 sign similarity amongst M1L2 DHH signers.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept	1	55	39.24	0<.001	0.42
L1-L2 sign similarity	4	55	2.54	0.05	0.16

Table 13
Mixed model results of gesture-sign similarity scores predicting wu-palmer scores.

	Df numerator	*Df denominator*	*F-value*	*p-value*	η^2
Intercept	1	144	350.00	0<.001	0.71
Gesture-sign similarity	3	144	4.03	0.01	0.009

between individual linguistic signs in the first and second language. However, this effect was of weak size, and limited to DHH signers. Its limitation to DHH signers could be explained by effects of language dominance. The hearing bimodal signers may be more dominant in spoken English than ASL, which may weaken item-level transfer effects from sign. Although item-specific cognate effects contributed to metaphor detection, the main effect of modality-specific language expertise and awareness was stronger than the item-specific effect, pointing in the direction of the systemic macro-level nature of this transfer effect. Hence, although item-level similarity affected metaphor inferencing positively, the overall macro-level effect of sign language expertise and resulting awareness of modality-specific metaphor inferencing strategies had a more pervasive effect on performance.

Amongst M2L2 non-signers, the degree of similarity between target signs and spontaneously elicited gestures for equivalent meanings predicted metaphor inferencing. The observed gesture-sign transfer effects are in line with a recent study by Karadöller & Ortega (2024). However, they contradict null results by Hofweber et al. (2023) who did not detect cognate effects of gestural repertoire on meaning assignment using an identical experimental paradigm and items. The difference likely suggests that the gesturally-mediated cognate effects apply selectively to the domain of SPACE-TIME metaphoricity in the context of our experimental stimuli. It is possible that a universal intuitiveness of SPACE-TIME metaphoricity resulted in higher levels of overlap between co-speech gesture and conventionalised linguistic metaphors, increasing the likelihood and potential for transfer.

Our study has important implications for educational practices. Firstly, we observe that metaphor comprehension in a novel sign language is modulated by learner profiles, calling for bespoke teaching approaches. Whilst M1L2 learners with existing knowledge of sign languages have an easier time decoding meaning of metaphorical signs, M2L2 sign novices find it hard to infer metaphorical abstract meaning in novel sign languages. Sign language education for M1L2 learners can thus draw upon the facilitative effect of metaphoricity, whilst sign language education for M2L2 learners needs to develop methods that generate macro-level awareness of strategies of metaphor formation in sign languages. Moreover, M2L2 sign novices may benefit from building bridges highlighting the similarity between spontaneously elicited gestures and linguistic target signs. Future research should explore different strategies in intervention studies, e.g. involving explicit instruction.

Our results underline that SPACE-TIME metaphors are partially arbitrary and require learning. As such, L2 metaphor comprehension is shown to be subject to transfer effects. In M2L2 sign language learning contexts, this challenge may be exacerbated due to the visuo-spatial nature of sign, which draws learners' attention to the salient spatial meaning, thereby creating a tendency to jump to a concrete SPACE interpretation. To overwrite the intuitive concrete meaning, sign language novices may thus require additional explicit explanatory support to recognise common sources of SPACE-TIME metaphoricity in the novel visual modality. A possible teaching strategy may be to point out similarities between intuitively produced gestures and L2 target signs, thus enhancing the recognition of SPACE-TIME metaphors.

Our study is limited in that the number of metaphorical items were small. Hence, future studies involving greater numbers of metaphorical and non-metaphorical items are required to confirm our observations. Future research could also consider a greater variety of metaphors. A further limitation stems from the fact that all our metaphorical items were also low in transparency. To tease apart effects of transparency and iconicity from effects of metaphoricity, future research should present participants with highly iconic metaphorical items. However, this is challenging because low levels of transparency of signs often stem from their metaphoricity, so metaphorical items are inherently associated with low transparency scores. It should be noted, though, that our transparency scores were based on a task completed by M2L2 sign novices. If the transparency scores had been derived from M1L2 signers' responses, the task may have yielded higher levels of transparency scores. Future research should consider this factor. A further important consequence of the (inherent) confound between metaphoricity and (low) transparency is that our study is limited to investigating meaning assignment, rather than learning. To detect evidence of learning, we would need to observe frequency effects, which are often limited to highly transparent items (Hofweber et al., 2023).

Further suggestions for future research can be derived from the nature of our sample. Firstly, our small sample size might limit the generalization of the findings. Secondly, our groups differed in terms of education levels and age. Whilst age was statistically accounted for through its inclusion as a covariate in the analyses, the education differences remain. Future research should therefore replicate our results with larger samples. Moreover, future research could manipulate participants' language profiles to paint a more nuanced picture of transfer effects in the visual modality. For instance, it would be interesting to explore whether the positive transfer effects we observe amongst the M1L2 native signers are subject to a sensitive period, or whether late bimodal bilinguals, e.g. hearing adults who acquired a sign language later in life, would benefit equally from their sign language expertise when decoding metaphors in a novel sign language. Finally, our M1L2 sample is heterogeneous regarding first sign languages, i.e. ASL versus BSL. Although the majority of participants have first language ASL, we aimed to control for the first language differences through the item-level cognate effects, which differentiate by first language. L1-L2 similarity scores were based on either ASL or BSL, depending on signers' first language.

To conclude, our study investigated the effects of metaphoricity on meaning assignment as a function of existing knowledge of modality-specific metaphor formation strategies. A dissociation was detected in that M1L2 learners performed better on meaning assignment to metaphorical than non-metaphorical L2 signs, whilst M2L2 learners performed less well on items involving metaphor inferencing. Moreover, cognate effects facilitated metaphor inferencing. In the linguistic domain, we report some evidence of positive transfer from L1-L2 sign formal similarity in M1L2 learners. More generally, M2L2 learners were helped by their intuitive gestural repertoires. Our results suggest the effects of metaphoricity are modulated by existing linguistic knowledge, calling for teaching approaches tailored to learner profiles. Moreover, our results call for a hybrid approach to metaphor inferencing in sign SLA. Whilst iconic form-referent relationships are to some extent universally accessible, further abstraction and metaphorical inferencing is modulated by existing language- and modality-specific expertise, i.e. subject to learning.

Funding Sources

This work was supported by a Tier 1 grant provided by Northeastern University Boston and London to the Principal investigators, i.e. Dr. Julia Hofweber and Dr. Zhenghan Qi.

CRediT authorship contribution statement

Julia Hofweber: Conceptualization, Methodology, Project administration, Resources, Supervision, Funding acquisition, Writing – review & editing. **Lizzie Aumônier:** Methodology, Project Administration, Investigation. **Theresa Latham:** Methodology, Investigation. **Katherine Trice:** Methodology, Project administration, Investigation, Software, Formal analysis, Data curation, Writing – review & editing. **Zhenghan Qi:** Conceptualization, Methodology, Project administration, Resources, Supervision, Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Athanasopoulos, P., Samuel, S., & Bylund, E. (2017). The psychological reality of spatio-temporal metaphors. *Studies in figurative thought and language*, 295–321.

Athanasopoulos, P., & Bylund, E. (2024). Semantics and cognition: Spatiotemporal metaphors and time perception in L2 users. *The Routledge Handbook of Second Language Acquisition, Morphosyntax, and Semantics*, 545–558.

Beréndi, M., Csábi, S., & Kovecses, Z. (2008). Using conceptual metaphors and metonymies in vocabulary teaching. *Applications of Cognitive Linguistics, 6*, 65.

Chen, Y. C., & Lai, H. L. (2014). The influence of cultural universality and specificity on EFL learners' comprehension of metaphor and metonymy. *International Journal of Applied Linguistics, 24*(3), 312–336.

Chen, T. Y., & Myers, J. (2017). Worldlikeness: A web-based tool for typological psycholinguistic research. *University of Pennsylvania Working Papers. Linguistics, 23* (1), 4.

Ding, H., Qin, W., Liang, M., Ming, D., Wan, B., Li, Q., & Yu, C. (2015). Cross-modal activation of auditory regions during visuo-spatial working memory in early deafness. *Brain, 138*(9), 2750–2765.

Dove, G. (2016). Three symbol ungrounding problems: Abstract concepts and the future of embodied cognition. *Psychonomic bulletin & review, 23*, 1109–1121.

Drown, L., Giovannone, N., Pisoni, D. B., & Theodore, R. M. (2022, June 16). *Validation of two measures for assessing English vocabulary knowledge on web-based testing platforms: Long-form assessments.* 10.1515/lingvan-2022-0115.

Gentner, D. (2001). Spatial metaphors in temporal reasoning. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 203–222). Cambridge: MIT Press.

Gullberg, M., Roberts, L., Dimroth, C., Veroude, K., & Indefrey, P. (2010). Adult language learning after minimal exposure to an unknown natural language. *Language Learning, 60*, 5–24.

Heng, L. I. (2016). The psychological reality of spatial metaphors for time—Evidence from gesture and sign language. *Journal of Psychological Science, 39*(5), 1080.

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

Hofweber, J., Aumônier, L., Janke, V., Gullberg, M., & Marshall, C. (2023). Which aspects of visual motivation aid the implicit learning of signs at first exposure? *Language Learning, 73*(S1), 33–63.

Howe, R. (2024). Do metaphors help or hinder second language vocabulary acquisition? *Journal of Language Teaching, 4*(1), 1–6.

Karadöller, D. Z., Peeters, D., Manhardt, F., Özyürek, A., & Ortega, G. (2024). Iconicity and gesture jointly facilitate learning of second language signs at first exposure in hearing nonsigners. *Language Learning, 74*(4), 781–813.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test, Second Edition (KBIT-2)* [Database record]. APA PsycTests. 10.1037/t27706-000.

Kövecses, Z. (2005). Metaphor in culture: Universality and variation. *Cambridge UP*.

Lai, V. T., & Desai, R. H. (2016). The grounding of temporal metaphors. *Cortex, 76*, 43–50.

Lakoff, G., & Johnson, M. (2008). *Metaphors we live by*. University of Chicago press.

Lee, Y. Y. (2022). A conceptual analysis of typological distance and its potential consequences on the bilingual brain. *International Journal of Bilingual Education and Bilingualism, 25*(9), 3333–3346.

Marschark, M. (2004). Metaphors in sign language and sign language users: A window into relations of language and thought. In *Figurative language comprehension* (pp. 309–334). Routledge.

Máthé, Z. (2022). Comparative Analysis of Genitival Components in Time Metaphors. *Hungarian Studies Yearbook, 4*(1), 160–182.

McCullough, S., & Emmorey, K. (2021). Effects of deafness and sign language experience on the human brain: Voxel-based and surface-based morphometry. *Language, cognition and neuroscience, 36*(4), 422–439.

Meir, I., & Cohen, A. (2018). Metaphor in sign languages. *Frontiers in psychology, 9*, 1025.

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2011). Effects of cognate status on word comprehension in second language learners: An ERP investigation. *Journal of cognitive neuroscience, 23*(7), 1634–1647.

Motamedi, Y., Little, H., Nielsen, A., & Sulik, J. (2019). The iconicity toolbox: Empirical approaches to measuring iconicity. *Language and Cognition, 11*(2), 188–207.

Nation, P. (2012). *The Vocabulary Size Test*. https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-tests/the-vocabulary-size-test/VST-version-A.pdf.

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, 112*(37), 11684–11689.

Oram, P. (2001). WordNet: An electronic lexical database. Christiane Fellbaum (Ed.). Cambridge, MA: MIT Press, 1998. Pp. 423. *Applied Psycholinguistics*, *22*(1), 131-134.

Ortega, G. (2017). Iconicity and sign lexical acquisition: A review. *Frontiers in psychology, 8*, 1280.

Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of memory and language, 68*(4), 315–332.

Pichler, D. C., & Koulidobrova, E. (2023). The role of modality in L2 learning: The importance of learners acquiring a second sign language (M2L2 and M1L2 learners). *Language Learning, 73*(S1), 197–233.

Pinker, S. (1997). *How the mind works*. New York, NY: Norton.

Que, M., Jiang, X., Yi, C., Gui, P., Jiang, Y., Zhou, Y. D., & Wang, L. (2018). Language and sensory neural plasticity in the superior temporal cortex of the deaf. *Neural Plasticity, 2018*(1), Article 9456891.

Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research, 27*(1), 107–127.

Scott, G. D., Karns, C. M., Dow, M. W., Stevens, C., & Neville, H. J. (2014). Enhanced peripheral visual processing in congenitally DHH humans is supported by multiple brain regions, including primary auditory cortex. *Frontiers in human neuroscience, 8*, 177.

Sehyr, Z. S., & Emmorey, K. (2019). The perceived mapping between form and meaning in American Sign Language depends on linguistic knowledge and task: Evidence from iconicity and transparency judgments. *Language and Cognition, 11*(2), 208–234.

Simon, M., Campbell, E., Genest, F., MacLean, M. W., Champoux, F., & Lepore, F. (2020). The impact of early deafness on brain plasticity: A systematic review of the white and gray matter changes. *Frontiers in neuroscience, 14*, 206.

Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge University Press.

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.

Further reading

Casasanto, D. (2013). Development of metaphorical thinking: The role of language. *Language and the creative mind*, 3–18.

Emmorey, K. (2023). Ten things you should know about sign languages. *Current directions in psychological science, 32*(5), 387–394.

Gu, Y., Zheng, Y., & Swerts, M. (2019). Which is in front of chinese people, past or future? the effect of language and culture on temporal gestures and spatial conceptions of time. *Cognitive Science, 43*(12), Article e12804.

Nilsson, A. L. (2018). Expressing time through space: Embodying conceptual metaphors in an L1 vs. an L2 signed language. *Translation and Interpreting Studies, 13*(1), 6–26.

Trettenbrein, P. C., Papitto, G., Friederici, A. D., & Zaccarella, E. (2021). Functional neuroanatomy of language without speech: An ALE meta-analysis of sign language. *Human brain mapping, 42*(3), 699–712.

Ortega, G., & Özyürek, A. (2020). Types of iconicity and combinatorial strategies distinguish semantic categories in silent gesture across cultures. *Language and Cognition, 12*(1), 84–113.