COPiL CAMBRIDGE OCCASIONAL PAPERS *in* LINGUISTICS
Volume **18**, Article 1: pp. 1–23, 2026 | ISSN: 2050-5949

'With more internalisation comes greater embodiment'. Investigating whether embodiment effects vary as a function of the practice condition

George Blumenthal
Department of Theoretical and Applied Linguistics

Abstract In this study, I investigate whether embodiment varies with the practice condition. Since the turn of the century, it has been repeatedly shown that language learning and processing involve perceptual simulation (e.g., Stanfield & Zwaan 2001). Recently, similar questions have been posed with respect to the L2 (e.g., Chen, Wang & Zhang 2020). The notion that L2 embodiment is a straightforward function of proficiency has been called into doubt (Norman & Peleg 2022). Simultaneously, in the pedagogical domain, research on meaning and form-focused learning has been conducted for decades (Çelik 2016). Recently, it was found that participants trained on a novel morphological system in a meaning-focused way exhibited linguistic relativity effects (Williams & Xue 2024), and native-like neurophysiological responses to syntactic violations (Morgan-Short, Finger, Grey & Ullman 2012a). Given the link between linguistic relativity, native-like brain signatures and embodiment (Xue 2024), drawing these hitherto unconjoined threads together, there is reason to suppose that learning style is an overlooked modulator of L2 embodiment. As hypothesised, I found that participants trained on a novel morphological system pertaining to relative object size in a meaning-focused way displayed robust evidence of mental simulation, as indexed by an SPVT, while the form-focused group exhibited no such evidence. I discuss the implications of my findings, proposing the embodiment-compositionality hypothesis.

1 Introduction

1.1 Embodiment in the L2

Amodal theories of cognition (e.g., Fodor 1975) posit that the symbols involved in language processing are wholly abstract and therefore divorced from perception. Embodied theories, which emerged at around the turn of the century, challenge this idea by positing that these symbols are, in fact, grounded in perception (e.g., Barsalou 1999).

One branch of embodiment research involves action-compatibility-effects (ACEs), a phenomenon where implied directionality in a sentence matches the bodily movement required in response. For example, when reading a sentence in which movement away from the body is implied, such as *close the drawer*, responses were faster when participants were required to make a similar movement away from the body (Glenberg & Kaschak 2002). The Embodied Cognition Hypothesis has also been bolstered neuroscientifically. In a functional magnetic resonance imaging (fMRI) study, the same areas in the premotor and motor cortices of the brain were shown to be active when, for instance, reading sentences which described *kicking*, as when participants actually performed a kicking action (Hauk, Johnsrude & Pulvermüller 2004).

Another influential branch of embodiment research is mental simulation, a phenomenon which amodal theories struggle to explain (but see Mahon & Caramazza 2008). Evidence for mental simulation is obtained through SPVTs (sentence-picture-verification-tasks). In a pioneer study, it was found that RTs were significantly faster when the implied orientation of the object described in the sentence (e.g., *I hammered a nail into the wall*) matched the orientation of the object depicted in the following image (e.g., a horizontally oriented nail) compared to items in which they did not match (e.g., a vertically oriented nail), suggesting that orientation is mentally simulated (Stanfield & Zwaan 2001). Similar studies incorporating SPVTs have produced compelling evidence that a range of perceptual properties are mentally simulated, such as shape (Zwaan, Stanfield & Yaxley 2002), colour (Connell 2007), distance (Vukovic & Williams 2014), visual clarity (Yaxley & Zwaan 2007), and size (de Koning, Wassenburg, Bos & Van der Schoot 2016). The latter (size), as we shall see, is most pertinent to the current study, in which I train participants on a novel morphological system pertaining to relative object size.

Interestingly, the presiding research question seems to have shifted from whether embodied processes are universal to comparing the relative degrees to which the L1 and L2 are embodied. Philosophically, this reflects an increasingly pragmatistic orientation in the sciences of Mind, whereby questions relating to 'Truth' seem to have given way to questions relating to pedagogical praxis.

Some evidence suggests that the L2 is embodied to a partial or lesser degree than the L1 (Norman & Peleg 2022, Chen, Su & Wang 2024, Vukovic & Shtyrov 2014, Qian 2016, Ahlberg, Bischoff, Kaup, Bryant & Strozyk 2018). One behavioural experiment employing the SPVT found that while in the L1 object size and shape were mentally simulated, only object size was simulated in the L2 (Koster, Cadierno & Chiarandini 2019). But findings are conflicted: other studies suggest that the L1 and L2 are embodied to a comparable extent (Ahn & Jiang 2018, Dudschig, de la Vega & Kaup 2014, Vanek, Matić Škorić, Košutar, Matějka & Stone 2024, De Grauwe, Willems, Rueschemeyer, Lemhöfer & Schriefers 2014). For example, it was shown that mental simulation of object colour does not differ between the L1 and L2 in another study that employed the SPVT (van Zuijlen, Singh & Gunawan 2024). Results of this kind seem to contradict the intuition that because we acquire the L1 in real-life scenarios and the L2 in formal classroom settings, a more robust connection between the L1 and the sensorimotor system is likely to develop (Chen et al. 2024, Monaco, Jost, Gygax & Annoni 2019, Norman & Peleg 2022, Zhao, Vanek, Yang & Wang 2025).

These conflicting findings naturally raise the question of which factors modulate L2 embodiment. Evidence suggests that it may be modulated by *proficiency* (e.g., Birba, Beltrán, Martorell Caro, Trevisan, Kogan, Sedeño, Ibáñez & García 2020), *Age of acquisition* (AOA) (Norman & Peleg 2022), and *amount of exposure*. For instance, in a study conducted on learners of German, motor simulations were found to correlate with L2 proficiency (Ahlberg et al. 2018). But others cast doubt on the extent to which proficiency modulates embodiment (Kogan, García-Marco, Birba, Cortés, Melloni, Ibáñez et al. 2020, Monaco, Mouthon, Britz, Sato, Stefanos-Yakoub, Annoni & Jost 2023, Chen et al. 2020, 2024). For example, in an SPVT, high-proficiency L1 Hebrew speakers of English did not exhibit evidence of mental simulation of object shape (Norman & Peleg 2022).

One overlooked potential modulator of L2 embodiment, I contend, is *learning style*. Given the high correlation between proficiency and other factors such as AoA and exposure (Saito 2015, Segal & Kavé 2024), I suspect it is difficult to isolate the causal influence of each component separately. Therefore, my study will investigate whether learning a novel linguistic system in a way that approximates L1 acquisition leads to a greater degree of embodiment than learning it in a way that approximates formal classroom instruction, whilst controlling for proficiency across groups.

1.2 Meaning and form-focused learning

In a different domain altogether, research on language learning has pertained to *form* and *meaning*- focused conditions (see Çelik 2016). Form-focused learning involves focusing on *grammatical properties*, and meaning-focused learning involves focusing on *content*. The goal of the former is generally to produce correct linguistic forms, and the goal of the latter is generally to enhance communication (Eriksson 2023). A form-focused task might, for example, involve grammatical categorisation with error-feedback (Ellis 2001), and a meaning-focused task may involve taking part in a conversation, or reading a story (Nation 1996). In practice, the distinction between the two kinds of learning is not binary. Rather, it resides on a cline from very form-focused, characterised by *extralinguistic explanation* and *explicit learning*, to very meaning-focused, characterised by an emphasis on meaningful input, and *implicit learning* (Williams & Xue 2024).

Research harnessing novel linguistic systems suggests that meaning-focused learning may lead to a higher degree of embodiment than form-focused learning. In a neuroscientific study, participants who learned an artificial language, Brocanto2, in a meaning-focused manner exhibited native-like brain signatures to syntactic violations, while those who learned it in a form-focused manner did not (Morgan-Short, Steinhauer, Sanz & Ullman 2012b). More recently, it was found that those who learned a novel morphosyntactic system in a meaning-focused way exhibited *linguistic relativity effects*, as indexed by (in)transitivity-based categorisations in a triads-matching task (Williams & Xue 2024). Crucially, in both experiments, both groups exhibited comparable awareness at the level of the understanding, which refers to knowledge of the generalisations that govern form-meaning-connections (Schmidt 1990). Put differently, while both groups displayed comparable proficiency in the novel linguistic systems, the tacit aspects of their knowledge systematically varied as a function of the training condition.

Drawing these threads together, in Morgan-Short et al.'s study (2012a), the existence of native-like brain signatures amongst the implicit (meaning-focused) group, but not the explicit (form-focused) group may suggest that meaning-focused learning leads to a more robust connection with the sensorimotor system. In addition, it is speculated that linguistic relativity effects and embodiment effects share similar underlying mechanisms (cf. Xue 2024). Linguistic relativity theorists posit that language has a causal influence on perception, and language embodiment theorists posit that perception has a causal influence on language comprehension. Both domains, in which the nature of cognition is probed through the lens of the relationship between language and perception, share clear theoretical affinities. Therefore, the absence of research at its intersection is surprising (Thierry, Abdel-Rahman & Athanasopoulos 2024).

How might embodiment effects vary as a function of the training condition? Depth of processing, (DoP), defined as "the meaningfulness extracted from the stimulus", furnishes a possible explanation (Craik 1973). Meaning-focused learning is thought to induce deeper processing than form-focused learning (Dragomir & Niculescu 2022). This, in turn, may result in a greater degree of internalisation of knowledge, defined as its degree of elaboration, or the number of details stored in LTM that are associated with a lexical item (Leow & Mercer 2015). These details may be perceptually simulated during language comprehension, thereby leading to a greater degree of embodiment. In sum, one possible explanation is that meaning-focused learning leads to deeper processing, which leads to a greater degree of internalisation and therefore a more robust connection with the sensorimotor system.

1.3 Current study

In my study, drawing these hitherto unconjoined research threads together, I plan to investigate whether embodiment effects vary with the practice condition. My primary hypothesis is that a significant interaction between group and mental simulation will show that meaning-focused learners embody a novel morphological system to a significantly greater degree than form-focused learners, while controlling for proficiency across groups. Here, I will investigate this hypothesis with respect to novel determiners pertaining to object size. Participants will undergo a grammaticality judgment task (GJT) that measures superficial knowledge, or awareness at the level of understanding. Empirical evidence shows that pseudowords are capable of being embodied (Günther, Nguyen, Chen, Dudschig, Kaup & Glenberg 2020). I do not expect to find evidence of mental simulation amongst form-focused learners in isolation.

2 Experiment

2.1 Size & Mental Simulation

Participants will be presented with novel determiners *gi*, meaning relatively large, and *ro*, meaning relatively small. The reason for my decision to include size, as opposed to, say, orientation, is that evidence for the mental simulation of size is robust (de Koning et al. 2016), and has been replicated in the L2 (Koster et al. 2019). In order to ensure that awareness at the level of understanding remains constant between groups, both groups will be extralinguistically instructed as to the meanings of the novel determiners at the beginning of the task. This constitutes a subtle difference between my study and others (Williams & Xue 2024, Morgan-Short et al. 2012b), in which meaning-focused groups did not undergo extralinguistic instruction. Hence the terminological shift from 'learning' to 'practice'. Then, those in the form-focused-group (FFG) will undergo a form-focused task in which grammatical components are emphasised, whilst those in the meaning-focused-group (MFG) will read and answer comprehension questions about a fictional story within which the novel determiners are incorporated. Then, mental simulation of the novel determiners will be measured by an SPVT in which relative object size is manipulated. In order to further investigate the aforementioned hypotheses, I will investigate whether RTs to (in)congruent fillers items vary along the aforementioned parameters.

2.2 Participants

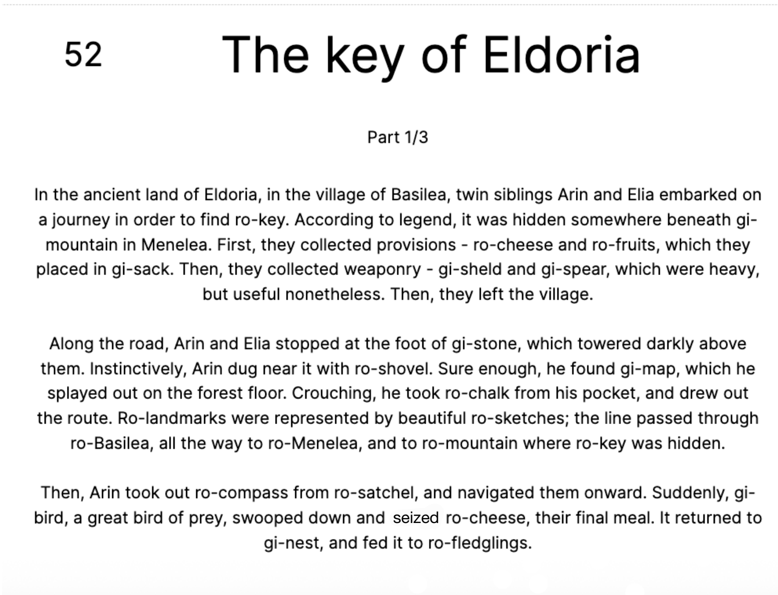
All participants were native speakers of English with no history of dyslexia. They were recruited via Prolific, Facebook, and emails sent to undergraduates studying linguistics and psychology. Overall, 150 participants started the experiment, and 25 failed to progress to the main task, leaving usable data available for 125 participants. Among these participants, age ranged from 18 to 73 years (M = 29.66, SD = 11.62). Mean age in the FFG group was 27.77 (SD = 7.04), and in the MFG was 33.33 (SD = 15.11). The two groups did not differ significantly in age.

2.3 Methodology

2.3.1 Materials

Stage 1: Training

Meanings of the novel determiners were specified relative to the average backpack. In another study, size was designated relative to the "average dog" (Chen, Guo, Tang, Zhu, Yang & Dienes 2011). Since dogs vary significantly in size, this did not strike me as a sensible criterion. Backpacks, however, are of relatively unambiguous magnitude. Participants were instructed that *gi* denoted 'larger than the average backpack', and *ro* denoted 'smaller than the average backpack'. In the FFG, participants were given 4 minutes to complete a *forced choice classification task*, namely a multiple-choice task in which they were to classify items as 'gi', 'ro', or ‘neither/either’. Feedback was not provided. In case participants reached the end within the time given, I designed the task such that it would cycle back to the beginning. Those in the FFG received 93 items. In the MFG, on the other hand, participants were presented with a short story named *The Key of Eldoria*, which was divided into 3 parts, as exemplified in Figure 1. Participants answered 12 comprehension questions, 4 after each part. In my study, I aimed to retain constancy across all variables apart from the practice condition. Importantly, this meant designing the practice materials such that both groups spent a comparable amount of time undergoing their respective tasks. Each of the 3 parts were displayed for 70 seconds. A timer was displayed in the top-left corner.


52 The key of Eldoria

Part 1/3

In the ancient land of Eldoria, in the village of Basilea, twin siblings Arin and Elia embarked on a journey in order to find ro-key. According to legend, it was hidden somewhere beneath gi-mountain in Menelea. First, they collected provisions - ro-cheese and ro-fruits, which they placed in gi-sack. Then, they collected weaponry - gi-sheld and gi-spear, which were heavy, but useful nonetheless. Then, they left the village.

Along the road, Arin and Elia stopped at the foot of gi-stone, which towered darkly above them. Instinctively, Arin dug near it with ro-shovel. Sure enough, he found gi-map, which he splayed out on the forest floor. Crouching, he took ro-chalk from his pocket, and drew out the route. Ro-landmarks were represented by beautiful ro-sketches; the line passed through ro-Basilea, all the way to ro-Menelea, and to ro-mountain where ro-key was hidden.

Then, Arin took out ro-compass from ro-satchel, and navigated them onward. Suddenly, gi-bird, a great bird of prey, swooped down and seized ro-cheese, their final meal. It returned to gi-nest, and fed it to ro-fledglings.

Figure 1 Extract from the story given to the meaning-focused practice group.

Stage 2: Grammaticality Judgement Task

In order to measure awareness at the level of understanding, or superficial knowledge, all participants were presented with a *grammaticality judgement task* (GJT), in which they were tasked with deciding whether items were well-formed or malformed. Participants were given 2 minutes to answer 20 questions. The structure of this task followed Zhao, Kormos, Rebuschat & Suzuki’s (2021) two-alternative forced choice test used to assess learning in the context of the incidental learning of novel determiners.

Stage 3: Sentence-picture-verification (SPVT)

Then, in order to measure mental simulations, all participants were presented with a SPVT. Each item contained a sentence followed by an image. In the sentence, an object was described as *gi-* or *ro-*, for instance, ‘The archaeologist examined gi-bone'. Then, an image was presented, in which either a large (*gi*) or small (*ro*) object was depicted. Participants were instructed to decide whether the object depicted in the image had been described in the preceding sentence. In the congruent condition, the size of the object described in the image matched that of the object depicted in the image, and in the incongruent condition, they mismatched. In the filler condition, on the other hand, the object depicted in the image had not been described in the sentence. For example, 'The pupil threw gi-snowball' was followed by an image of a hammer. An example of a fully counterbalanced trial item is given in Table 1. A practice phase with 8 items was administered.

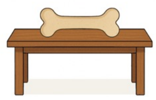
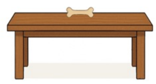
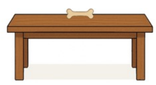
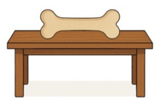
Condition	**Sentence**	**Image**
Congruent	The archaeologist examined gi-bone.	
Incongruent	The archaeologist examined gi-bone.	
Congruent	The archaeologist examined ro-bone.	
Incongruent	The archaeologist examined ro-bone.	

Table 1 Example of a fully counterbalanced item.

The SPVT included a total of 40 items: 20 critical trials, and 20 fillers. Using a 2x2 Latin square design, I counterbalanced all items, so that in each of the 4 lists was a unique variation of each trial combination. Within each condition, there were an equal number of size-matched combinations (*gi*-*gi*, *ro*-*ro*), where the first term of the combination refers to the determiner incorporated in the sentence, and the second term of the combination refers to the size of the object depicted in the image. Filler items remained invariant across lists. Anticipating that participants may have developed a strategy whereby they ignored the novel determiners and simply attended to each individual object, 12 probe trials were included in which after participants had indicated whether an object had appeared in a sentence, they were asked which determiner had appeared in the sentence. Feedback was provided in these trials.

All sentences consisted of a person, followed by a verb phrase, followed by a description of the object that would appear (or not, in the case of filler items) in the following image, for example 'The adventurer looked at gi-compass'. All images were generated by ChatGPT (OpenAI 2025). The inclusion of some object whose size remained constant over the course of the task was necessary so that size could be gauged relative to the average backpack. Therefore, following de Koning et al. (2016), I included a table in all images, relative to which the size of relevant objects could be gauged. Images were edited using Adobe Express (`new.express.adobe.com`).

2.3.2 Procedure

All tasks were implemented on Gorilla (`www.gorilla.sc`), an online experiment builder (Anwyl-Irvine, Massonnié, Flitton, Kirkham & Evershed 2020). All participants underwent the experiment online. An overview is shown in Figure 2. First, participants filled out a primary questionnaire, whose purpose was to ensure that they were native speakers of English, and had no history of dyslexia. Then, all participants were extralinguistically instructed as to the meanings of the novel determiners. Then, participants were randomised as to whether they underwent meaning or form-focused practice, the former of which contained a narrative comprehension task, the latter of which a grammatical categorisation task. There was a 50% chance they would enter each group. Participants in the FFG were given 4 minutes to complete their task. Those in the MFG were given 70 seconds to read each of 3 sections of the story, after which comprehension questions followed. No time limit was given for the completion of comprehension questions.

Then, all participants underwent the GJT, which measured awareness at the level of understanding. Then, they underwent the SPVT, which measured mental simulation of relative object size. Participants were instructed to press the spacebar in order to indicate they read and understood the sentence. Then, a fixation cross appeared (see below), after which an image appeared. Participants were instructed to press 'M' if the object described in the sentence appeared in the image, and 'Z' if it had not. They were told to rest their fingers on the relevant keys for the duration of the experiment, and to be as quick and intuitive as possible. There was a 25% chance that each of the 4 lists would be chosen. All stimuli were fully randomised.

Studies have shown that if an image is shown soon after the sentence (0-200ms), there is a *mismatch effect*. That is, in congruent conditions, where the relevant property of the object depicted in the image matches that of the object described in the sentence, in this case relative size, RTs will be slower than in incongruent conditions, where they mismatch. On the other hand, if an image is shown later (1200ms+), there is a *match effect*, that is, RTs in congruent conditions are shorter

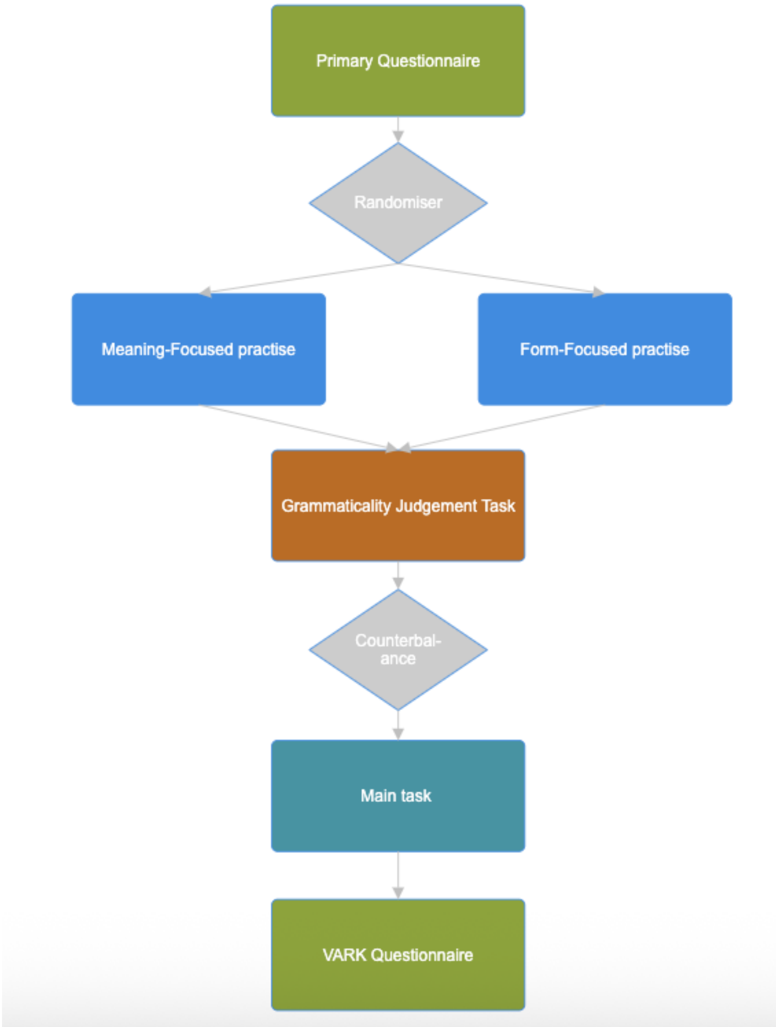


Figure 2 Experiment flow.

(see e.g., Liu, Wang & Han 2024). Opting for the latter, after the sentence was processed, a fixation cross appeared in the middle of the screen for 1500ms, at which point the image appeared.

2.4 Results

All data analysis was executed on R (R Core Team 2024). As a side note, in order to ensure that my results were not skewed by the inclusion of participants of a higher than conventional age range, I statistically tested whether the match effects of those aged 60+ constituted outlier data. I found the means of all three to be within 1SD (363ms) of the overall mean (168ms) (z-scores of -0.81, -0.57, and 0.63 respectively).

2.4.1 Data filtering

Firstly, I filtered out fillers and practice items. After this, I removed 2 participants for not completing enough trials, one of whom completed 1 trial, the other 2. Out of those left, all trials were complete and accuracy was 100%. This is likely to have been a consequence of the lengthy (8 trials) practice phase, in which feedback was provided.

Then, I calculated separate means per condition (congruent/incongruent) per participant. Following de Koning et al. (2016), I removed data points > 2.5SD above and below combined participant means (a removal of 5.32% oof the data), and recalculated them. Similarly to other studies in which data 4000ms or higher was removed (e.g., Kaan 2014), I removed 2 participants with combined mean RTs of over 4000ms. Of the 58 left, 27 were in the FFG, and 31 were in the MFG. Mean age in the FFG was 27.77 (SD = 7.04), and in the MFG was 33.33 (SD = 15.11). Group ages between groups were not statistically different.

After data filtering procedures were applied, each participant's score was calculated out of 20. Then, mean GJT scores were calculated per group from those of their participants. Those in the FFG (n = 27) scored a mean of 18.44 (SD = 2.82), out of 20 (92.2%), and those in the MFG (n = 31) scored a mean of 17.71 (SD = 4.34) out of 20 (88.55%). As hypothesised, there was no significant difference between the scores of both groups, $t(52.04) = 0.77$, $p = .442$.

2.4.2 SPVT

Firstly, I aimed to test the hypothesis that mean RTs, regardless of group, differed by condition. A paired t-test revealed that RTs were significantly shorter in the congruent (M = 1413ms, SD = 552ms), 95% CI [1268, 1558], than incongruent condition (M = 1581ms, SD = 692ms), 95% CI [1399, 1763], $t(57) = -3.628$, $p > .000612$, yielding a mean difference of 168ms, SD = 353ms, 95% CI [75, 261], $d = .476$, indicating a small to moderate effect size (Sullivan & Feinn 2012).

Then, aiming to test whether this effect varied as a function of the practice condition, I conducted a 2×2 mixed ANOVA with condition as a within-subjects factor and group as a between-subjects factor. There was no significant effect of group $F(1, 56) = 1.157$, $p = .287$, indicating that overall between-group RTs did not significantly differ. In line with the above, there was a significant effect for condition $F(1, 56) = 14.024$, $p < .000428$, $\eta_p^2 = .200$, the latter value indicating a large effect according to Richardson (2011). Vitally, I found a significant interaction between group and condition $F(1, 56) = 4.527$, $p = .038$, $\eta_p^2 = .075$. The partial eta-squared value indicates a medium effect (Richardson 2011). This provides sufficient evidence to reject the null hypothesis. Mean RTs per group per condition are exhibited in Table 2 and Figure 3. Mean match effect per group is shown in Figure 4.

In order to investigate the hypothesis that the meaning-focused group mentally simulated the pseudowords to a significantly greater extent than the form-focused group, I conducted paired t-tests on both groups separately. Amongst the FFG, I found an insignificant difference between congruent (M = 1373ms, SD = 513ms), 95% CI [1170, 1576], and incongruent (M = 1438ms, SD = 520ms), 95% CI [1232, 1644], conditions, yielding a mean match effect of 65ms (SD = 256ms), 95% CI [-36, 166], $t(26) = 1.31$, $p = .200$, $d = .253$. Amongst the MFG, on the other hand, I found a significant difference between congruent (M = 1448ms, SD = 590ms), 95% CI [1231, 1644], and incongruent (M = 1706ms, SD = 800ms), 95% CI [1412, 2000], conditions, $t(31) = 3.57$, $p = .00122$, $d = .64$, indicating a medium effect size, yielding a mean match effect of 258ms (SD = 403ms), 95% CI [111, 406].

Moreover, I conducted a Welch's independent-samples t-test, revealing that the match effect was significantly larger in the MFG than the FFG (see above), yielding a mean difference of 194ms, 95% CI [17, 369], $t(51.5) = 2.21$, $p = .031$.

Group	Mean Congruent	Mean Incongruent	Match Effect
Meaning-Focused	1448ms (SD= 590ms)	1706ms (SD = 800ms)	258 (SD = 403ms)
Form-Focused	1373ms (SD = 513ms)	1438ms (SD = 520ms)	65 (SD = 256ms)

Table 2 Participant-level RTs by condition by group.

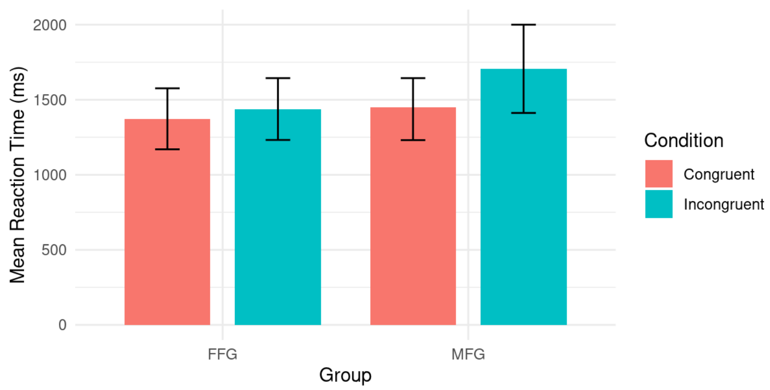


Figure 3 Participant-level RTs by condition by group. Error bars represent 95% confidence intervals.

2.4.3 Fillers

Filler items, to recapitulate, included a sentence incorporating 'gi' or 'ro', followed by an image in which mismatched the object described in the preceding sentence. I divided all filler items into congruent fillers, in which the size of the object depicted in the image matched that of the mismatched object described by the sentence, and incongruent fillers, in which they mismatched. The same participants were included as above. Following the same data filtering procedures, a unique participant mean for congruent and incongruent fillers was calculated. Data points 2.5SDs away from each participants' combined means were removed, and means were recalculated. Again, there were no incorrect answers to eliminate.

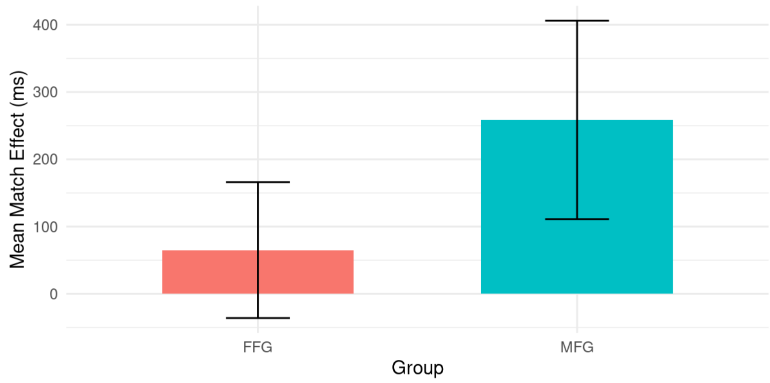


Figure 4 Match effect per group. Error bars represent 95% confidence intervals.

Conducting a *t*-test across all participants, I found a significant difference between RTs to congruent filler items (M = 1183 ms, SD = 516ms), 95% CI [1047, 1319], and incongruent filler items (M = 1262ms, SD = 481ms), 95% CI [1135, 1388], yielding a mean difference of 79ms, $t(57) = 3.2154$, $p = .02147$, $d = .422$, 95% CI [30, 128]. I conducted a 2×2 ANOVA with condition as a within-subjects factor, and group as a between-subjects factor. I found no effect of group, $F(1, 56) = 2.176$, $p = .146$, $\eta_p^2 = .037$, indicating that overall RTs did not differ significantly by group; a significant effect for condition, $F(1, 56) = 10.18$, $p = .00233$, $\eta_p^2 = .154$, and, finally, no significant interaction between group and condition, $F(1, 56) = 0.022$, $p = .883$, $\eta_p^2 < .001$. Mean RTs per group per filler condition are displayed in Table 3 and Figure 5. In sum, I found a significant difference between congruent and incongruent filler items across all participants. This effect did not vary as a function of the practice condition. It is important to point out, however, that fillers were not counterbalanced, that is, they were fixed across lists. Evidence, therefore, should be interpreted with caution.

Group	**Congruent**	**Incongruent**
Meaning (MFG)	1272ms (SD = 627ms)	1348ms (SD = 569ms)
Form (FFG)	1080ms (SD = 331ms)	1163ms (SD = 340ms)

Table 3 Mean RT per filler condition per group.

2.5 Discussion

I trained two groups of participants on novel determiners pertaining to relative object size. As expected, both groups scored indistinguishably in a grammaticality judgement task (GJT), indicating comparable awareness at the level of understanding. My first hypothesis, namely that there would be a match effect across all participants, thereby indicating mental simulation of the novel determiners, was borne

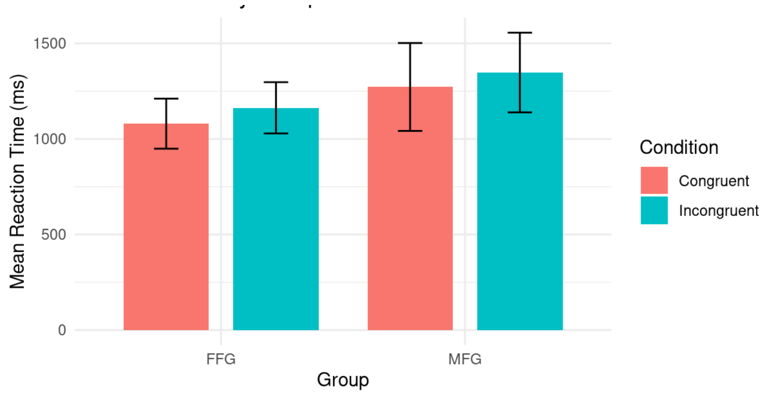


Figure 5 Mean RT per filler condition per group. Error bars represent 95% confidence intervals.

out. My principal hypothesis, namely that those in the meaning-focused group (MFG) would exhibit mental simulation effects to a significantly greater degree than those in the form-focused group (FFG), was also borne out. In sum, the significant interaction between group and condition suggests that robustness of mental simulation varied systematically between learning conditions. Interestingly, I also found that all participants rejected congruent filler items significantly faster than incongruent ones. Incidentally, the match effect was of approximately the same magnitude as the one exhibited by the FFG in critical trials, which was found to be insignificant. Since filler items were not counterbalanced, however, this finding should be interpreted with due circumspection.

Minimally, regardless of the practice condition, I have demonstrated compelling evidence that novel pseudowords are capable of being embodied. This contradicts studies showing null little or no evidence of mental simulation in the L2 (Chen et al. 2020), and corroborates those in which the opposite was found (Ahn & Jiang 2018).

My findings corroborate those of (Williams & Xue 2024), in which it was found that *linguistic relativity effects* were modulated by the learning condition. But in their study, there was no guarantee that those in both groups spent an equal amount of time practising. I have systematically built on this work by a) ensuring that a comparable amount of time is spent on the determiners between groups, and b) showing that embodiment effects, similarly to linguistic relativity effects, can be modulated by the practice condition, elucidating an interesting similarity between the two domains of research that merits further investigation (cf. Xue 2024). My findings corroborate those of Morgan-Short et al. (2012a), which found that only those who practised a novel linguistic system in a meaning-focused manner exhibited native-like brain signatures in response to syntactic violations, while those who practised it in a form-focused manner did not. I have systematically built on their work by demonstrating that differences in the underlying aspects of knowledge between groups, which vary as a function of the practice condition, can be gauged behaviourally, not only neuroscientifically, as they suggest.

In sum, the MFG showed robust evidence of mental simulation of the novel morphological system. Tentative evidence suggests that the FFG may have undergone less detailed mental simulations. In what follows, I will explain the ramifications of my results on models of second language learning, and outline a theoretical hypothesis that explains them.

2.5.1 Implications for L2 learning

There is significant controversy regarding which factors modulate L2 embodiment (see section 1). Of course, due caution should be exercised with respect to extrapolating the conclusions of my experiment, which harnesses an artificial language paradigm, to bilingual cognition at large (see Vigliocco, Convertino, De Felice, Gregorians, Kewenig, Mueller, Veselic, Musolesi, Hudson-Smith, Tyler, Flouri & Spiers 2024). Still, my findings suggest that *teaching style*, that is, the extent to which meaning or form-focused learning is incorporated into one's curriculum, may be a longitudinal predictor of L2 embodiment. Meaning-focused practice has been shown to induce native-like brain signatures following months of no exposure (Morgan-Short et al. 2012a), providing preliminary evidence for this hypothesis.

In Wang & Zhao's (2024) adaptation of (Bergen & Chang 2013) model of mental simulation in the L1, the first stage (contextual analysis) is modulated by *length of immersion in L2 context*, as well as *L2 proficiency*. The second stage (contextual resolution) is modulated by *amount of L2 classroom instruction*. The third stage, *embodied simulation*, is also modulated by *instructional context*. Under this framework, embodied simulation is reducible to a) proficiency and b) amount of time spent practising the L2. Their model, which includes *amount of instruction*, is blind to the ways in which this instruction may be delivered. *Learning style*, namely the extent to which meaning-focused materials are incorporated into one's curriculum, naturally suggests itself from my results as a modulator of L2 embodiment that should be incorporated into their model. Note also that in their own study, no interaction was found between L2 proficiency and mental simulation, which would seem to contradict the core claims of their model. Finally, it is fatalistic to assume a binary distinction between *L2 immersion* and *classroom instruction*, for my results suggest that the latter can be delivered in a way that approximates the former.

But this begs the question: if the goal of L2 acquisition is to attain higher levels of proficiency, then why bother with embodiment effects? Meaning-focused learning strategies have been linked with improved critical thinking skills (Nation 2001), willingness to communicate (Dehqan, Azizi & Miri 2022), collocational knowledge (word order) (Noroozi 2023), and enjoyment (Nation 2001: 8), which is crucial in order to sustain motivational levels. If, as I have demonstrated, meaning-focused learning drives L2 embodiment, and is beneficial in the aforementioned ways, then embodying the L2 can be described as beneficial insofar as it incorporates meaning-focused learning strategies. In fact, I would speculate that L2 embodiment may play a causal role in the aforementioned benefits that come to be attributed to meaning-focused learning.

These insights reflect calls for a shift of paradigm from form to meaning-focused learning (Dragomir & Niculescu 2022, Williams & Xue 2024). My claim is not that form-focused learning is inherently inferior. Form-focused teaching has been repeatedly shown, for example, to be useful for grammar acquisition (e.g., Azizpour & Alavinia 2021), and fluency & accuracy (Chen & Li 2022). Furthermore, deep processing plausibly requires greater cognitive effort, so harnessing meaning-focused tasks in all situations is unlikely to be a sustainable learning strategy. Rather, a harmonious unification of both should be sought. Longitudinal studies should investigate the nature of this combination.

Much of the research at the intersection between embodied cognition and language learning involves leveraging 'embodied styles of learning', by which I mean direct use of the body during the learning process, in order to improve vocabulary acquisition. For example, committing motor-congruent actions (Casasanto & de Bruin 2019), and observing semantically congruent gestures (Hughes-Berheim, Cheimariou, Shelley-Tremblay, Doheny & Morett 2022) have been shown to facilitate novel pseudoword learning. But my findings suggest that embodied styles of learning (see Glenberg, Jaworski, Rischal & Levin 2007, Gómez & Glenberg 2022) are not required in order to embody the L2. Given that embodied learning strategies can often be resource, time, and energy-intensive, and moreover that non-able-bodied children and adults may not be able to engage with them (cf. van Grunsven 2025), this insight is practically significant.

2.5.2 Embodiment-Compositionality Hypothesis (EC)

In the philosophy of language, compositionality is the principle that the meaning of a sentence is constituted by the sum of the meanings of its parts (Szabó 2024). Taking inspiration from this, I suggest that perceptual (mental and motor) simulations may be structured compositionally. Meaning-focused learning, I contend, may lead to fully integrated and therefore *compositional simulations*, while form-focused learning may lead to non-integrated and therefore *non-compositional simulations*. Recall (see Introduction) the hypothesis that meaning-focused learning induces deeper processing of the materials. This, I suggest, may be a consequence of emotional salience or personal connection, and should be further investigated. Regardless, as a result of deeper processing of meaning-focused materials, a larger number of perceptual symbols relating to a linguistic label may be incorporated in long-term-memory (LTM). Activation may spread from the linguistic label to the perceptual symbols around it, leading to fully compositional simulations.

According to this interpretation, in my study, while form-focused participants mentally simulated spatial schemata corresponding with the pseudoword and its referent (the object) separately, meaning-focused participants mentally simulated the pseudoword in combination with its referent in a fully unified manner. For example, when presented with stimulus 'gi-dog' (*gi-* means 'big'), meaning-focused participants mentally simulated *a large dog*, while form-focused participants mentally simulated an unintegrated spatial schema corresponding with 'largeness' and 'a dog' separately (*largeness* + *dog*). Critical trials required fully integrated simulations for congruency effects, explaining why the meaning-focused group exhibited demonstrable effects while the form-focused group did not. The EC also explains why both groups indiscriminately exhibited match effects with respect to filler items. To recapitulate, in congruent filler items, which required 'no' responses, only the size but not the object matched (for example, 'gi-dog' followed by a large pineapple). Therefore, unintegrated mental simulations could plausibly lead to congruency effects among filler items, explaining why the effect did not vary as a function of the practice condition. This should be further investigated in an experiment which counterbalances filler items.

The EC is supported by studies that find partial evidence of embodiment amongst L2 learners. For instance, it was found that L2 learners, like native speakers, relax the zygomatic muscle when 'smiling' is described. But unlike native speakers, they also do so when reading negated versions are processed (i.e., 'I am not smiling) (Foroni 2015). Participants in this study "learnt after the age of L2 through *scholastic programs*" (my emphasis). As such, they are likely to have undergone predominantly form-focused learning. My interpretation is that L2 learners may have undergone non-compositional motor simulations. More specifically, unlike native speakers, they were unable to integrate negation into their motor simulations of sentence stimuli, leading them to activate the zygomatic muscle in response to negated sentences in which 'smiling' occurred. The notion that form-focused learners may undergo rudimentary kinds of simulation is further supported by a study which demonstrated that novel pseudowords without correlates in the L1, learned exclusively via extralinguistic instruction, were immediately embodied in the sensorimotor system, as indexed by an action-compatibility-effect (Günther et al. 2020).

2.5.3 Against a Behaviouristic Approach to Linguistic Knowledge

In *Word and Object*, Quine proposes the following thought experiment (2013). Two men, one of whom is partially colour blind, learn to use the term "red" in entirely different ways. For the colour-blind man, "red" is learned via "elaborate special combinations of supplementary conditions" such as intensity, shape and so forth, and the other learns "red" via "regulation photochemical [effects]". For Quine, therefore, the "uniformity that unites us in communication" overlies "*a chaotic subjective diversity of connections between words and experience*" (my emphasis).

In my study, precisely these tacit aspects of linguistic knowledge were measured. According to Quine's (albeit idiosyncratic form of) behaviourism, the attribution of mental states may only be justified on the basis of "systematic efficiency" rather than reality (cf. Bromberger & Halle 2000). It was, no doubt, prescient to emphasise the manner in which divergent "private mechanisms" give rise to what appears to be uniform knowledge. But contra Quine, I have shown that such tacit aspects of linguistic knowledge are amenable to systematic empirical investigation. In other words, the heterogeneity of internalised knowledge, which he considers a "chaotic subjective diversity", can be indexed by behavioural measures. It would seem they are not so "chaotic" after all. Divergent "private mechanisms", in this case, were parametrised by different practice conditions. My contention is that evidence from my experiment, and others (Williams & Xue 2024, Morgan-Short et al. 2012b), suggests that varying levels of internalisation of knowledge of a linguistic system do, in fact, merit the attribution of differing mental states. In short, I have shown that tacit aspects of linguistic knowledge are capable of being charted. At the same time, research in this area remains in its nascent stages.

The general blueprint of my experiment, namely teaching participants novel determiners in different ways, and gauging embodiment effects, constitutes a fecund domain of research that should be explored extensively. Firstly, regardless of form and meaning-focused styles of learning, whether implicit learning may induce L2 embodiment should be explored (cf. Williams 2005). Secondly, the same logic of my experiment should be applied to other object properties (e.g., colour, see van Zuijlen et al. 2024). Thirdly, while my study investigated the robustness of *mental simulations* with respect to *concrete referents*, future studies should harness Conceptual Metaphor Theory to investigate whether the practice condition modulates the robustness of *motor simulation* with respect to *abstract referents* (e.g., whether meaning-focused learning of determiners pertaining to relative social power leads to a more robust internal representation of POWER = UP).

3 Conclusion

I investigated the hypothesis that meaning-focused practice is more conducive than form-focused practice to L2 embodiment. I found compelling evidence that those who practiced novel determiners pertaining to relative size in a meaning-focused, but not form-focused manner, mentally simulated the novel determiners. Upon further investigation, by dividing filler items into congruent and incongruent fillers, I found preliminary evidence suggesting that those in the form-focused group may have mentally simulated the novel determiners in some attenuated capacity. It is a common platitude that it is not *what* is known, but rather the *way* something is known that really matters. The internal frameworks I bring to the world may radically differ. Nonetheless, from an outside perspective, they are sufficiently similar to facilitate communication (most of the time). In this article, I probed these tacit aspects of linguistic knowledge. Exploration of the tacit aspects of linguistic knowledge constitutes an exciting area of future research. In sum, *with greater internalisation comes greater embodiment.*

References

Ahlberg, Daniela K., Heike Bischoff, Barbara Kaup, Doreen Bryant & Jan V. Strozyk. 2018. Grounded cognition: Comparing Language × Space interactions in first language and second language. *Applied Psycholinguistics* 39(2). 437–459. doi:10.1017/S014271641700042X.

Ahn, Sunyoung & Nan Jiang. 2018. Automatic semantic integration during L2 sentential reading. *Bilingualism: Language and Cognition* 21. 375–383. doi:10.1017/S1366728917000256.

Anwyl-Irvine, Alexander L., Jessica Massonnié, Adam Flitton, Natasha Kirkham & Jo K. Evershed. 2020. Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods* 388–407. doi:10.3758/s13428-019-01237-x.

Azizpour, Sh. & P. Alavinia. 2021. The impact of focus on form and focus on forms instruction on grammar acquisition of the subjunctive by iranian advanced EFL learners. *Teaching English Language* 15(1). 215–249. doi:10.22132/TEL.2021.134368.

Barsalou, Lawrence W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22(4). 577–660. doi:10.1017/S0140525X99002149.

Bergen, Benjamin K. & Nancy Chang. 2013. Embodied construction grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 168–190. Oxford University Press. doi:10.1093/oxfordhb/9780195396683.013.0010.

Birba, A., David Beltrán, Miguel Martorell Caro, Piergiorgio Trevisan, Boris Kogan, Lucas Sedeño, Agustín Ibáñez & Adolfo M. García. 2020. Motor-system dynamics during naturalistic reading of action narratives in first and second language. *NeuroImage* 216. 116820. doi:10.1016/j.neuroimage.2020.116820.

Bromberger, Sylvain & Morris Halle. 2000. The ontology of phonology (revised). In *Phonological Knowledge: Conceptual and Empirical issues*, Oxford University Press. doi:10.1093/oso/9780198241270.003.0002.

Casasanto, Daniel & Angela de Bruin. 2019. Metaphors we learn by: Directed motor action improves word learning. *Cognition* 182. 177–183. doi:10.1016/j.cognition.2018.09.015.

Çelik, B. 2016. Comparing the effectiveness of form-focused and meaning-focused instructions in EFL teaching. *Journal of Education in Black Sea Region* 1. doi:10.31578/jebs.v1i1.4.

Chen, D., J. Su & R. Wang. 2024. Differences in perceptual representations in multilinguals' first, second, and third language. *Frontiers in Human Neuroscience* doi:10.3389/fnhum.2024.1408411.

Chen, D., R. Wang & J. Zhang. 2020. Perceptual representations in L1, L2 and L3 comprehension: Delayed sentence–picture verification. *Journal of Psycholinguistic Research* 49. 41–57. doi:10.1007/s10936-019-09670-x.

Chen, M. & W. Li. 2022. The influence of form-focused instruction on the L2 Chinese oral production of korean native speakers. *Frontiers in Psychology* 13. 790424. doi:10.3389/fpsyg.2022.790424.

Chen, W., Xiuyan Guo, Jinghua Tang, Lei Zhu, Zhiliang Yang & Zoltan Dienes. 2011. Unconscious structural knowledge of form–meaning connections. *Consciousness and Cognition* 20(4). 1751–1760. doi:10.1016/j.concog.2011.03.003.

Connell, Louise. 2007. Representing object colour in language comprehension. *Cognition* 102(3). 476–485. doi:10.1016/j.cognition.2006.02.009.

Craik, Fergus I. M. 1973. A "Levels of Analysis" View of Memory. In *Communication and Affect: Language and Thought*, 46–65. New York: Academic Press.

De Grauwe, S., Roel M. Willems, Shirley-Ann Rueschemeyer, Kristin Lemhöfer & Herbert Schriefers. 2014. Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia* 56.

334–349. doi:10.1016/j.neuropsychologia.2014.02.003.

Dehqan, Mahmood, Danial Azizi & Fatemeh Miri. 2022. Meaning-focused output and meaning-focused input instruction and willingness to communicate: Effects and perceptions. *Mextesol Journal* 46. doi:10.61871/mj.v46n3-3.

Dragomir, Isabela-Anda & Brandusa-Oana Niculescu. 2022. Shallow and deep processing – An integrated cognitive architecture for foreign language learning. *Land Forces Academy Review* 27. 216–220. doi:10.2478/raft-2022-0028.

Dudschig, Carolin, Iria de la Vega & Barbara Kaup. 2014. Embodiment and second-language: Automatic activation of motor responses during processing spatially associated L2 words and emotion L2 words in a vertical stroop paradigm. *Brain and Language* 132. 14–21. doi:10.1016/j.bandl.2014.02.002.

Ellis, Rod. 2001. Introduction: Investigating form-focused instruction. *Language Learning* 51. 1–46. doi:10.1111/j.1467-1770.2001.tb00013.x.

Eriksson, A. 2023. Form-focused or meaning-focused? grammar tasks in EFL textbooks for English 5 in upper secondary school. Publication details incomplete in source list.

Fodor, Jerry. 1975. *The Language of Thought*. Harvard University Press.

Foroni, Francesco. 2015. Do we embody second language? evidence for "partial" simulation during processing of a second language. *Brain and Cognition* 99. 8–16. doi:10.1016/j.bandc.2015.06.006.

Glenberg, Arthur M., Bennett Jaworski, Michael Rischal & Juliet Levin. 2007. What brains are for: Action, meaning, and reading comprehension. In *Reading Comprehension Strategies: Theories, Interventions, and Technologies*, 221–240. Lawrence Erlbaum Associates Publishers.

Glenberg, Arthur M. & Michael P. Kaschak. 2002. Grounding language in action. *Psychonomic Bulletin & Review* 9(3). 558–565. doi:10.3758/BF03196313.

Gómez, Ligia & Arthur Glenberg. 2022. Embodied classroom activities for vocabulary acquisition. In *Movement Matters: How Embodied Cognition Informs Teaching and Learning*, 77–90. The MIT Press. doi:10.7551/mitpress/13593.003.0011.

van Grunsven, J. 2025. Disabled body-minds in hostile environments: Disrupting an ableist cartesian sociotechnical imagination with enactive embodied cognition and critical disability studies. *Topoi* 44. 505–515. doi:10.1007/s11245-024-10080-5.

Günther, Friederike, T. Nguyen, L. Chen, Carolin Dudschig, Barbara Kaup & Arthur M. Glenberg. 2020. Immediate sensorimotor grounding of novel concepts learned from language alone. *Journal of Memory and Language* 115. 104172. doi:10.1016/j.jml.2020.104172.

Hauk, Olaf, Ingrid Johnsrude & Friedemann Pulvermüller. 2004. Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41(2). 301–307. doi:10.1016/S0896-6273(03)00838-9.

Hughes-Berheim, Sarah S., Spyridoula Cheimariou, John F. Shelley-Tremblay, Mary M. Doheny & Laura M. Morett. 2022. Extending gesture's impact on word learning to reading: A self-paced reading study. *Discourse Processes* 59(8). 646–667. doi:10.1080/0163853X.2022.2132080.

Kaan, Edith. 2014. Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism* 4(2). doi:10.1075/lab.4.2.05kaa.

Kogan, Boris, Ernesto García-Marco, Agustina Birba, C. Cortés, Lucio Melloni, Agustín Ibáñez et al. 2020. How words ripple through bilingual hands: Motor–language coupling during L1 and L2 writing. *Neuropsychologia* 146. 2–9. doi:https://doi.org/10.1016/j.neuropsychologia.2020.107563.

de Koning, Björn B., Sanne I. Wassenburg, Linda T. Bos & Menno Van der Schoot. 2016. Size does matter: Implied object size is mentally simulated during language comprehension. *Discourse Processes* 54(7). 493–503. doi:10.1080/0163853X.2015.1119604.

Koster, Diarra, Teresa Cadierno & Marco Chiarandini. 2019. Mental simulation of object orientation and size: A conceptual replication with second language learners. *Journal of the European Second Language Association* 2(1). 38–48. doi:10.22599/jesla.39.

Leow, Ronald & Johnathan Mercer. 2015. Depth of processing in L2 learning: Theory, research, and pedagogy. *Journal of Spanish Language Teaching* 2. 1–14. doi:10.1080/23247797.2015.1026644.

Liu, D., L. Wang & Y. Han. 2024. Mental simulation of colour properties during language comprehension: Influence of context and comprehension stages. *Cognitive Processing* 25(4). 587–600. doi:10.1007/s10339-024-01201-4.

Mahon, Bradford Z. & Alfonso Caramazza. 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris* 102(1–3). 59–70. doi:10.1016/j.jphysparis.

Monaco, E. M., J. Mouthon, J. Britz, M. Sato, I. Stefanos-Yakoub, J.-M. Annoni & L. B. Jost. 2023. Embodiment of action-related language in the native and a late foreign language – An fMRI-study. *Brain and Language* 244. doi:10.1016/j.bandl.2023.105312.

Monaco, Elena, Laura B. Jost, Pascal M. Gygax & Jean-Marie Annoni. 2019. Embodied semantics in a second language: Critical review and clinical implications. *Frontiers in Human Neuroscience* 13. 110. doi:10.3389/fnhum.2019.00110.

Morgan-Short, Kara, Ingrid Finger, Sarah Grey & Michael T. Ullman. 2012a. Second language processing shows increased native-like neural responses after months of no exposure. *PLoS One* 7(3). doi:10.1371/journal.pone.0032974.

Morgan-Short, Kara, Karsten Steinhauer, Cristina Sanz & Michael T. Ullman. 2012b. Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience* 24(4). 933–947. doi:10.1162/jocn_a_00119.

Nation, I. S. P. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, Paul. 1996. The four strands of a language course. *TESOL in Context* 6(1). 7–12. https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/paul-nations-publications/publications/documents/1996-Four-strands.pdf.

Norman, T. & O. Peleg. 2022. The reduced embodiment of a second language. *Bilingualism: Language and Cognition* 25(3). 406–416.

doi:10.1017/S1366728921001115.

Noroozi, M. 2023. The effect of blended learning through meaning-focused input and output activities on learning collocations. *Mextesol Journal* 46. 1–13. doi:10.61871/mj.v46n4-16.

OpenAI. 2025. Chatgpt [large language model]. https://chat.openai.com/chat.

Qian, W. 2016. Embodied cognition processing and representation of power words by second language learners with different proficiency levels. *Chinese Journal of Applied Linguistics* 39. doi:10.1515/cjal-2016-0030.

Quine, W. V. 2013. *Word and object*. MIT Press new edition edn.

R Core Team. 2024. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Richardson, J. T. E. 2011. Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review* 6(2). 135–147. doi:10.1016/j.edurev.2010.12.001.

Saito, Kazuya. 2015. The role of age of acquisition in late second language oral proficiency attainment. *Studies in Second Language Acquisition* 37(4). 713–743. doi:10.1017/S0272263115000248.

Schmidt, Richard. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11. 129–158. doi:https://doi.org/10.1093/applin/11.2.129.

Segal, D. & G. Kavé. 2024. Exposure duration is the main determinant of bilinguals' vocabulary knowledge. *Applied Psycholinguistics* 45(6). 1000–1015. doi:10.1017/S0142716424000389.

Stanfield, Robert A. & Rolf A. Zwaan. 2001. The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science* 12(2). 153–156. doi:10.1111/1467-9280.00326.

Sullivan, Gail M. & Richard Feinn. 2012. Using effect size—or why the *p* value is not enough. *Journal of Graduate Medical Education* 4(3). 279–282. doi:10.4300/JGME-D-12-00156.1.

Szabó, Z. 2024. Compositionality. The Stanford Encyclopedia of Philosophy. Edward N. Zalta and Uri Nodelman (eds.). https://plato.stanford.edu/archives/fall2024/entries/compositionality/.

Thierry, Guillaume, Rasha Abdel-Rahman & Panos Athanasopoulos. 2024. An introduction to the cognitive neuroscience of language embodiment and relativity. *Language Learning* 74(51). 5–19. doi:10.1111/lang.12643.

Vanek, N., A. Matić Škorić, S. Košutar, Matějka & K. Stone. 2024. Looks at what isn't there: Eye movements on a blank screen when processing negation in a first and a second language. *Frontiers in Human Neuroscience* 18. 1457038. doi:10.3389/fnhum.2024.1457038.

Vigliocco, Gabriella, Linda Convertino, Serena De Felice, L. Gregorians, V. Kewenig, M. A. E. Mueller, S. Veselic, Mirco Musolesi, Andrew Hudson-Smith, Nick Tyler, Eirini Flouri & Hugo J. Spiers. 2024. Ecological brain: Reframing the study of human behaviour and cognition. *Royal Society Open Science* 11(11). 240762. doi:10.1098/rsos.240762.

Vukovic, N. & Yury Shtyrov. 2014. Cortical motor systems are involved in second-language comprehension: Evidence from rapid mu-rhythm desynchronisation. *NeuroImage* 102. 695–703. doi:10.1016/j.neuroimage.2014.08.039.

Vukovic, N. & John Williams. 2014. Automatic perceptual simulation of first language meanings during second language sentence processing in bilinguals. *Acta Psychologica* 145. 98–103. doi:10.1016/j.actpsy.2013.11.002.

Wang, M. & H. Zhao. 2024. Perceptual representations in L1 and L2 spatial and abstract language processing: Applying an innovative sentence-diagram verification paradigm. *Frontiers in Human Neuroscience* 18. 1425576. doi:10.3389/fnhum.2024.1425576.

Williams, John. 2005. Learning without awareness. *Studies in Second Language Acquisition* 27(2). 269–304. doi:10.1017/S0272263105050138.

Williams, John & Y. Xue. 2024. Learning without awareness revisited and reconsidered: A conceptual replication and extension. *Studies in Second Language Acquisition* 46(4). 1231–1257. doi:10.1017/S0272263124000500.

Xue, Y. 2024. Linguistic relativity: Loci and conditions of emergence. doi:10.17863/CAM.114899. Apollo - University of Cambridge Repository.

Yaxley, Richard H. & Rolf A. Zwaan. 2007. Simulating visibility during language comprehension. *Cognition* 105(1). 229–236. doi:10.1016/j.cognition.2006.09.003.

Zhao, C., J. Kormos, P. Rebuschat & S. Suzuki. 2021. The role of modality and awareness in language learning. *Applied Psycholinguistics* 42(3). 703–737. doi:10.1017/S0142716421000023.

Zhao, H., N. Vanek, J. Yang & M. Wang. 2025. Editorial: Role of perceptual and motor representations in bilingual and second language processing. *Frontiers in Human Neuroscience* 19. 1555254. doi:10.3389/fnhum.2025.1555254.

van Zuijlen, S., S. Singh & K. Gunawan. 2024. Automatic mental simulation in native and non-native speakers. *Memory & Cognition* 52. 1152–1163. doi:10.3758/s13421-024-01533-8.

Zwaan, Rolf A., Robert A. Stanfield & Richard H. Yaxley. 2002. Language comprehenders mentally represent the shape of objects. *Psychological Science* 13(2). 168–171. doi:10.1111/1467-9280.00430.

Abbreviations

ACE	Action-compatibility effect	AoA	Age of acquisition
ANOVA	Analysis of variance	CI	Confidence interval
DoP	Depth of processing	EC	Embodiment-Compositionality Hypothesis
FFG	Form-focused group	fMRI	Functional magnetic resonance imaging
GJT	Grammaticality judgement task	L1/2	First/second language
LTM	Long-term memory	MFG	Meaning-focused group
RT	Reaction time	SD	Standard deviation
SPVT	Sentence-picture-verification task		

George Blumenthal
Department of Theoretical and Applied Linguistics